

FOURTH EDITION

THE HANDBOOK OF
FORENSIC
PSYCHOLOGY

EDITED BY
IRVING B. WEINER
RANDY K. OTTO

WILEY

The Handbook of Forensic Psychology

The Handbook of Forensic Psychology

Fourth Edition

Edited by
Irving B. Weiner
Randy K. Otto

WILEY

Cover design: Wiley

This book is printed on acid-free paper.

Copyright © 2013 by John Wiley & Sons, Inc. All rights reserved

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at www.wiley.com/go/permissions.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with the respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor the author shall be liable for damages arising herefrom.

For general information about our other products and services, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley publishes in a variety of print and electronic formats and by print-on-demand. Some material included with standard print versions of this book may not be included in e-books or in print-on-demand. If this book refers to media such as a CD or DVD that is not included in the version you purchased, you may download this material at <http://booksupport.wiley.com>. For more information about Wiley products, visit www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Weiner, Irving B.

The handbook of forensic psychology / Irving B. Weiner, Randy K. Otto. – Fourth edition.

1 online resource.

Includes bibliographical references and index.

ISBN 978-1-118-34841-3 (cloth : alk. paper)

ISBN 978-1-118-73471-1 (ebk.)

ISBN 978-1-118-73483-4 (ebk.)

1. Forensic psychology—Handbooks, manuals, etc. I. Otto, Randy K. II. Title.

RA1148

614'.15—dc23

2013018068

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Contents

Preface	ix
Contributors	xiii
PART ONE CONTEXT OF FORENSIC PSYCHOLOGY	
1 History of Forensic Psychology <i>Curt R. Bartol and Anne M. Bartol</i>	3
2 Defining Forensic Psychology <i>Randy K. Otto and James R. P. Ogloff</i>	35
3 Accessing the Law and Legal Literature <i>David DeMatteo, Michael E. Kessler, and Heidi Strohmaier</i>	57
4 Practicing Ethical Forensic Psychology <i>Irving B. Weiner and Allen K. Hess</i>	85
5 Training in Forensic Psychology <i>Daniel A. Krauss and Bruce D. Sales</i>	111
PART TWO APPLYING PSYCHOLOGY TO CIVIL PROCEEDINGS	
6 Conducting Child Custody and Parenting Evaluations <i>Philip M. Stahl</i>	137
7 Conducting Personal Injury Evaluations <i>Lisa Drago Piechowski</i>	171
8 Identifying and Treating Educational Disabilities <i>Daniel J. Reschly</i>	197
9 Assessing Civil Capacities <i>Michele Galietta, Alexandra Garcia-Mansilla, and Barbara Stanley</i>	219
10 Conducting Child Abuse and Neglect Evaluations <i>Lois O. Condie</i>	237

PART THREE APPLYING PSYCHOLOGY TO CRIMINAL PROCEEDINGS

- 11 Assessing Competency to Stand Trial 281
Patricia A. Zapf, Ronald Roesch, and Gianni Pirelli
- 12 Assessing Criminal Responsibility 315
Patricia A. Zapf, Stephen L. Golding, Ronald Roesch, and Gianni Pirelli
- 13 Specific Intent and Diminished Capacity 353
Charles R. Clark

PART FOUR SPECIAL APPLICATIONS

- 14 Assessing Violence Risk 385
Kevin S. Douglas, Stephen D. Hart, Jennifer L. Groscup, and Thomas R. Litwack
- 15 Psychology and Law Enforcement 443
Ellen M. Scrivner, David M. Corey, and Lorraine W. Greene
- 16 Evaluating and Assisting Jury Competence in Civil Cases 469
Jennifer K. Robbennolt, Jennifer L. Groscup, and Steven Penrod
- 17 Evaluating Eyewitness Testimony of Adults 513
Stephen J. Ross, Colin G. Tredoux, and Roy S. Malpass
- 18 Evaluating Eyewitness Testimony of Children 561
Sue D. Hobbs, Jonni L. Johnson, Gail S. Goodman, Daniel Bederian-Gardner, Michael J. Lawler, Ivan D. Vargas, and Macaria Mendoza
- 19 Employing Polygraph Assessment 613
William G. Iacono and Christopher J. Patrick
- 20 Applying Hypnosis in Forensic Contexts 659
Alan W. Scheflin

PART FIVE COMMUNICATING EXPERT OPINIONS

- 21 Writing Forensic Reports 711
Irving B. Weiner
- 22 Testifying in Court 733
Randy K. Otto, Susan L. Kay, and Allen K. Hess

PART SIX INTERVENING WITH OFFENDERS

- 23 Practicing Psychology in Correctional Settings 759
Paul Gendreau and Claire Goggin

24	Treating Criminal Offenders	795
	<i>Robert D. Morgan, Daryl G. Kroner, Jeremy F. Mills, and Ashley B. Batastini</i>	
25	Assessing and Treating Sex Offenders	839
	<i>W. L. Marshall, Douglas Boer, and Liam E. Marshall</i>	
	Appendix: Specialty Guidelines for Forensic Psychology	867
	Author Index	881
	Subject Index	905

Preface

THE potential for psychologists to assist the legal system has been recognized since the early 20th century, but only within the past 50 years has psychology begun to realize this potential in meaningful ways. This progress has included newly developed professional organizations, such as the American Psychology-Law Society and the International Association for Correctional and Forensic Psychology; graduate, internship, and fellowship programs in the specialty area (listed in www.ap-ls.org/education/GraduatePrograms.php); organizations devoted to certifying qualified practitioners, such as the American Board of Forensic Psychology and the American Board of Police and Public Safety Psychology; such scientific journals as *Law and Human Behavior*, *Behavioral Sciences and the Law*, and *Criminal Justice and Behavior*; and books devoted to the interface of psychology and law. This specialty area has continued to grow rapidly since the previous edition of the *Handbook of Forensic Psychology* was published in 2006, with increasing numbers of psychologists becoming involved in forensic practice and research and a steady flow of new ideas and information becoming available.

This fourth edition of the *Handbook of Forensic Psychology*, like its predecessors, aims to provide an authoritative and comprehensive resource for understanding the theoretical foundations of forensic psychology, becoming familiar with the expanding research base in this specialty, and learning to apply forensic concepts artfully in everyday practice. To this end, the contributors to this volume, as in the prior three editions, are accomplished scholars and practitioners in their respective areas. Some are prominent academicians who conduct research and offer consultation. Others are actively engaged service providers who also make significant contributions to the literature. Several have degrees in law as well as psychology. These authors were asked to delineate the enduring issues in an area of their specialty and frame these issues in the light of contemporary research and prevailing conceptual formations.

Although similar in focus and structure to previous editions, the present volume has been substantially rewritten and updated to enhance its value. The content and sequence of the chapters have been reframed to increase their relevance to the practice of forensic psychology and encompass both recent research findings and developments in statutory and case law. As testimony to the fresh perspectives in this fourth edition, the Table of Contents identifies 48 authors and co-authors, of whom 24 are new contributors to the *Handbook*.

The present volume comprises six parts. Part One concerns the context of forensic psychology and begins with chapters on the history of forensic psychology and on defining the nature of forensic psychology. Chapter 3 then provides information about and guidelines for accessing the legal literature. Chapter 4 alerts forensic psychologists to ethical and legal considerations that should guide their work, with specific attention to the American Psychological Association ethics code and the Specialty Guidelines for Forensic Psychology. Chapter 5 describes training models and resources in forensic psychology for faculty developing programs of instruction and for students and general practitioners seeking specialized education or supervised experience in forensic psychology.

Part Two comprises five chapters concerning applications of psychology in civil proceedings. Chapter 6 addresses family law procedures and issues related to conducting evaluations of children and their parents involved in disputed custody. Chapter 7 discusses personal injury litigation, with particular attention to considerations in psychological assessment. Chapter 8 reviews the impact of recent congressional legislation on identifying and treating educational disabilities. Chapter 9 examines issues related to the assessment of persons' competence to execute a variety of legal rights in civil contexts. Chapter 10 concludes this section with guidelines for conducting evaluations in cases of alleged child abuse or neglect.

Part Three deals with applying psychology in criminal proceedings and covers three critical considerations of concern to triers-of-fact. Chapter 11 provides guidelines for assessing competence to stand trial. Chapter 12 traces the development and current applications of the concepts of criminal responsibility and legal insanity. Chapter 13 delineates the related nuances of criminal intent and diminished capacity.

Part Four presents information on seven special applications of forensic psychology. Chapter 14 leads off with a discussion of violence risk research and assessment, and Chapter 15 follows with an overview of emerging roles for psychologists in law enforcement. Chapter 16 reviews considerations related to evaluating jury decision making and promoting juror competence. Chapters 17 and 18 review developments related to the evaluation of testimony given by adults and children. Chapters 19 and 20 then provide accounts of the development of lie detection and hypnosis and describe current and emerging trends in forensic uses of these procedures.

Part Five of the *Handbook* looks at effective communication of expert opinion in forensic cases. Chapter 21 focuses on the essentials of writing appropriate and useful reports, and Chapter 22 discusses the admissibility of expert testimony and key considerations in communicating one's work and opinions to legal decision makers.

Part Six concludes the *Handbook* with consideration of some important aspects of providing services to offenders. Chapter 23 discusses principles of effective correctional rehabilitation in both prison settings and the community, and Chapter 24 provides a behind-the-bars guide to conducting psychotherapy with offenders. Chapter 25 continues this theme with specific attention to diagnostic and treatment

procedures useful in working with sexual offenders. Finally, the Appendix provides readers with the full text of the Specialty Guidelines for Forensic Psychology.

We would like to thank our authors for their valuable contributions to this volume, both those who revised chapters that appeared in previous editions and those who joined us for the first time in this edition. We also want to acknowledge with appreciation the guidance and support of the John Wiley & Sons editorial staff, particularly Patricia Rossi, Executive Editor, and Kara Borbely, Editorial Program Coordinator. Finally, we recognize and honor the contributions of our colleague Allen Hess, who served as co-editor of the first three editions before his untimely death. Readers familiar with previous editions of the *Handbook* will recognize his fingerprints on this edition as well.

IRVING B. WEINER
RANDY K. OTTO

Contributors

Anne M. Bartol, PhD

Bartol Consulting and Writing
Glenville, New York

Curt R. Bartol, PhD

Bartol Consulting and Writing
Glenville, New York

Ashley B. Batastini, MS

Department of Psychology
Texas Tech University
Lubbock, Texas

Daniel Bederian-Gardner, MA

Department of Psychology
University of California
Davis, California

Douglas Boer, PhD

Department of Psychology
University of Waikato
Hamilton, New Zealand

Charles R. Clark, PhD, ABPP

Independent Practice
Ann Arbor, Michigan

Lois O. Condie, PhD, ABPP

Department of Neurology
Harvard Medical School
Boston, Massachusetts

David M. Corey, PhD, ABPP

Independent Practice
Lake Oswego, Oregon

David DeMatteo, JD, PhD

Department of Psychology
Earle Mack School of Law
Drexel University
Philadelphia, Pennsylvania

Kevin S. Douglas, LLB, PhD

Department of Psychology
Simon Fraser University
Burnaby, British Columbia, Canada

Michele Galietta, PhD

John Jay College of Criminal Justice
City University of New York
New York, New York

Alexandra R. Garcia-Mansilla, PhD

Queens TAC Mental Health Jail
Diversion Program
New York, New York

Paul Gendreau, PhD

Department of Psychology
University of New Brunswick
Saint John, New Brunswick,
Canada

Claire Goggin, PhD

Department of Criminology and
Criminal Justice
St. Thomas University
Fredericton, New Brunswick,
Canada

Stephen L. Golding, PhD, ABPP

Department of Psychology
University of Utah
Salt Lake City, Utah

Gail S. Goodman, PhD

Department of Psychology
University of California
Davis, California

Lorraine W. Greene, PhD, ABPP

University Counseling Center
Tennessee State University
Nashville, Tennessee

Jennifer L. Groscup, JD, PhD

Department of Psychology
Scripps College
Claremont, California

Stephen D. Hart, PhD

Department of Psychology
Simon Fraser University
Burnaby, British Columbia, Canada

Allen K. Hess, PhD

Department of Psychology
Auburn University at Montgomery
Montgomery, Alabama

Sue D. Hobbs, MA

Department of Psychology
University of California
Davis, California

William G. Iacono, PhD

Department of Psychology
University of Minnesota
Minneapolis, Minnesota

Jonni L. Johnson, MA

Department of Psychology
University of California
Davis, California

Susan L. Kay, JD

Vanderbilt University Law School
Nashville, Tennessee

Michael E. Keesler, JD, MS

Department of Psychology
Drexel University
Philadelphia, Pennsylvania

Daniel A Krauss, JD, PhD, ABPP

Department of Psychology
Claremont McKenna College
Claremont, California

Daryl G. Kroner, PhD

Department of Criminology &
Criminal Justice
Southern Illinois University
Carbondale, Illinois

Michael J. Lawler, PhD

School of Health Sciences
University of South Dakota
Vermillion, South Dakota

Thomas R. Litwack, PhD, JD

John Jay College of Criminal Justice
City University of New York
New York, New York

Roy S. Malpass, PhD

Department of Psychology
University of Texas at El Paso
El Paso, Texas

Liam E. Marshall, PhD

Rockwood Psychological Services
Kingston, Ontario, Canada

W. L. Marshall, PhD

Rockwood Psychological Services
Kingston, Ontario, Canada

Macaria Mendoza, MA
 Center for Public Policy Research
 University of California
 Davis, California

Jeremy F. Mills, PhD
 Department of Psychology
 Carleton University
 Ottawa, Ontario, Canada

Robert D. Morgan, PhD
 Department of Psychology
 Texas Tech University
 Lubbock, Texas

James R. P. Ogloff, JD, PhD
 School of Psychology, Psychiatry,
 and Psychological Medicine
 Monash University and Institute of
 Forensic Mental Health
 Victoria, Australia

Randy K. Otto, PhD, ABPP
 Department of Mental Health Law
 and Policy
 University of South Florida
 Tampa, Florida

Christopher J. Patrick, PhD
 Department of Psychology
 Florida State University
 Tallahassee, Florida

Steven Penrod, JD, PhD
 John Jay College of Criminal Justice
 City University of New York
 New York, New York

Lisa Drago Piechowski, PhD, ABPP
 Department of Clinical Psychology
 American School of Professional
 Psychology
 Washington, DC

Gianni Pirelli, PhD
 Department of Psychology
 Greystone Park Psychiatric Hospital
 Morris Plains, New Jersey

Daniel J. Reschly, PhD
 Peabody College of Education and
 Human Development
 Vanderbilt University
 Nashville, Tennessee

Jennifer K. Robbennolt, JD, PhD
 College of Law
 University of Illinois
 Champaign, Illinois

Ronald Roesch, PhD
 Department of Psychology
 Simon Fraser University
 Burnaby, British Columbia, Canada

Stephen J. Ross, PhD
 Interdisciplinary Arts and Sciences
 University of Washington, Tacoma
 Tacoma, Washington

Bruce D. Sales, PhD, JD
 Department of Criminal Justice
 University of Indiana
 Bloomington, Indiana

Alan W. Schefflin, LLM
 Santa Clara University School of Law
 Santa Clara, California

Ellen M. Scrivner, PhD
 Executive Fellow
 Police Foundation
 Washington, DC

Philip M. Stahl, PhD, ABPP
 Independent Practice
 Queen Creek, Arizona

Barbara Stanley, PhD

Department of Neuroscience
New York State Psychiatric
Institute
New York, New York

Heidi Strohmaier

Department of Psychology
Drexel University
Philadelphia, Pennsylvania

Colin G. Tredoux, PhD

Department of Psychology
University of Cape Town
Cape Town, South Africa

Ivan D. Vargas, JD

Center for Public Policy Research
University of California
Davis, California

Irving B. Weiner, PhD, ABPP

Department of Psychiatry and
Neurosciences
University of South Florida
Tampa, Florida

Patricia A. Zapf, PhD

John Jay College of Criminal Justice
City University of New York
New York, New York

PART ONE

CONTEXT OF FORENSIC
PSYCHOLOGY

CHAPTER 1

History of Forensic Psychology

CURT R. BARTOL AND ANNE M. BARTOL

IN the course of writing this chapter over four editions of this *Handbook*, we have learned a few lessons. In the first edition, we asserted that psychologists do not care about the history of their profession but are instead drawn to contemporary issues and theories. We learned that this was a simplistic generalization, so in subsequent editions we acknowledged that our initial statement had been rash. Psychologists (perhaps most of them) *do* care about history, as is apparent from numerous articles published in professional journals reviewing historical trends, the continuing publication of a journal devoted to the history of psychology, and special interest divisions of professional organizations, such as Division 26, Society for the History of Psychology, of the American Psychological Association (APA). We have also learned that there is some danger in proclaiming an event or a person a historic “first” or a “father,” because these proclamations may be challenged, usually with kindness but not always with good humor.

Psychology, like other disciplines, needs historical insights. It needs to understand whence it came in order to assess where it is going. A perusal of journals and books published at the turn of the 20th century, for example, may spark interest in a concept long forgotten or a predecessor whose theories and research deserve to be revisited. Yet delving into early works reminds us of false starts and the occasional damage they did, such as the work of Henry H. Goddard (1914) on feeble-mindedness during the early 1900s and the self-promotion of Hugo Münsterberg. However, we have also learned that hindsight is imperfect; people are sometimes overlooked, and the historical discoveries may be incomplete. We thus approach this chapter once again with humility. To paraphrase the phrase that “journalism is the first rough draft of history,” we say here that this chapter is our fourth rough draft of the history of forensic psychology, with emphasis on its American origins.

In these early years of the 21st century, forensic psychology remains a young branch of applied psychology, having been recognized by the APA as a specialty in 2001 and recertified in 2008. Even before that, in 1991, Specialty Guidelines for

Forensic Psychologists (Committee on Ethical Guidelines for Forensic Psychologists [hereafter Committee], 1991) were adopted by the American Psychology–Law Society, which is Division 41 of the APA. These Guidelines were recently revised, renamed Specialty Guidelines for Forensic Psychology (APA, 2013), and accepted by the APA Council of Representatives. (The Specialty Guidelines are reprinted as the appendix to this volume with permission of the APA.) Interestingly, although forensic psychology was initially viewed as primarily clinical in nature—such as by providing assessments to the courts—its scope has broadened to encompass the practice of psychology as it provides expertise to the law in a very wide range of contexts (see APA, 2013; Committee, 1991).

This broad view of forensic psychology was not always supported. According to Ronald Roesch, for example (cited in Brigham, 1999, p. 279), “Most psychologists define the area more narrowly to refer to clinical psychologists who are engaged in clinical practice within the legal system.” A few years later, Brigham and Grisso (2003) modified this somewhat, noting “*Many* psychologists define forensic psychology more narrowly to refer to clinical psychologists who are engaged in clinical practice within the legal system. The distinction here is between psychologists who bring scientific information to the courts for their consideration in cases and psychologists who evaluate individuals and testify about them in reference to a legal question” (p. 392, emphasis added). In recognizing forensic psychology as a specialty in 2001, the APA itself adopted the narrow approach, to include “the primarily clinical aspects of forensic assessment, treatment, and consultation” (Otto & Heilbrun, 2002, p. 8). However, as noted, the Specialty Guidelines take a broader view.

In this chapter, forensic psychology is being viewed broadly. It is both (1) the *research endeavor* that examines aspects of human behavior directly related to the legal process (e.g., eyewitness memory and testimony, jury decision making, and criminal behavior) and (2) the *professional practice* of psychology within or in consultation with a legal system that encompasses both criminal and civil law and the numerous areas where they intersect. Therefore, the term *forensic psychology* refers broadly to the *production* of psychological knowledge and its *application* to the civil and criminal justice systems. It includes activities as varied as these: courtroom testimony, child custody evaluations, law enforcement candidate screening, treatment of offenders in correctional facilities, assessment of plaintiffs with disability claims, research and theory building in the area of criminal behavior, and the design and implementation of intervention and prevention programs for youthful offenders. A review of the table of contents of this *Handbook* indicates a similarly broad focus.

In the pages to follow, after an introductory section covering seminal contributions, we review developments in four major areas of forensic psychology: legal psychology, correctional psychology, police psychology, and criminal psychology. Readers will undoubtedly recognize that there is considerable overlap in these categories and in the subheadings. Correctional psychology, for example, presupposes some understanding of criminal psychology. Assessment, which we cover under legal psychology, is an essential tool of the trade for psychologists, and it

underlies all practice. Nonetheless, for purposes of identifying historical trends and landmarks, discussion of these four distinctive areas is warranted.

We focus on forensic psychology rather than forensic psychiatry, which has its own well-documented and rich history, probably centered on the early work of Isaac Ray, who is considered by some the father of forensic psychiatry (Brigham & Grisso, 2003). We also do not delve into the origins of the sociology of law, referred to as sociological jurisprudence, or the legal realism movement within the law itself. This movement, born during the first third of the 20th century, advocated a partnership between the law and the social sciences (Ogloff, Tomkins, & Bersoff, 1996).

In addition, we emphasize the work of forensic psychologists in the United States and, to a lesser extent, Canada,¹ although we give due recognition to the work of European psychologists, who dominated the field prior to World War I. We review the achievements of psychologists from the end of the 19th century and extend our discussion into the 1970s, when forensic psychology came of age (Loh, 1981). The reader interested in more detail about the issues and individuals discussed might check landmark summaries of psychology and law published by Whipple (1909, 1910, 1911, 1912, 1913, 1914, 1915, 1917), Hutchins and Slesinger (1929), Louisell (1955, 1957), Tapp (1976), Loh (1981), and Monahan and Loftus (1982). More recently, Brigham and Grisso (2003) and Mülberger (2009) have published historical pieces on this topic, the latter with a strong emphasis on German influences. On the whole, however, developments from the 1980s forward are addressed in the works of other contributors to this *Handbook*.

LEGAL PSYCHOLOGY

Legal psychology refers to psychological theory, research, and practice directly pertinent to the law and legal issues. It focuses on psycholegal research and contacts with judges, lawyers, and other law-related professionals in a wide range of contexts. The origins of legal psychology can be traced to the work of experimental psychologists in Europe in the 19th century, particularly in relation to the psychology of testimony (Mülberger, 2009; Sporer, 1982, 2008) and most particularly to the testimony of children, whose memory of events was considered unreliable (Lipmann, 1911). We discuss this work shortly.

U.S. ORIGINS

Do chestnut or oak trees lose their leaves earlier in autumn? Do horses in the field stand with head or tail to the wind? In which direction do the seeds of an apple point? What was the weather one week ago today?

When J. McKen Cattell posed these questions to 56 college students at Columbia University in March 1893, he was probably conducting one of the first American

1. We are grateful to Dr. Craig Bennell, Department of Psychology, Carleton University, for recommending additional readings on the history of forensic psychology in Canada.

studies, albeit an informal one, on the psychology of testimony. The questions he asked his students were similar to those that “might naturally be asked in a court of justice” (Cattell, 1895, p. 761). His subjects were allowed 30 seconds to consider their answers, then told to write their responses and indicate their degree of confidence in each answer.

When Cattell conducted his informal and preliminary study, it was reasonably well established that eyewitness accounts of events were unreliable and incomplete. As we will see shortly, both French and German psychologists were familiar with the powerful influence of suggestion over sensation and perception, having conducted substantial research in these areas. The specific conditions under which testimony was inaccurate were not known, however. Cattell (1895) noted: “An unscrupulous attorney can discredit the statements of a truthful witness by cunningly selected questions. The jury, or at least the judge, should know how far errors in recollection are normal and how they vary under different conditions” (p. 761). But Cattell himself was surprised at both the degree of inaccuracy he uncovered and the wide range of individual differences in the levels of confidence expressed by the students. Answers to the weather question, for example, were “equally distributed over all kinds of weather which are possible at the beginning of March” (p. 761). Some students were nearly always sure they were correct, even when they were not, while others were consistently uncertain and hesitant in their answers, even when they were correct.

Cattell’s study probably was the genesis of modern forensic psychology in the United States, because it sparked the interest of other researchers in the psychology of testimony, which remains to this day a dominant research interest among legal psychologists. Joseph Jastrow immediately replicated Cattell’s “experiment” at the University of Wisconsin and obtained similar results (Bolton, 1896). Aside from this brief flirtation, however, American psychologists did not immediately embrace the study of legal issues. Psychologists in Europe seemed more intrigued—they had long been interested in the psychological concepts involved. First, Alfred Binet (1900) replicated Cattell’s project in France. In addition, he summarized relevant experiments on the psychology of testimony that were being conducted in Europe, and he eventually called for a “science psycho-judiciaire” (Binet, 1905; Binet & Clarparedé, 1906).

EUROPEAN ORIGINS

Most significant for the historical development of forensic psychology was the apparent fascination Cattell’s experiment and Binet’s work held for (Louis) William Stern (1902, 1910, 1939), who had received his doctorate in psychology at the University of Berlin under the tutelage of Hermann Ebbinghaus. In 1901, Stern collaborated with the criminologist F. v. Liszt in an attempt to lend realism to the Cattell design. Stern and Liszt conducted a “reality experiment” in a law class, staging a bogus quarrel between two students over a scientific controversy.

As Stern later recounted it, the argument accelerated until one student drew a revolver (Stern, 1939). At this point, the professor intervened and asked for written and oral reports from the class about aspects of the dispute. Although the witnesses were law students who, Stern asserted, should have known the pitfalls of testifying, none could give a faultless report. The number of errors per individual ranged from 4 to 12. Moreover, the researchers found that inaccuracies increased with respect to the second half of the scenario, when excitement and tension were at their peak. They concluded—tentatively—that “affective reactions inhibit exact observation and reliable remembrance” (Stern, 1939, p. 11).

By his own account, Stern (1939) was more interested in basic research than its application. “Indeed, when I began in 1901 to examine the correctness of recollection among my students, I was determined by theoretical interests in the realm of memory rather than by any practical considerations. Yet once confronted with the results, I realized the importance of this research beyond the borders of mere academic psychology” (p. 4).

Throughout that first decade of the 20th century, Stern was an active researcher in the psychology of testimony. He also helped establish and edited the first journal on the psychology of testimony, *Beträge zur Psychologie der Aussage (Contributions to the Psychology of Testimony)*, which was published in Leipzig. The journal was superseded in 1907 by the much broader *Zeitschrift für Angewandte Psychologie (Journal of Applied Psychology)*, edited by Stern and his colleague Otto Lipmann. In a cautionary note about his research, Stern stressed that most witnesses did not intentionally falsify their reports. Rather, the subtle and common problem created was one of unintentional falsification: “Subjective sincerity does not guarantee objective truthfulness,” he wrote (1939, p. 13). In his research, Stern concluded among other things that: (1) leading and suggestive questions contaminate the accuracy of eyewitness accounts of critical events; (2) there are important differences between adult and child witnesses; (3) lineups are of limited value when the members are not matched for age and physical appearance; and (4) interceding events between an initial event and its recall can have drastic effects on memory. Therefore, modern forensic psychology began as legal psychology with empirical research on the psychology of testimony.

During these early years, European psychologists interacted much more regularly with the law than their American counterparts did. Despite the fact that Stern and Binet, for example, did not initially intend that their research on suggestibility and reliability of observation be applied to the law, they eventually did recommend such an application. Thus European, particularly German, psychologists conducted experimental research, lectured, and consulted with jurists, particularly in the latter half of the 19th century and into the 20th (Mülberger, 2009; Sporer, 1982).

Courtroom Testimony. Pinpointing the origins of courtroom testimony by psychologists in Europe is not easy. Sources differ, often depending on the nature of the forum (e.g., civil versus criminal court, preliminary hearing versus trial) or its

context (informal conversation with a judge versus formal testimony). Hale (1980) suggests that the earliest testimony by a psychologist in a criminal court occurred in 1896, when Albert von Schrenck-Notzing testified at the trial of a Munich man accused of murdering three women. The murders had received extensive and sensational press coverage in the months prior to the trial, and Schrenck-Notzing (1897) opined that this pretrial publicity, through a process of suggestion, probably led numerous witnesses to retroactive memory-falsification. Witnesses could not distinguish between what they had seen and what the press reported had happened. Schrenck-Notzing supported this opinion with social framework testimony (Monahan & Walker, 1988) in the form of accounts of laboratory research on memory and suggestibility. Although the accused was convicted on the basis of solid evidence, Schrenck-Notzing's direct application of the psychology of suggestion to court processes helped stimulate the interest of both German jurists and psychologists (Hale, 1980).

However, Karl Marbe, a psychology professor at the University of Wurzburg, credited himself with the first court appearance, 15 years later. "The first German psychological legal expert opinion was my testimony in a case of sexual assault in Wurzburg in 1911, in which I had to discuss the question of the testimony of children" (Marbe, 1936, p. 184). In that case, several German adolescent girls had accused their teacher of sexually molesting them. Marbe persuaded the jury that the girls' statements were unreliable, and the teacher was exonerated.

Also in 1911, several psychologists testified in a Belgian murder trial in which a man was accused of raping and killing a 9-year-old girl. Two of the child's playmates had apparently seen the murderer but gave inconsistent and contradictory accounts. Among the psychologists retained by the defense was Julian Varendonck, who designed a series of experiments based on questions suggested by information obtained at the preliminary hearing. Varendonck's subjects were children of approximately the same age as the two witnesses (8 to 10). He found that they were inaccurate in their recall of important events. Over the objection of the prosecution, he was allowed to present the results of these experiments as well as the general research on the psychology of testimony that was available at that time. Whipple (1912) wrote that Varendonck's testimony "elicited violent outbursts from the court authorities, but it reached the jury and induced a verdict of 'not guilty'" (p. 268), Whipple added that the psychology of testimony had "found its way formally into the court room and saved a man's life." The jury found the defendant not guilty.

Varendonck, it should be noted, was vehemently opposed to *any* use of child witnesses in the courtroom. In contrast, both Binet (1900) and Stern (1939) believed that errors in recollection, whether by children or adults, were more a reflection of leading, suggestive courtroom questioning than of any "natural" tendency to distort reality.

In 1912, Marbe became one of the earliest European psychologists to testify at a *civil* trial, offering expert opinion on the psychological issue of reaction times as applied to a train wreck near Müllheim. Marbe was asked to testify as to the probable

effect of alcohol both on the mental status of the engineer and the reaction time of the fireman and guard applying the brakes. Based on reaction time experiments, Marbe testified that the train could not have been stopped in time to avert a disaster. As he did in the criminal case, Marbe appears to take credit for paving the way for other psychologists: "Since that time, through my agency and that of others, a mass of psychological expert testimony has been submitted, bearing continually upon new circumstances" (Marbe, 1936, p. 184).

Although Mülberger (2009) wrote that other psychologists were testifying in civil courts even before Marbe's time, it is difficult to find written documentation of who they might have been. Marbe, along with Stern, has been credited with developing forensic psychology in Germany (Sprung & Sprung, 2001). In essence, it is not difficult to find illustrations of psychologists who had impact on the nascent field of legal psychology, but ranking their contributions chronologically must be done with caution.

European psychologists at the turn of the 20th century and until World War I also were delving into the area of guilt deception, the precursor of the lie detection of today. In 1904, psychologists in Germany, Austria, and Switzerland were busy developing a lie detection test for use in criminal investigations. The test was a word association/reaction time task in which key words were embedded in a list of innocuous words. Presumably, the slower the reaction time in recognizing the key words, the more likely the respondent was trying to deceive. Barland (1988), who has reviewed this history in impressive detail, notes that this approach did not catch on because it was inefficient, time consuming, and often yielded inconclusive results.

DEVELOPMENTS IN THE UNITED STATES

At the turn of the 20th century, American psychologists remained comparatively uninterested in applying research on topics related to law. One reason was that they were just beginning to explore the broad psychological landscape and had little inclination to specialize in law-related matters. This reticence was probably also due to the influence of Wilhelm Wundt, who had trained many of the American pioneers in his Leipzig laboratory (Cattell being the first). Wundt, a philosopher and an experimentalist, was wary of applying psychology until sufficient research had been conducted. He believed that the premature use of partial information could be disastrous. His students often took this caveat quite seriously, although some, like Cattell, eventually began to link the laboratory to the world outside.

One of Wundt's not-so-cautious students was the German psychologist Hugo Münsterberg, who arrived in the United States in 1892 at the invitation of William James to direct the psychology laboratory at Harvard University. Münsterberg spent 24 years trying to persuade the public that psychology had something to offer virtually every area of human endeavor. Now acknowledged by many as the father of applied psychology, he believed psychological knowledge could be applied to education, industry, advertising, music, art, and, of course, law. His claims were

often exaggerated, however, and his proposals were rarely empirically based. He usually published in popular magazines rather than in scholarly journals (some of his colleagues called his a “Sunday-supplement psychology”). He also incessantly promoted himself and his native Germany, a practice that alienated him increasingly from his colleagues and the public as World War I approached. In fact, his ardent pro-German stance may have had as much to do with the public’s antipathy toward him as his abrasive personality.

Not surprisingly, the legal community vehemently resisted his intrusion into its territory (Hale, 1980), and there was much ado about this. Charles C. Moore (1907), a well-known attorney, referred to Münsterberg’s work as “yellow psychology” (a term that mirrored the sensational, often inaccurate *yellow journalism* of that era) and concluded that it provided nothing new or helpful to the court. Most noteworthy, the great legal commentator John Henry Wigmore (1909) found it necessary to assail Münsterberg in a satirical and devastating law review article. Wigmore’s attack was prompted by the publication of Münsterberg’s (1908) controversial best-seller *On the Witness Stand*, in which he proclaimed that the time was ripe to apply psychology to the practical needs of the legal system. The book—which was essentially a compilation of already published columns—dealt with a wide spectrum of topics, ranging from witness accuracy and jury persuasion to hypnosis and lie detection.

In 1914, Münsterberg published a study on group decision making, using Harvard and Radcliffe students as subjects, which he titled “The Mind of the Juryman.” In a conclusion not atypical of the times, he stated that “the psychologist has every reason to be satisfied with the jury system as long as the women are kept out of it” (p. 202). He based his conclusion on a finding that the female students in his study were less accurate in their final decisions than the male students. Interestingly, as will be noted shortly, one of his own students later arrived at a very different conclusion.

Münsterberg, always willing to give speeches, gave his inaugural lecture at Radcliffe College in 1894 and his last at the same location in 1916, when he suddenly died of a heart attack midsentence while lecturing his general psychology class (Landy, 1992). Landy wrote that “at the time of his death . . . Münsterberg was an object of public scorn and was well on the way to professional ostracism. By 1919, less than 3 years after his death, there was hardly any reference to any of his more than 10 books and dozens of articles in basic and applied psychology” (p. 787). Benjamin (2003) noted that Münsterberg “was one of the most despised individuals in America” (p. 734). Interestingly, in a recent article, Sporer (2008) correctly pointed out that much valuable information about early contributions of other individuals in legal psychology has been lost because of excessive focus on Münsterberg.

In similar fashion, Bornstein and Penrod (2008) sought to resurrect the long-ignored work of George Frederick Arnold, a civil servant in the British Empire who published *Psychology Applied to Legal Evidence and Other Constructions of Law* in 1906, 2 years before Münsterberg’s *On the Witness Stand*. Bornstein and Penrod admirably compared the value of these respective texts, noting that Arnold, even though he

was not an academician, displayed an impressive familiarity with the psychological literature of the day. They noted also that his style was dry and “reads like the serious academic tome that it is” (p. 763), whereas Münsterberg’s style was directed at a general, less serious audience. Bornstein and Penrod are to be commended for bringing attention to this obscure work, but the fact remains that Arnold’s overall contributions were not as far reaching as those of Münsterberg.

Münsterberg has been accused of being more an opportunist than a trailblazer, however (Kuna, 1978). It is tempting to blame his brashness, his apparently despicable demeanor, and his pro-German views for the tenuous and occasionally hostile initial relationship between psychology and law. Nonetheless, he undeniably pushed his reluctant American colleagues into the practical legal arena and made a seminal contribution to applied psychology in general and forensic psychology in particular.

World War I placed in abeyance most of the exploration in applying psychology to law, although the war and early postwar years saw a few landmarks in American forensic psychology, including the gradual acceptance of psychologists as expert witnesses. The first psychologists, along with other social scientists, were also appointed to law school faculties during these years.

Psychologist Donald Slesinger, a protégé of Robert M. Hutchins, made his mark during the years immediately following World War I. Although he had no formal legal training, Slesinger was appointed by Acting Dean Hutchins as a one-year Sterling Fellow to the Yale Law School in 1927. The following year, he became a research assistant. In 1929, he was appointed associate professor, teaching a course in the psychology of evidence, which appears to qualify him as the first psychologist granted faculty status in an American law school. In 1930, Slesinger followed Hutchins to the University of Chicago, where he served as professor of law and, briefly, as dean of the law school.

Several years earlier, psychologist William Marston had been the first to receive a faculty appointment as professor of legal psychology when he joined the faculty at American University in 1922. Marston was by far the most influential psychologist associated with the legal system during this era. He was a student of Münsterberg but did not have his mentor’s penchant for alienating the legal community and much of the American public. He received a law degree in 1918 and a PhD in Psychology in 1921, both from Harvard. Marston’s interests were multifaceted. (He was even the originator, cartoonist, and producer of the successful comic strip *Wonder Woman*, under the pen name Charles Moulton.) Although admitted to the Massachusetts bar, Marston soon gave up his law practice to concentrate on psychology.

As a laboratory assistant in psychology at Radcliffe College, Marston (1917) had discovered a significant positive correlation between systolic blood pressure and lying, which became the basis of the modern polygraph. In fact, Marston was the psychologist who testified in the landmark case *Frye v. U.S.* (1923), the case that set the original standard for the acceptance of expert testimony in federal courts.

Although his continuing work in lie detection (Marston, 1920, 1921, 1925) represents one of his major contributions to the forensic area, it was by no means the only one. He frequently consulted with attorneys, police, and other criminal justice personnel, and his evidence was determinative in the acquittals of several defendants accused of murder. It is likely, therefore, that Marston—along with Lewis Terman and psychologists associated with the New York City Psychopathic Clinic (both to be discussed later in the chapter)—qualifies as one of the first psychological consultants to the criminal justice system in the United States.

Marston also conducted the first serious research on the jury system (Winick, 1961). Using subjects in simulated jury conditions, he found in a series of studies (Marston, 1924) that written evidence was superior to oral evidence; free narration, though less complete, was more accurate than cross-examination or direct questioning; a witness's caution in answering was a good indicator of accuracy; and female jurors considered evidence more carefully than male jurors (compare with Münsterberg's conclusions about female jurors, mentioned earlier). Because of his legal background and his cautious style, Marston's ideas and research were more acceptable to the legal community than Münsterberg's had been, although there is little evidence that the legal system put his findings to extensive use. This is not surprising because some of his recommendations (e.g., free recall rather than directed questions and cross-examinations) were inapposite to the adversarial process in the United States, and others would have required fundamental changes in court procedures. Interestingly, the German psychologist Stern, discussed earlier, had cautioned his colleagues that experimental research in psychology might be of more relevance to the inquisitorial process used in European courts, where a neutral jurist asked questions of witnesses, than to the adversarial process in the United States (Stern, 1939).

Also during this time period, various reviewers took on the task of documenting the progress of legal psychology. Hutchins and Slesinger, for example, coauthored numerous summary articles on its status (1927, 1928a, 1928b, 1928c, 1929). Slesinger wrote another article with Marion Pilpel in 1929, surveying 48 articles written by psychologists on issues relating to the law that had appeared in professional journals up to that time. Eleven were concerned with the psychology of testimony, 10 with deception, 7 with intelligence and crime, and 6 with criminal behavior. The remainder focused on general topics such as the scientific method or legal research. Fifteen of the 48 articles had been written by German psychologists.

Like applied psychology in general, legal psychology was somewhat dormant between the two world wars and did not regain its energy until the late 1940s and 1950s. In addition to Marston's work, the period did see scattered research on how juries formed opinions and verdicts (Weld & Danzig, 1940; Weld & Roff, 1938), a master's thesis on the relationship between narrative and interrogative methods of questioning (Cady, 1924), another study on questioning and testimony (Snee & Lush, 1941), and a survey of legal and psychological opinions about the validity of some of Wigmore's rules of evidence (Britt, 1940).

According to Loh (1981), there was some interest in psychology and law during the late 1920s and the 1930s. However, this interest was almost exclusively on the part of lawyers, who produced such books as *Legal Psychology* (Brown, 1926), *Psychology for the Lawyer* (McCarty, 1929), and *Law and the Social Sciences* (Cairns, 1935). Wigmore (1940), the foremost authority on rules of evidence, paved the way for the use of test data in the courtroom. He observed that the psychometrist introducing test evidence would stand “on the same footing as the expert witness to insanity” (cited by McCarty, 1956, p. 9), as long as such tests are recognized as valid and feasible by the general scientific community.

In 1931, psychologist Harold Burt (who referred to Münsterberg as his mentor at Harvard) wrote *Legal Psychology*, possibly the first textbook in the area. Disputing this claim, Mülberger (2009) commented that the German psychologist Otto Lipmann had published a psychological textbook for jurists long before this (in 1908). The truth may depend on the meaning of the word *textbook*. Lipmann (1908) clearly deserves credit for his work, which was a compilation of the lectures he gave to students studying law. Lipmann’s book was specifically intended to educate current and future judges and lawyers, whereas Burt’s book was intended for both lawyers and students of applied psychology. Nevertheless, although Burt’s book made a valuable contribution to the academic psychological literature, it had little discernible influence on the legal profession or on applied psychology in general. In 1935, Edward S. Robinson published *Law and the Lawyers*, which predicted that jurisprudence would become one of the family of social sciences and argued that all of its fundamental concepts must be brought into line with psychological knowledge. The book was lambasted by lawyers and essentially ignored by psychologists. In hindsight, later scholars found Robinson’s ideas much more palatable (e.g., Horowitz & Willging, 1984; Loh, 1981).

EXPERT TESTIMONY

It is generally believed that American psychologists have served as expert witnesses since the early 1920s (Comment, 1979), but, like their European counterparts, they consulted with lawyers and the courts, perhaps particularly the civil courts, before that time. Included in this latter category are the juvenile courts, which were a hybrid of the civil and the criminal, dealing with matters of both child protection and delinquency. Psychological consultation with juvenile courts was common from their inception in 1899 (Brigham & Grisso, 2003). Consultation with and testimony in criminal courts was much less common, as we discuss shortly.

According to Rogers (1910, 1918), the results of experimental research on visual perception were routinely accepted in trademark infringement cases. In *Coca-Cola Company v. Chero-Cola Company* (1921), for example, an experimental psychologist was asked whether the trademarks used by the two companies were so similar as to be likely to cause confusion in the public mind and ultimately deceive the consumer. This was apparently considered a “safe” undertaking, as the psychologists

were not infringing on the territory of the “medical experts”—physicians and psychiatrists—who routinely testified on matters of criminal responsibility. As Louisell (1955) noted, however, because trial court records are generally unavailable and only appellate decisions are published, the testimony of psychologists, particularly in civil cases, may have been less rare than the paucity of documentation would indicate. We do know that psychological testimony was almost inevitably *rejected* in criminal cases involving the defendant’s mental state. “As a general rule, only medical men—that is, persons licensed by law to practice the profession of medicine—can testify as experts on the question of insanity; and the propriety of this general limitation is too patent to permit discussion” (*Odom v. State*, 1911; cited in Comment, 1979, fn. 14).

The first published case in which an American psychologist qualified as an expert appears to be *State v. Driver* in 1921. The occasion was only a partial victory for forensic psychology, however. A West Virginia trial court accepted the chief psychologist of the State Bureau of Juvenile Research as an expert on the matter of juvenile delinquency. However, it rejected his testimony, based on psychological test data, that a 12-year-old alleged victim of an attempted rape was a “moron” (in retrospect, an unfortunate term coined by Henry H. Goddard, who is discussed later) and could not be presumptively believed. In agreeing with the trial court, the West Virginia Supreme Court noted, “It is yet to be demonstrated that psychological and medical tests are practical, and will detect the lie on the witness stand” (*State v. Driver*, p. 488). Although some commentators interpreted *Driver* as a major loss for psychologists wishing to achieve status as expert witnesses, Louisell (1955) noted that the decision was not a rejection of psychologists per se, only of the particular evidence offered by one psychologist.

Nevertheless, it was not until much later, in the 1940s and 1950s, that psychologists testified in courts of law on a regular basis, at least in some jurisdictions. They offered opinions and presented data relevant to subjects as diverse as the influence of pretrial publicity on potential witnesses and juries, the effects of pornography on adolescents, the effect of certain educational practices on children, and the likely influence of advertisements on consumers (Greenberg, 1956; Loh, 1981; Louisell, 1955). This is not to say that there was widespread acceptance of the idea that psychologists deserved a niche in the courtroom. Resistance to the idea, or at best a cautious approach, consistently characterized much of the legal literature (Comment, 1979).

In the early 1940s and the post–World War II era, appellate courts also began to hand down rulings that allowed psychologists to offer expert testimony in trial courts on the issue of mental responsibility for criminal and tortious conduct. Loh (1981) attributed this eventual acceptance to an increase in professionalization, “the rapid growth of mental health professions during this period, and the formulation of legal doctrines of insanity consistent with modern psychiatry” (p. 323).

One important decision, perhaps the first *influential* decision, was *People v. Hawthorne* (1940), a Michigan case. Hawthorne had been tried for the murder of his wife’s lover and had pleaded not guilty by reason of insanity. The trial

court refused to qualify as an expert witness a professor of psychology from Michigan State Normal College who had a doctoral degree and an impressive list of credentials. In finding that the trial court had erred in not accepting the psychologist as an expert, the Michigan Supreme Court ruled that the standard for determining expert status was not a medical degree but the extent of the witness's knowledge. It advised trial courts to evaluate carefully the merits of a potential witness's claim to expertise, noting that a psychologist's ability to detect insanity could not be presumed inferior to that of a "medical man." The dissenters, however, believed that insanity is a disease and therefore only a person with medical training should qualify as an expert.

Later, in *Hidden v. Mutual Life Insurance Co.* (1954), the Fourth Circuit Court of Appeals allowed psychological expertise to be applied to a *civil* case relating to mental status. The plaintiff argued that a disabling nervous condition prevented him from engaging in any gainful occupation and entitled him to disability benefits. A clinical psychologist with a doctoral degree administered a battery of projective tests and testified on his behalf. Not only did he report on the test results, but he also gave the opinion that the plaintiff deserved the benefits. When the lawyer for the insurance company objected, the trial judge instructed the jury to disregard the entire opinion testimony on the grounds that the psychologist did not qualify as an expert. The circuit court of appeals ruled that the psychologist should have been qualified as an expert to express his opinion about the plaintiff's mental condition.

While some psychologists were struggling to be accepted as experts on questions of mental status, competence, and criminal responsibility, others during this era were joining the crucial legal battle against school segregation by testifying and consulting with attorneys in the state cases that would ultimately culminate in the 1954 landmark ruling *Brown v. Board of Education* (Kluger, 1975). David Krech and Helen Trager, social psychologists who had published articles on racial attitude tests, and Horace B. English, an expert on child psychology, were among many who testified for the plaintiffs at some of the school segregation trials. Psychologist Henry Garrett, a former president of the APA, testified on behalf of the state (Jackson, 2000). Perhaps the most widely publicized—and since then highly critiqued—contribution on behalf of the plaintiffs was that of Kenneth Clark and Mamie Clark, who conducted the now-famous "doll research" to gauge the effects of segregation. Kenneth Clark then gave social framework testimony reporting the results of this research (Kluger, 1975). When the National Association for the Advancement of Colored People (NAACP) appealed *Brown* and three other segregation cases to the U.S. Supreme Court, Kenneth Clark, Isidor Chein, and Stuart W. Cook wrote the Social Science Statement that included signatures of 32 eminent social scientists (Jackson, 2000).

This was not, however, the first social science brief to be submitted to an appellate court. According to Brigham and Grisso (2003), that distinction belongs to the brief submitted to the Oregon Supreme Court in *Muller v. Oregon* (1908). In that

case, Louis Brandeis—who later became a prominent justice of the U.S. Supreme Court—argued in support of the state that work hours of women should be limited because social science data demonstrated their inherent weakness.

History has not been kind to the scientists in either case. Brandeis's patriarchal argument in the *Muller* case would be deplored and roundly denounced today, both for its tenor and for its lack of empirical support and rigor. Social scientists in the *Brown* case were criticized for their naive methodology, lack of objectivity, and faulty conclusions based on insufficient scientific evidence (Jackson, 2000). In his historiographical inquiry, however, Jackson noted that the doll experiments were but one prong of many studies that psychologists and other social scientists referenced in their trial testimony and in the brief submitted to the Supreme Court. He also argued convincingly that critiques of these social scientists reflected a misreading of their testimony, their research, and their evaluation of relevant evidence. (See also Brigham & Grisso, 2003, for an enlightening discussion of psychology's involvement in both of these cases.)

During the same era, psychologists were continuing to make enough inroads testifying on the issue of criminal responsibility that psychiatrists felt the need to protect their turf. In 1954, the Council of the American Psychiatric Association, the Executive Council of the American Psychoanalytical Association, and the American Medical Association joined in a resolution stating that only physicians were legitimate experts in the field of mental illness for purposes of courtroom testimony. Other individuals could participate only if their testimony was coordinated by medical authority. The resolution greatly influenced trial courts (Miller, Lower, & Blechmore, 1978), which became reluctant to accept independent psychological testimony.

Finally, in *Jenkins v. United States* (1962), the Court of Appeals for the District of Columbia gave its own direct, although conditional, support to the use of psychologists as experts on the issue of mental illness. Although the court was sharply divided, its decision remains the predominant authority for the use of psychologists in the area of criminal responsibility. Following that opinion, federal courts and increasingly more state courts certified psychologists as expert witnesses in both criminal and civil cases.

COGNITIVE AND PERSONALITY ASSESSMENT

During the years in which Münsterberg was proselytizing about psychology's usefulness in the courtroom, particularly involving expert testimony, another American psychologist was more quietly making inroads into a different forensic area, one specifically related to juvenile courts. As we noted earlier, consultation with these courts was common, but it was chiefly in the area of assessment. In 1909, clinical psychologist Grace M. Fernald worked with psychiatrist William Healy to establish the first clinic designed for youthful offenders, the Juvenile Psychopathic Institute.

It was initially developed to serve the newly established Juvenile Court of Chicago by offering diagnoses of “problem” children. Fernald, who received her doctorate from the University of Chicago in 1907, was probably the first clinical psychologist to work under the supervision of a psychiatrist (Napoli, 1981) as well as one of the earliest psychologists to specialize in the diagnosis and treatment of children and adolescents who appeared before the juvenile courts. The institute, which extended its services rapidly to include treatment and research as well as diagnosis, became a public agency in 1914, the Institute for Juvenile Research. Arguably, it also provided the earliest formal internships in forensic psychology in the country (Resnick, 1997).

Fernald and Healy used the relatively new Stanford-Binet Intelligence Scale to assess delinquents, but they soon realized the importance of obtaining “performance” measures as well. This prompted them to develop the Healy-Fernald series of 23 performance tests, which they began using in 1911. The two eventually went their separate ways. Fernald became a specialist in intellectual disability and intelligence and testing and taught psychology at the University of California–Los Angeles for 27 years, until her retirement in 1948. Healy, along with psychologist Augusta Bronner, went on to establish the Judge Baker Clinic in Boston in 1917. During the first third of the 20th century, most psychologists providing regular services to the courts were psychometrists associated with clinics. The term *forensic psychology* had not been minted, and *legal psychologists* were in the halls of academe or consulting sporadically with judges and lawyers. Thus, it seems that much of the forensic work of psychologists during this period consisted of cognitive and personality assessments of individuals, both juveniles and adults, who were to come before the courts. The drudgery of day-to-day testing (often under the watchful eyes of a physician or psychiatrist) made applied psychology unappealing as a profession. Often, however, it was where female psychologists were most accepted. In the 1930s, for example, fewer than one-third of all American psychologists were women, but women made up over 60% of all applied psychologists (Napoli, 1981).

In one of the first published accounts of the work of these early psychometrists, E. I. Keller (1918) described some of the challenges they faced. He noted that in December 1916, a psychopathic laboratory was established at the New York City Police Department for the express purpose of examining persons detained before trial. The staff included psychiatrists, neurologists, social workers, and psychologists, whose task was to conduct hasty pretrial evaluations. (Because these psychologists worked out of the police department but conducted evaluations for the courts, they could be considered both legal and police psychologists.) According to Keller, who was a consulting psychologist at the clinic, detainees arrived for testing at 9 A.M. “The disadvantage is the lack of time, for all prisoners [*sic*] must be examined in time to get them to court by noon or earlier, and many courts are situated in distant parts of the city” (p. 85). Staff members had little time in which

to conduct the evaluation and prepare a report that would help the court in its decision making.

The work of Henry H. Goddard during this time must—in hindsight—be regarded with embarrassment. A student of noted psychologist G. Stanley Hall, Goddard paved the way for the massive intelligence testing of immigrants and residents of mental institutions, prisons, and juvenile training schools. His followers consulted with the juvenile courts and dutifully administered these tests to the children of the poor who arrived at their door. Goddard’s warning that “feeble-minded” individuals should not be allowed to roam about freely in society because of their innate proclivity toward antisocial behavior contributed significantly to the incarceration of individuals during their reproductive periods and the sterilization of residents in both juvenile and adult facilities (Kelves, 1984).

Psychologists continued to work in court clinics during the second third of the 20th century, performing a variety of tasks related to the assessment process (see Box 1.1). In addition, as we described earlier, they gradually became more involved in providing expert testimony, not only on the results of their assessments but also on research that was relevant to legal issues. Other psychologists continued to offer services to inmates and staff of jails and prisons, an endeavor that apparently began early in the 20th century. It is to this second aspect of forensic psychology that we now turn.

Box 1.1 Help Wanted: Court Psychologist

An article in Volume 1 of the *American Psychologist* (Shartle, 1946) carried the following job description for a court psychologist.

COURT PSYCHOLOGIST

(Clinical Psychologist)

Duties

Interviews offenders referred by the court to determine the causes of the crime, the attitudes and conflicts, and the educational, vocational, and social background of the client. Also may interview parents and guardians.

Administers and interprets individual intelligence, performance, and personality tests including projective techniques.

Writes complete case histories including interview information and test interpretations. Presents case histories and recommended treatment to colleagues including medical and other officers of the court. May testify in court.

Qualifications include MA in psychology with a PhD preferred, relevant course work (e.g., abnormal, clinical, psychometrics, criminology, medical subjects), previous experience, and emotional maturity.

Interestingly, Shartle noted that, although few psychologists were employed in such positions, there was indication that employment in the field would increase. However, “higher positions” in the court were not usually open to psychologists.

CORRECTIONAL PSYCHOLOGY

Lindner (1955) pinpointed 1913 as the date when psychological services were first offered in a U.S. correctional facility, specifically a women's reformatory in the state of New York. Watkins (1992) identified the psychologist as Eleanor Rowland, who was asked to devise a test battery to identify offenders who would benefit from educational programs and be safely returned to society (Rowland, 1913). However, the main function of psychologists employed in some capacity in the state and federal correctional systems during these years was apparently the detection of "feble-mindedness" among offenders, a condition thought to lead to a life of crime (Giardini, 1942; Watkins, 1992). Again, the work of Goddard and his followers is relevant.

Concurrently, however, some psychologists—like Rowland—became involved in a different endeavor: the classification of inmates into various groups for determining where they were to be placed (custody decisions) and what services might be provided (treatment decisions). The first prison classification system developed by psychologists was apparently instituted in New Jersey in 1918 (Barnes & Teeters, 1959; Watkins, 1992). New Jersey also became the first state to hire a full-time correctional psychologist. The first state in the United States to provide comprehensive psychological examinations of all admissions to its prison system and applications for parole was Wisconsin, in 1924 (Bodemar, 1956).

In the late 1930s, Darley and Berdie (1940) surveyed 13 federal and 123 state prisons and learned that they employed a total of 64 psychologists who called themselves "prison psychologists." Although all considered themselves clinical psychologists, only about half had doctorates in psychology. Later, Raymond Corsini (1945) expressed concern that there was as yet "no history of prison psychology." He estimated that during the 1940s, there were approximately 200,000 individuals confined in U.S. correctional facilities who were served by a mere 80 psychologists. Their work consisted of (1) testing (personality, aptitude, and academic progress); (2) providing educational, vocational, and personal guidance (usually at the inmate's request); and (3) maintaining working relationships with all members of the prison staff (see Box 1.2). In one of the most comprehensive surveys undertaken during the early 1940s, questionnaires were sent to 4,580 psychologists (3,209 men and 1,371 women) in an effort to discover the nature of the profession (Bryan & Boring, 1946). Of the 3,241 questionnaires returned in 1940, 76 men and 20 women indicated they were employed as full-time psychologists in prisons or correctional institutions. Of the 3,106 questionnaires returned by the same group in 1944, 53 men and 27 women said they were employed in prisons or correctional institutions. Although these data support Corsini's estimation that between 80 and 100 psychologists were employed in the nation's correctional facilities during the early to mid-1940s, it is interesting to note that, by the mid-1940s, approximately one-third of prison psychologists were women.

Box 1.2 Help Wanted: Correctional Psychologist

1940s VERSION

In Volume 1 of the *American Psychologist*, Shartle (1946) described the work of a prison psychologist.

PSYCHOLOGIST, PENAL INSTITUTION

(Prison Psychologist)

Duties

Administers intelligence, aptitude, and other tests to either all inmates or certain groups depending on institutional policy. Writes an interpretation of test results for the prisoner's records.

Interviews each prisoner to determine background, attitudes, and personality traits for use in guidance, education, possibilities for parole, and placement. Results of interview are written and may be submitted in form of case study with test results or other reports.

Makes recommendations for parole and supplies technical information at staff meetings. Gives information in consultation with administrative officers or with specialists in the field of medicine, psychiatry, sociology, education, occupational training, or parole.

Assists in planning or revising programs for medically sponsored cases including psychiatric and severe physical disability cases.

Participates in research. Investigates problems of penal psychology or test construction and prepares reports of finding.

Again it was noted that opportunities in the field were limited and the number of openings not numerous. However, several states were planning postwar expansion in buildings and services.

Psychologists entered the Canadian correctional system much later, perhaps as late as the early 1950s. Watkins (1992) notes that Canadian correctional psychology made its first appearance in the literature in 1952 in a series of newsletters published by the Ontario Psychological Association. The newsletters focused on psychology in the Ontario provincial corrections programs and the federal correctional service. The first correctional psychologist in the federal system in Canada was employed in 1955 at St. Vincent de Paul Penitentiary (later renamed Laval Institution) in Quebec (Watkins, 1992). Correctional psychologists in Canada were at first employed primarily to classify inmates for security placement and were usually not a component of the mental health treatment afforded to inmates. In the United States, their role appears to have been broader (see Box 1.2). Since these early days, however, Canada in many ways has outpaced American corrections—particularly state prison systems—both in developing risk assessment instruments and providing rehabilitation services to inmates (Wormith & Luong, 2007).

Classification, however, has always been an important enterprise for psychologists working in correctional settings. Reliable offender classification was (and is) both an important service to offer to correctional administrators and in many respects a prerequisite to effective treatment. In both the United States and Canada,

from the mid-20th century on, psychologists became increasingly involved in developing and testing more sophisticated classification systems. One of the earliest of these “modern” systems was the Jesness (1971) Classification System. Best known, however, was the system proposed by Edwin Megargee and based on the Minnesota Multiphasic Personality Inventory (MMPI). Megargee (1977), using his research on overcontrolled and undercontrolled personalities as a springboard, identified 10 “inmate types.” Prison officials then made use of these groupings to assign inmates to custody levels, job assignments, and rehabilitation programs. Megargee’s system is still in use in some prison systems, and Clements (1996) observed that Megargee deserves much credit for providing correctional psychologists with an excellent list of seven criteria for a good classification system.

In the 1960s and early 1970s, correctional psychology as a subdiscipline of forensic psychology began to expand. Even to this day, though, many if not most psychologists working in corrections prefer to be called correctional psychologists rather than forensic psychologists (Magaletta, Patry, Dietz, & Ax, 2007). This may be because they see their primary function as one of providing services to inmates, not to the legal system. Until the 1960s and 1970s, although there were exceptions, psychologists in correctional facilities focused more on classification than on treatment, although important treatment models were proposed by psychologists such as Herbert Quay and Marguerite Warren (Brodsky, 2007). Nevertheless, treatment was not the predominant activity, both because the demand for diagnostic services was great and the obstacles relative to respecting confidentiality and achieving the trust of inmates were difficult to surmount.

Perhaps even more relevant was the suspicion directed toward psychologists by both administrative and correctional staffs. In an essay reviewing this period in the history of correctional psychology, Brodsky (2007) cited examples of military psychologists being given punitive assignments or civilian psychologists being obstructed from providing meaningful treatment services to inmates—in some cases even reporting for work to find themselves no longer employed, their possessions waiting for them at the prison gate. “With the exception of psychologists in the Federal Bureau of Prisons, psychologists working in American prisons reported organizational impediments to conducting meaningful assessments and offering meaningful treatment” (p. 864).

In the 1960s, rehabilitation as a correctional goal began to gain favor, and—in some but certainly not all prison settings—psychologists spent more time working directly with offenders and providing treatment services. Although positions were plentiful, the turnover rate was high, primarily because psychologists often had not received adequate preparation for responding to the unique challenges of these environments (Watkins, 1992).

One noteworthy innovation that was introduced in federal prisons during this era was the unit management system, which was initially conceptualized by Daniel Glaser (1964) and later promoted by Robert Levinson (Toch, 1992). Unit management divided prison populations into small groups of prisoners and staff

members based on the programming needs of the former and the expertise of the latter. Some units—those in which more intensive treatment services could be provided—became “therapeutic communities.” Other units provided education, training, or work experiences, together with some counseling (Toch, 1992). Although unit management lost support in the United States during the punitive 1980s and 1990s (with overcrowding having its obvious effects), the concept survives in some state and federal facilities, particularly where substance abuse treatment is provided.

Many correctional psychologists worked in the trenches during the 1960s and early 1970s and made significant contributions. Stanley Brodsky was instrumental in launching modern correctional psychology in the United States, but many other individuals (e.g., Robert Levinson, Ascher Pacht, Hans Toch, Edwin Megargee, and Marguerite Warren) made significant contributions as well. Canada has its own group of pioneers who have had great impact on correctional philosophy and practice on an international level. They include psychologists Paul Gendreau (coauthor of Chapter 23 in this volume), Karl Hanson, Don Andrews, and many others whose work is cited in the excellent historical reviews and summaries of Watkins (1992) and Wormith and Luong (2007).

In the United States, Brodsky’s term as president of the American Association for Correctional Psychology (AACP) helped provide the impetus to move correctional psychology into a recognized and viable profession. (The AACP was actually born in 1953 with the name Society of Correctional Psychologists and underwent several name changes during the late 1950s through the early 1970s [Bartol & Freeman, 2005; Brodsky, 2007]. It is now called the International Association for Correctional and Forensic Psychology.) During 1972 and 1973, with Brodsky at the helm, the AACP played a key role in setting up a series of conferences on psychology in the criminal justice system, with emphasis on corrections. The proceedings were published in a volume edited by Brodsky (1973), *Psychologists in the Criminal Justice System*. The publication of this influential book could arguably be the official launch date of modern correctional psychology, even though the AACP itself predated Brodsky’s book. Brodsky also became the founding editor of the international journal *Criminal Justice and Behavior*, launched in 1974 and sponsored by the AACP. Brodsky’s leadership and enthusiasm also helped build one of the earliest doctoral programs specifically designed to prepare clinical psychologists to work in the criminal justice system, particularly corrections, at the University of Alabama. In the late 1970s, the APA approved a clinical internship in corrections at the Wisconsin Department of Corrections. Today, such programs exist in a variety of colleges and universities, many of which provide internship opportunities for students in state prisons as well as the Federal Bureau of Prisons.

POLICE PSYCHOLOGY

Those who prefer a narrow definition of forensic psychology do not typically include police psychology in its purview. We have done so because police are sworn to

uphold the law and are in many cases the gatekeepers to entry into criminal and juvenile courts, if not civil courts. Thus, psychologists who consult with police in numerous capacities (e.g., investigation, candidate screening, hostage-taking incidents, interviewing strategies) are connected with the legal system.

It is difficult to pinpoint precisely when police psychology began, primarily because individual psychologists have provided a variety of services to law enforcement without their work being formally recognized. Viteles (1929) noted that police departments in Germany used psychologists in a variety of capacities as early as 1919. In the United States, in keeping with the psychometric movement of the early 20th century, contributions centered around assessment, particularly cognitive assessment administered to candidates for law enforcement positions.

Four discernible but overlapping historical trends in American police psychology can be identified: (1) cognitive and aptitude screening, (2) personality assessment and the search for the “police personality,” (3) stress management and other clinical services, and (4) fairness in screening and selection (Bartol & Bartol, 2004). The first trend—1916 to 1960—is characterized by attempts of psychologists to assess the intellectual skills required to be an effective police officer. The second trend—1952 to 1975—focused on the development of personality measures capable of distinguishing effective from less effective officers. During the second trend, there also were many unsuccessful attempts to identify a “police personality.” The third trend—1974 to 1994—was characterized by psychologists becoming increasingly involved in the identification and treatment of stress and other emotional reactions often experienced by police officers and their loved ones. Such topics of interest included the use of excessive force, police decision making, post-shooting traumatic reaction, fitness for duty evaluations, and police suicide.

The fourth trend—1980 to the present—refers to the legal requirements that all persons should have an equal chance of being selected on the basis of individual merit and qualifications. Topics during this trend include the Americans with Disabilities Act of 1990, gender issues in policing, and minority/ethnic/racial composition of law enforcement agencies. Because this chapter focuses on early history, we briefly sketch only the first two trends. It should be noted, however, that police psychologists today are actively involved in consultation with law enforcement and with research in a variety of areas that reflect and transcend the above trends. Many belong to professional organizations, such as the APA’s Division 18, Psychologists in Public Service and its subgroup Police and Public Safety (see Chapter 15 in the present volume).

COGNITIVE AND APTITUDE SCREENING

Lewis Terman (1917) was the first American psychologist to use “mental tests” as screening devices in the selection of law enforcement personnel. On October 31, 1916, at the request of the city manager of San Jose, California, he administered an abbreviated form of the Stanford-Binet to 30 police and fire department applicants.

They ranged in age from 21 to 38, with a median age of 30. Only four had attended high school, and none had attended beyond the sophomore year. Terman found that most of the applicants functioned near the dull-normal range of intelligence (68–84 on the Stanford revision of the Binet-Simon Intelligence Scale); only three obtained an IQ over 100, the score considered average for the general population. Based on his experience with the intellectual capabilities of school-age children, Terman suggested, somewhat arbitrarily, that applicants with an IQ under 80 were not fit for police work or firefighting. The city manager agreed, and 10 applicants were immediately excluded from further consideration.

A contemporary of Terman, psychologist Louis Thurstone, was also interested in the value of intellectual testing in police screening. Thurstone (1922) administered the newly developed Army Intelligence Examination (Army Alpha) to 358 male members of the Detroit Police Department. The Army Alpha, developed by Robert Yerkes, E. L. Thorndike, and Lewis Terman and adopted by the U.S. Army in 1917, was probably the first exclusively American test of intelligence (Resnick, 1997). Police officers at all ranks scored below average on the Army Alpha; in fact, the more experienced the police officer, the lower was his intelligence score. The average score for the 307 patrol officers was 71.44; the sergeants averaged 54.71; and the 17 lieutenants, 57.80 (Army Alpha mean = 100, standard deviation of 15). Thurstone concluded that law enforcement did not attract intelligent individuals, and the more intelligent individuals who entered police service left for other occupations where their abilities and intelligence were better utilized.

Law enforcement officers were vindicated somewhat, however, when Maude A. Merrill (1927) administered the Army Alpha to a group of already employed officers and applicants. They scored at the average level (the sample's mean IQ was 104). The differences between her findings and those of Terman and Thurstone were probably due to department leadership factors, recruitment procedures, and selection ratios (Terrio, Swanson, & Chambelin, 1977). Intelligence testing continued throughout much of the middle part of the 20th century and may still exist in some departments today. However, questions about the validity of such assessment and understandable resistance from police unions persuaded most agencies to turn to a different form of assessment, the personality assessment.

PERSONALITY ASSESSMENT

In the years between the two world wars, psychologists gradually became more involved in the screening of law enforcement personnel and began to incorporate personality assessment into that enterprise. Wilmington, Delaware, and Toledo, Ohio, appear to share the distinction of being the first two cities to require ongoing psychological screening for use in police selection, in the form of mental and personality tests (Gottesman, 1975; Oglesby, 1957). The year was 1938. Thus, personality tests came on the scene at about this time. It was not until the late 1950s and 1960s, though, that personality assessment overtook cognitive tests in the screening of law

enforcement personnel. While the aforementioned psychologists were among the first to study the cognitive capacities of police officers and candidates, there is no indication that they consistently participated in the screening and selection of law enforcement personnel. At this point, we have no information about who might have been the first psychologist to assume this regular role. As late as 1939, Donald Paterson (1940) could identify only one psychologist, L. J. O'Rourke, who had actively investigated the validity of the nation's civil service examination system, even though routine competitive exams were administered as far back as 1883.

During the late 1940s and the 1950s, psychologists continued to consult with police departments. The psychological screening processes initiated by the Wilmington and Toledo police departments was adopted by other cities; Jacksonville in 1947, Berkeley in 1949, Oakland in 1950, New Orleans in 1952, and Pasadena, Philadelphia, Milwaukee, and Cleveland in 1953 (Gottesman, 1975; Oglesby, 1957). In June 1952, the Los Angeles Police Department (LAPD) began to administer a battery of psychological tests (MMPI, Rorschach, and a psychological interview; Rankin, 1957, 1959). The 1957 Rankin article was the first to appear in the literature attesting to any ongoing program of psychological assessment for police applicants (Gottesman, 1975).

During the late 1960s, personality assessment, psychological screening, and police psychology in general received an immense boost when the President's Commission on Law Enforcement and the Administration of Justice (1967) strongly recommended widespread use of psychological measures to determine the emotional stability of all officer candidates. This recommendation was followed by the strong endorsement in 1968 by the National Advisory Commission on Civil Disorder that psychological screening would improve the emotional quality of individuals entering law enforcement (Scrivner, 1994). In keeping with commission recommendations, Congress provided Law Enforcement Assistance Administration funds for law enforcement agencies to retain the services of mental health professionals. In 1973, the Police Task Force Report of the National Commission on Criminal Justice Standards and Goals encouraged the establishment of a behavioral sciences unit or consultant for all law enforcement agencies.

Even before then, though, psychologists were offering services to law enforcement on an as-needed basis, consulting in such areas as stress management, crisis management with the mentally disordered, and domestic violence. According to Nietzel (2000), the first project to train police in crisis intervention techniques in domestic disputes was developed in the late 1960s by Morton Bard, consulting with the New York City Police Department.

At about the same time, in December 1968, Martin Reiser was hired by the Los Angeles Police Department (LAPD) as a full-time police psychologist. The evidence to date indicates that Reiser was the first full-time psychologist whose responsibilities were strictly police related. Reiser (1982) himself is not entirely certain he was the first full-time police psychologist in the country. In 1969, he presented a paper at the Western Psychological Association Convention in Vancouver entitled

“The Police Department Psychologist.” This presentation may represent the “official” launch of contemporary North American police psychology. The paper was published in 1972. Reiser continued to be the most prolific writer on police psychology during the early 1970s. In 1972, in cooperation with the California School of Professional Psychology and the Los Angeles Police department (LAPD), he helped establish what is believed to be the first clinical internship in police psychology in the United States. By 1977, at least six other law enforcement agencies employed full-time psychologists (Reese, 1986, 1987).

CRIMINAL PSYCHOLOGY

In the early years of the 20th century, psychologists began to offer psychological perspectives on criminal behavior and to speculate about the causes of crime. Like the police psychology discussed earlier, criminal psychology typically is not considered in the narrow definitions of forensic psychology, primarily because it appears more theoretical than clinical in nature. However, in its youth, criminal psychology was essentially clinical in nature, as the theories often centered on the measurable mental capacities of offenders. Furthermore, forensic psychology devoid of a theoretical base—such as that provided by criminal psychology—is difficult to justify and support.

Psychologists like Goddard had repeatedly found that most juvenile and adult offenders were “mentally deficient,” which led to the conclusion that a primary “cause” of crime and delinquency was intellectual limitation. In large part, this belief reflected the pervasive influence of Darwinism, which contended that humans differ only in degree from their animal brethren (and that some humans are closer to their animal ancestry than others). The “mentally deficient” were considered both intellectually and morally less capable of adapting to modern society. They presumably resorted to more “primitive” ways of meeting their needs, such as crime. These unfortunate conclusions, which did not take into account social conditions, cultural differences, or socialization processes, lent support to unconscionable practices such as lengthy incarceration of the disadvantaged, confused, and powerless.

In the history of psychology, few scholars have ventured to offer comprehensive theories on crime or delinquent behavior. Those who have (e.g., Eysenck, 1964) have often been strongly influenced by Darwinian thinking. Therefore, theoretical orientations focusing on mental deficiency or biological and constitutional dispositions dominated early psychological criminology.

In the early 1960s, a psychological criminology distinct from psychiatric and more extensive than psychometrics began to show signs of life. Hans Toch (1961), who was also making significant contributions to correctional psychology, edited one of the first books on psychological criminology, *Legal and Criminal Psychology*. Some may argue that Hans Gross published the first criminal psychology book in 1897 (*Kriminalpsychologie*), the same year in which he was appointed professor in ordinary for criminal law and justice administration at the University of Czernowitz

in Austria. One writer has asserted that Gross was the originator of the discipline of criminal psychology (Undeutsch, 1992). However, Gross was a lawyer by training, in practice, and in spirit and eventually became a successful judge. His book details his observations of offenders, witnesses, jurors, and judges but relies very little on psychological research. This is not surprising, of course, because psychology in 1897 was far from being an integrated discipline with a rich body of knowledge. Nevertheless, it is significant that Toch's book, published more than 60 years later, represents the earliest attempt to integrate, in an interdisciplinary fashion, the empirical research of psychologists relevant to criminal behavior and legal issues.

British psychologist Hans J. Eysenck, in *Crime and Personality* (1964), formulated the first comprehensive theoretical statement on criminal behavior advanced by a psychologist. Eysenck's theory focused on the personality characteristics of extraversion and introversion, which he believed could be attributed to both a biological predisposition to seek (extravert) or avoid (introvert) sensation *and* the learning experiences obtained in one's social environment. Although Eysenck's theory was circulated and tested extensively in the late 1960s and 1970s, it has been shifted aside today, replaced by popular developmental approaches. Shortly after Eysenck proposed his theory, Edwin Megargee (1966) put forth his own heuristic statements regarding undercontrolled and overcontrolled personalities and their relationships to violence, a theory that then served as a basis for his classification system referred to earlier. Toch (1969) followed with *Violent Men*. The relationship between aggression and violence was studied seriously under the leadership of Leonard Berkowitz (1962), Albert Bandura (1973; Bandura & Walters, 1959), and later Robert Baron (1977). Following psychiatrist Hervey Cleckley's (1941/1964) groundbreaking work on psychopaths, they became subjects of vigorous theory building and research in the hands of Canadian psychologist Robert Hare (1970) and others (e.g., Quay, 1965). Psychopathy continues to be a rich research area on the etiology of criminal behavior to this day.

1970s AND BEYOND

Since the 1970s, we have witnessed a literature and research explosion in all areas of forensic psychology. Some 30 years ago, Loh (1981) observed that forensic psychology had "come of age." Most recently, Heilbrun and Brooks (2010) noted that "[t]he field has matured: the recognition of the importance of the foundational science [of forensic psychology] is stronger, and we are closer to identifying best practices across a range of legal contexts that are addressed by forensic psychology research and practice" (p. 227). In 1965, just over 100 English-language articles and books related to forensic psychology had been published (Tapp, 1976). By the mid-1970s, the numbers were well into the thousands. Professional journals exclusively devoted to forensic psychological research and issues were beginning to emerge in North America. *Criminal Justice and Behavior* led the way in 1974, followed by *Law and Psychology Review* (a journal published by law students and

graduate psychology students at the University of Alabama) beginning in 1975, *Law and Human Behavior* in 1977, *Behavioral Sciences & the Law* in 1982, and *Psychology, Public Policy, and Law* in 1995. Great Britain followed suit with *Criminal Behavior and Mental Health* (launched in 1990), *Psychology, Crime, & Law* (1994), the British Psychological Society's *Legal and Criminological Psychology* (1996), and the *Journal of Forensic Psychology Practice* (2001). In addition to these, other interdisciplinary scholarly and scientific journals relevant to forensic psychology have emerged in recent years (e.g., *Journal of Forensic Sciences*, *American Journal of Forensic Psychiatry*, *Journal of Psychiatry and Law*).

During the 1970s, interdisciplinary and specialized training in forensic psychology was introduced at the doctoral, master's, internship, postdoctoral, and continuing education levels (Ogloff et al., 1996; see also Krauss & Sales, Chapter 5 this volume). The first interdisciplinary, successful psychology and law program was developed by Bruce Sales at the University of Nebraska–Lincoln in 1974 (Ogloff et al., 1996). Other universities soon followed in this endeavor, some more successfully than others. In the late 20th century and into 21st, thoughtful articles addressing the content of education and training programs in forensic psychology have been published (e.g., DeMatteo, Marczyk, Krauss, & Burl, 2009; Helmus, Babchishin, Camilleri, & Olver, 2011; Ogloff et al., 1996).

Another indication of the growth in forensic psychology is professional certification of practitioners in the field, a development that began in the late 1970s. Beginning in 1978, board certification in forensic psychology was provided by the American Board of Forensic Psychology (Otto & Heilbrun, 2002). In recent years, other board certifications have emerged, such as the American Board of Forensic Examiners. In 2001, as noted earlier, the APA voted to recognize forensic psychology as a specialty, and Specialty Guidelines for Forensic Psychologists and Psychology were adopted in 1991 and 2011, respectively. Forensic psychology has seen a rapid expansion in other parts of the globe besides North America, particularly in Europe and Australia. Blackburn (1996), in the first issue of *Legal and Criminological Psychology*, asserted, "The growth in the number of forensic psychologists has been among the most prominent developments in the burgeoning application of psychology to law during the last two decades" (p. 3). He noted that, although the growth was most apparent in the United States, there was a parallel growth throughout Europe in the latter part of the 20th century.

After an uncertain beginning and some stagnation between the two world wars, forensic psychology is now well established. Despite some continuing concerns about its definition (should it be broad or narrow?), it is importantly clinical in nature but also critically dependent on theory and research. All indicators suggest that forensic psychology has an extremely promising future as we continue into the 21st century. In the following chapters, other contributors assess forensic psychology's current status and the promise it holds for a future generation of researchers, practicing psychologists, theorists, and legal practitioners.

REFERENCES

- American Psychological Association. (2013). Specialty guidelines for forensic psychology. *American Psychologist, 68*(1), 7–19. doi:10.1037/a0029889
- Bandura, A. (1973). *Aggression: A social learning analysis*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A., & Walters, R. H. (1959). *Adolescent aggression*. New York, NY: Ronald Press.
- Barland, G. H. (1988). The polygraph test in the U.S.A. and elsewhere. In A. Gale (Ed.), *The polygraph test: Lies, truth and science* (pp. 73–96). London, England: Sage.
- Barnes, H. E., & Teeters, N. K. (1959). *New horizons in criminology* (2nd ed.). New York, NY: Prentice-Hall.
- Baron, R. A. (1977). *Human aggression*. New York, NY: Plenum Press.
- Bartol, C. R., & Bartol, A. M. (2004). *Introduction to forensic psychology*. Thousand Oaks, CA: Sage.
- Bartol, C. R., & Freeman, N. (2005). History of the American Association for Correctional Psychology. *Criminal Justice and Behavior, 32*, 123–142.
- Benjamin, L. T. (2003). Behavioral science and the Nobel prize: A history. *American Psychologist, 58*, 731–741.
- Berkowitz, L. (1962). *Aggression: A social-psychological analysis*. New York, NY: McGraw-Hill.
- Binet, A. (1900). *La suggestibilité* [Suggestibility]. Paris, France: Schleicher.
- Binet, A. (1905). La science du termoignage [The science of testimony]. *L'Année Psychologique, 11*, 128–137.
- Binet, A., & Clarparedo, E. (1906). La psychologie judiciaire [Legal psychology]. *L'Année Psychologique, 12*, 274–302.
- Blackburn, R. (1996). What is forensic psychology? *Legal and Criminological Psychology, 1*, 3–16.
- Bodemar, O. A. (1956). Correctional psychology in Wisconsin. *Journal of Correctional Psychology, 1*, 7–15.
- Bolton, F. E. (1896). The accuracy of recollection and observation. *Psychological Review, 3*, 286–295.
- Bornstein, B. H., & Penrod, S. D. (2008). Hugo Who? G. F. Arnold's alternative early approach to psychology and law. *Applied Cognitive Psychology, 22*, 759–768.
- Brigham, J. C. (1999). What is forensic psychology anyway? *Law and Human Behavior, 23*, 273–298.
- Brigham, J. C., & Grisso, J. T. (2003). Forensic psychology. In D. K. Freedheim (Ed.), *History of psychology* (pp. 391–411). Vol. 1 in I. B. Weiner (Ed.-in-Chief), *Handbook of psychology*. Hoboken, NJ: Wiley.
- Britt, S. H. (1940). The rules of evidence: An empirical study in psychology and law. *Cornell Law Quarterly, 25*, 556–580.
- Brodsky, S. L. (1973). *Psychologists in the criminal justice system*. Urbana, IL: University of Illinois Press.
- Brodsky, S. L. (2007). Correctional psychology and the American Association of Correctional Psychology: A revisionist history. *Criminal Justice and Behavior, 34*, 862–869.
- Brown, M. (1926). *Legal psychology*. Indianapolis, IN: Bobbs-Merrill.
- Brown v. Board of Education, 347 U.S. 483 (1954).
- Bryan, A. I., & Boring, E. G. (1946). Women in American psychology: Statistics from the OPP questionnaire. *American Psychologist, 1*, 71–79.

- Burt, H. E. (1931). *Legal psychology*. New York, NY: Prentice-Hall.
- Cady, H. M. (1924). On the psychology of testimony. *American Journal of Psychology*, *35*, 110–112.
- Cairns, H. (1935). *Law and the social sciences*. New York, NY: Harcourt, Brace.
- Cattell, J. M. (1895). Measurements of the accuracy of recollection. *Science*, *2*, 761–766.
- Cleckley, H. (1941/1964). *The mask of sanity*. Saint Louis, MO: C. V. Mosby.
- Clements, C. B. (1996). Offender classification: Two decades of progress. *Criminal Justice and Behavior*, *23*, 121–143.
- Coca-Cola Company v. Chero-Cola Company. 273 Fed. 755 (App. D.C. 1921).
- Comment. (1979). The psychologist as expert witness: Science in the courtroom? *Maryland Law Review*, *38*, 539–615.
- Committee on Ethical Guidelines for Forensic Psychologists. (1991). Specialty guidelines for forensic psychologists. *Law and Human Behavior*, *15*, 655–665.
- Corsini, R. (1945). Functions of the prison psychologist. *Journal of Consulting Psychology*, *9*, 101–104.
- Darley, J. G., & Berdie, R. (1940). The fields of applied psychology. *Journal of Consulting Psychology*, *4*, 41–52.
- DeMatteo, D., Marczyk, G., Krauss, D. A., & Burl, J. (2009). Educational and training models in forensic psychology. *Training and Education in Professional Psychology*, *3*, 184–201.
- Eysenck, H. J. (1964). *Crime and personality*. London, England: Routledge & Kegan Paul.
- Frye v. United States, 54 App. D.C. 46, 47 293 F. 1013, 1014 (1923).
- Giardini, G. I. (1942). The place of psychology in penal and correctional institutions. *Federal Probation*, *6*, 29–33.
- Glaser, D. (1964). *The effectiveness of a prison and parole system*. Indianapolis, IN: Bobbs-Merrill.
- Goddard, H. H. (1914). *Feeble-mindedness: Its causes and consequences*. New York, NY: Macmillan.
- Gottesman, J. (1975). *The utility of the MMPI in assessing the personality patterns of urban police applicants*. Hoboken, NJ: Stevens Institute of Technology.
- Greenberg, J. (1956). Social scientists take the stand: A review and appraisal of their testimony in litigation. *Michigan Law Review*, *54*, 953–970.
- Gross, H. (1897/1905). *Kriminalpsychologie* [Criminal psychology]. Graz: Teuschner & Lubensky.
- Hale, M. (1980). *Human science and social order: Hugo Münsterberg and origins of applied psychology*. Philadelphia, PA: Temple University Press.
- Hare, R. D. (1970). *Psychopathy: Theory and research*. New York, NY: Wiley.
- Heilbrun, K., & Brooks, S. (2010). Forensic psychology and forensic sciences: A proposed agenda for the next decade. *Psychology, Public Policy, and Law*, *16*, 219–253.
- Helmus, L., Babchishin, K. M., Camilleri, J. A., & Olver, M. E. (2011). Forensic psychology opportunities in Canadian graduate programs: An update of Simourd and Wormith's (1995) survey. *Canadian Psychology*, *52*, 122–127.
- Hidden v. Mutual Life Insurance Co., 217 F.2d 818 (4th Cir. 1954).
- Horowitz, I. A., & Willging, T. E. (1984). *The psychology of law: Integration and applications*. Boston, MA: Little, Brown.
- Hutchins, R. M., & Slesinger, D. (1927). Some observations on the law of evidence: Consciousness of guilt. *University of Pennsylvania Law Review*, *77*, 725–740.
- Hutchins, R. M., & Slesinger, D. (1928a). Some observations on the law of evidence: The competency of witnesses. *Yale Law Journal*, *37*, 1017–1028.

- Hutchins, R. M., & Slesinger, D. (1928b). Some observations on the law of evidence: Memory. *Harvard Law Review*, 41, 860–873.
- Hutchins, R. M., & Slesinger, D. (1928c). Some observations on the law of evidence: Spontaneous exclamations. *Columbia Law Review*, 28, 432–440.
- Hutchins, R. M., & Slesinger, D. (1929). Legal psychology. *Psychological Review*, 36, 13–26.
- Jackson, J. P. (2000). The triumph of the segregationists? A historiographical inquiry into psychology and the *Brown* litigation. *History of Psychology*, 3, 239–261.
- Jenkins v. United States, 307 F.2d 637 (D.C. Cir. 1962) *en banc*.
- Jesness, C. (1971). Jesness Inventory Classification System. *Criminal Justice and Behavior*, 15, 78–91.
- Keller, E. I. (1918). Psychopathic laboratory at police headquarters, New York City. *Journal of Applied Psychology*, 2, 84–88.
- Kelves, D. J. (1984, October 15). Annals of eugenics, II. *New Yorker*, 52–125.
- Kluger, R. (1975). *Simple justice*. New York, NY: Knopf.
- Kuna, D. P. (1978). One-sided portrayal of Münsterberg. *American Psychologist*, 33, 700.
- Landy, F. J. (1992). Hugo Münsterberg: Victim or visionary? *Journal of Applied Psychology*, 77, 787–802.
- Lindner, H. (1955). The work of court and prison psychologists. In G. J. Dudycha (Ed.), *Psychology for law enforcement officers* (pp. 131–150) Springfield, IL: Charles C Thomas.
- Lipmann, O. (1908). *Grundriss der psychologie für juristen* [Outline of psychology for jurists]. Leipzig, Germany: J. A. Bath.
- Lipmann, O. (1911). Pedagogical psychology of report. *Journal of Educational Psychology*, 2, 253–261.
- Loh, W. D. (1981). Perspectives on psychology and law. *Journal of Applied Social Psychology*, 11, 314–355.
- Louisell, D. W. (1955). The psychologist in today's legal world: Part I. *Minnesota Law Review*, 39, 235–260.
- Louisell, D. W. (1957). The psychologist in today's legal world: Part II. *Minnesota Law Review*, 41, 731–750.
- Magaletta, P. R., Patry, M. W., Dietz, E. F., & Ax, R. F. (2007). What is correctional about clinical practice in corrections? *Criminal Justice and Behavior*, 34, 7–21.
- Marbe, K. (1936). Work in commercial high schools and practical attitude. In C. Murchison (Ed.), *A history of psychology in autobiography*, Vol. 3 (pp. 181–213). Worcester, MA: Clark University Press.
- Marston, W. M. (1917). Systolic blood pressure changes in deception. *Journal of Experimental Psychology*, 2, 117–163.
- Marston, W. M. (1920). Reaction-time symptoms of deception. *Journal of Experimental Psychology*, 3, 72–87.
- Marston, W. M. (1921). Psychological possibilities in deception tests. *Journal of the American Institute of Criminal Law and Criminology*, 11, 551–570.
- Marston, W. M. (1924). Studies in testimony. *Journal of Criminal Law and Criminology*, 15, 5–32.
- Marston, W. M. (1925). Negative type reaction-time symptoms of deception. *Psychological Review*, 32, 241–247.
- McCary, J. L. (1956). The psychologist as an expert witness in court. *American Psychologist*, 11, 8–13.
- McCarty, D. G. (1929). *Psychology for the lawyer*. New York, NY: Prentice-Hall.

- Megargee, E. I. (1966). Undercontrolled and overcontrolled personality types in extreme antisocial aggression. *Psychological Monographs*, 80(3).
- Megargee, E. I. (1977). A new classification system for criminal offenders. *Criminal Justice and Behavior*, 4, 107–114.
- Merrill, M. A. (1927). Intelligence of policemen. *Journal of Personnel Research*, 5, 511–515.
- Miller, H. L., Lower, J. S., & Bleachmore, J. (1978). The clinical psychologist as an expert witness's questions of mental illness and competency. *Law and Psychology Review*, 4, 115–125.
- Monahan, J., & Loftus, E. F. (1982). The psychology of law. *Annual Review of Psychology*, 33, 441–475.
- Monahan, J., & Walker, L. (1988). Social science research in law: A new paradigm. *American Psychologist*, 43, 465–472.
- Moore, C. C. (1907). Yellow psychology. *Law Notes*, 11, 125–127.
- Mülberger, A. (2009). Teaching psychology to jurists: Initiatives and reactions prior to World War I. *History of Psychology*, 12, 60–86.
- Muller v. Oregon, 208 U.S. 412 (1908).
- Münsterberg, H. (1899/1994). Psychology and history. *Psychological Review*, 101, 230–236.
- Münsterberg, H. (1908). *On the witness stand: Essays on psychology and crime*. New York, NY: McClure.
- Münsterberg, H. (1914). The mind of the jurymen. In H. Münsterberg (Ed.), *Psychology and social sanity* (pp. 229–249). New York, NY: Doubleday, Page.
- Napoli, D. S. (1981). *Architects of adjustment*. Port Washington, NY: Kennikat.
- Nietzel, M. (2000). Police psychology. In A. E. Kazdin (Ed.), *Encyclopedia of psychology*, Vol. 6 (pp. 223–226). Washington, DC: American Psychological Association.
- Odom v. State, 174 Ala. 4, 7, 56 So. 913, 914 (1911).
- Oglesby, T. W. (1957). Use of emotional screening in the selection of police applicants. *Public Personnel Review*, 18, 228–231, 235.
- Ogloff, J. R. P., Tomkins, A. J., & Bersoff, D. N. (1996). Education and training in psychology and law/criminal justice. *Criminal Justice and Behavior*, 23, 200–235.
- Otto, R. K., & Heilbrun, K. (2002). The practice of forensic psychology: A look toward the future in light of the past. *American Psychologist*, 57, 5–19.
- Paterson, D. G. (1940). Applied psychology comes of age. *Journal of Consulting Psychology*, 4, 1–9.
- People v. Hawthorne, 293 Mich. 15, 291 N.W. 205 (1940).
- President's Commission on Law Enforcement and the Administration of Justice. (1967). *Task force report: The police*. Washington, DC: U.S. Government Printing Office.
- Quay, H. C. (1965). Psychopathic personality: Pathological stimulation-seeking. *American Journal of Psychiatry*, 122, 180–183.
- Rankin, J. H. (1957). Preventive psychiatry in the Los Angeles Police Department. *Police*, 2, 24–29.
- Rankin, J. H. (1959). Psychiatric screening of police recruits. *Public Personnel Review*, 20, 191–196.
- Reese, J. T. (1986). Foreword. In J. T. Reese & H. Goldstein (Eds.), *Psychological services for law enforcement* (p. v). Washington, DC: U.S. Government Printing Office.
- Reese, J. T. (1987). *A history of police psychological services*. Washington, DC: U.S. Government Printing Office.

- Reiser, M. (1972). *The police psychologist*. Springfield, IL: Charles C Thomas.
- Reiser, M. (1982). *Police psychology: Collected papers*. Los Angeles, CA: LEHI.
- Resnick, R. J. (1997). A brief history of practice—expanded. *American Psychologist*, 52, 463–468.
- Robinson, E. S. (1935). *Law and the lawyers*. New York, NY: Macmillan.
- Rogers, O. (1910). The unwary purchaser: A study in the psychology of trademark infringement. *Michigan Law Review*, 8, 613–644.
- Rogers, O. (1918). An account of some psychological experiments on the subject of trademark infringements. *Michigan Law Review*, 18, 75–95.
- Rowland, E. (1913). Report of experiments at the State Reformatory for Women at Bedford, New York. *Psychological Review*, 20, 245–249.
- Schrenck-Notzing, A. (1897). *Über suggestion und erinnerungsfälschung im Berchthold-process* [Concerning suggestion and falsification of memory in the Berchtold case]. Leipzig, Germany: Johann Ambrosius Barth.
- Scrivner, E. M. (1994). *The role of police psychology in controlling excessive force*. Washington, DC: National Institute of Justice.
- Shartle, C. L. (1946). Occupations in psychology. *American Psychologist*, 1, 559–582.
- Slesinger, D., & Pilpel, M. E. (1929). Legal psychology: A bibliography and a suggestion. *Psychological Bulletin*, 12, 677–692.
- Snee, T. J., & Lush, D. E. (1941). Interaction of the narrative and interrogatory methods of obtaining testimony. *Journal of Psychology*, 11, 225–236.
- Sporer, S. L. (1982). A brief history of the psychology of testimony. *Current Psychological Reviews*, 2, 323–339.
- Sporer, S. L. (2008). Lessons from the origins of eyewitness testimony research in Europe. *Applied Cognitive Psychology*, 22, 737–757.
- Sprung, L., & Sprung, H. (2001). History of modern psychology in Germany in 19th- and 20th-century thought and society. *International Journal of Psychology*, 36, 364–376.
- State v. Driver, 88 W. Va. 479, 107 S. E. 189 (1921).
- Stern, L. W. (1902). Zur psychologie der aussage [To the psychology of testimony]. *Zeitschrift für die gesamte Strafrechtswissenschaft*, 23, 56–66.
- Stern, L. W. (1910). Abstracts of lectures on the psychology of testimony. *American Journal of Psychology*, 21, 273–282.
- Stern, L. W. (1939). The psychology of testimony. *Journal of Abnormal and Social Psychology*, 40, 3–20.
- Tapp, J. L. (1976). Psychology and the law: An overture. *Annual Review of Psychology*, 27, 359–404.
- Terman, L. M. (1917). A trial of mental and pedagogical tests in a civil service examination for policemen and firemen. *Journal of Applied Psychology*, 1, 17–29.
- Terrio, L., Swanson, C. R. Jr., & Chambelin, N. C. (1977). *The police personnel selection process*. Indianapolis, IN: Bobbs-Merrill.
- Thurstone, L. L. (1922). The intelligence of policemen. *Journal of Personnel Research*, 1, 64–74.
- Toch, H. (Ed.). (1961). *Legal and criminal psychology*. New York, NY: Holt, Rinehart & Winston.
- Toch, H. (1969). *Violent men: An inquiry into the psychology of violence*. Chicago, IL: Aldine.
- Toch, H. (1992). Functional unit management: An unsung achievement. *Federal Prisons Journal*, 2, 15–19.

- Undeutsch, U. (1992). Highlights of the history of forensic psychology in Germany. In F. Losel, D. Bender, & T. Bliesener (Eds.), *Psychology and law: International perspectives* (pp. 509–518). Berlin, Germany: De Gruyter.
- Viteles, M. S. (1929). Psychological methods in the selection of patrolmen in Europe. *Annals of the American Academy*, 146, 160–165.
- Watkins, R. E. (1992). *An historical review of the role and practice of psychology in the field of corrections*. Ottawa, Canada: Correctional Service of Canada.
- Weld, H. P., & Danzig, E. R. (1940). A study of the way in which a verdict is reached by a jury. *American Journal of Psychology*, 53, 518–536.
- Weld, H. P., & Roff, M. (1938). A study in the formation of opinion based upon legal evidence. *American Journal of Psychology*, 51, 609–628.
- Whipple, G. M. (1909). The observer as reporter: A survey of the “psychology of testimony.” *Psychological Bulletin*, 6, 153–170.
- Whipple, G. M. (1910). Recent literature on the psychology of testimony. *Psychological Bulletin*, 7, 365–368.
- Whipple, G. M. (1911). Psychology of testimony. *Psychological Bulletin*, 8, 307–309.
- Whipple, G. M. (1912). Psychology of testimony and report. *Psychological Bulletin*, 9, 264–269.
- Whipple, G. M. (1913). Psychology of testimony and report. *Psychological Bulletin*, 10, 264–268.
- Whipple, G. M. (1914). Psychology of testimony and report. *Psychological Bulletin*, 11, 245–250.
- Whipple, G. M. (1915). Psychology of testimony. *Psychological Bulletin*, 12, 221–224.
- Whipple, G. M. (1917). Psychology of testimony. *Psychological Bulletin*, 14, 234–236.
- Wigmore, J. H. (1909). Professor Münsterberg and the psychology of testimony: Being a report of the case of *Cokestone v. Münsterberg*. *Illinois Law Review*, 3, 399–445.
- Wigmore, J. H. (1940). *Evidence in trials at common law*. Boston, MA: Little, Brown.
- Winick, C. (1961). The psychology of juries. In H. Toch (Ed.), *Legal and criminal psychology* (pp. 96–120). New York, NY: Holt, Rinehart & Winston.
- Wormith, J. S., & Luong, D. (2007). Legal and psychological contributions to the development of corrections in Canada. In R. K. Ax & T. J. Fagan (Eds.), *Corrections, mental health, and social policy: International perspectives* (pp. 129–173). Springfield, IL: Charles C Thomas.

CHAPTER 2

Defining Forensic Psychology

RANDY K. OTTO AND JAMES R. P. OGLOFF

THERE IS NO CONSENSUAL DEFINITION OF FORENSIC PSYCHOLOGY

PERHAPS it is surprising, given the relatively long history and growth of forensic psychology over the past 40 years, that there is no uniform or consensual definition for this specialty area, and most differences involve how narrowly or broadly the field is defined. Definitions range from expansive ones—that include any application of psychology to any legal matters—to those that are narrower and typically are limited to clinical and counseling psychologists' involvement in legal matters as examiners, treatment providers, or consultants. Examples of more expansive definitions include those offered by Huss (2009), who defined forensic psychology as “any application of psychology to the legal system” (p. 5); the American Psychological Association (APA; 2013), which in its Specialty Guidelines for Forensic Psychology (reprinted as the appendix to this volume with permission of the APA) indicated that “forensic psychology refers to professional practice by any psychologist working within any sub-discipline of psychology (e.g., clinical, developmental, social, cognitive) when applying the scientific, technical, or specialized knowledge of psychology to the law to assist in addressing legal, contractual, and administrative matters” (p. 7); and the American Board of Forensic Psychology, which described forensic psychology as “the application of the science and profession of psychology to questions and issues relating to law and the legal system” (www.abfp.com). In contrast, in its petition to the APA's Committee for the Recognition of Specialties and Proficiencies in Professional Psychology to establish forensic psychology as a psychological specialty, the Forensic Specialty Council (which comprises representatives from the American Psychology–Law Society [Division 41 of APA, hereinafter AP-LS], the American Board of Forensic Psychology, and the American Academy of Forensic Psychology) offered a more narrow definition: “Forensic psychology is the professional practice by psychologists within the areas of clinical psychology, counseling psychology, school psychology or another specialty recognized by

the American Psychological Association, when they are engaged as experts and represent themselves as such, in an activity primarily intended to provide professional psychological expertise to the judicial system" (2008; also see www.apa.org/ed/graduate/specialize/forensic.aspx).

Brigham (1999) observed that these differences are more than semantic, and the varying definitions have distinct advantages and disadvantages. Whereas adopting a broad definition, according to Brigham, could promote growth and coherence in the field and facilitate development of graduate training programs, he acknowledged that grouping clinical psychologists with nonclinicians (e.g., developmental, social, and experimental psychologists) could prove complicated and perhaps problematic given the very different training and licensure requirements that often apply. Indeed, Brigham reported that disagreement among members of AP-LS about how broad or narrow the definition should be initially led the group to abandon a cooperative effort with the American Academy of Forensic Psychology to jointly sponsor a petition to have APA formally recognize forensic psychology as a specialty.

Adding to the confusion, how the terms *forensic psychology* or *psychology and law* are defined differs internationally. Conventionally, the entire field—including both clinical or applied areas and research areas—has been referred to as psychology and law or law and psychology in North America and continental Europe, whereas the term *forensic psychology* has been more commonly employed in the United Kingdom, Australia, and New Zealand. In countries that use the term *forensic psychology* in an expansive way, those who work in the applied areas of psychology and the law typically are referred to as clinical forensic psychologists (Ogloff, 2011).

Regardless of the differences that exist concerning the definition of forensic psychology, it is uniformly agreed that forensic psychology involves the application of psychological knowledge and expertise to the legal system. To this end, forensic psychologists work at the interface of psychology and the law. As forensic psychologists work with legal actors, including attorneys, judges, and others in the justice system, a number of tensions exist. Although commentators have characterized these tensions differently (e.g., Brockman & Rose, 2011; Haney, 1980; Melton et al., 2007), a number of common themes emerge. Drawing on the framework provided by Haney (1980), we present eight differences between psychology and law that may contribute to tensions between the disciplines.

1. The emphasis in law is *stare decisis* (i.e., legal precedent), whereas in psychology the emphasis is on *creativity*. In the law, past cases and matters such as constitutional interpretation rather than innovation or creativity are painstakingly relied on for the development of legal arguments. The model adopted in law is one of legal precedent. In contrast, in psychology, the model is one of innovation, and psychologists, in both research and applied work, are encouraged continually to explore new ideas and methods.

2. Law is hierarchical whereas psychology is empirical. Decisions within the legal system are hierarchically based and authoritative, with lower courts bound by the decisions of higher courts. In psychology, however, it is the accumulation of a body of consistent and supporting data that confirms the validity of a particular position or claim, not its authoritative declaration.
3. Law relies on the *adversarial* method, whereas psychology relies on the *experimental* method. The law seeks “justice,” which equates to procedural fairness. It is hoped that just procedures will assist in arriving at the truth; however, knowing that the truth is elusive, it is seen as more important to ensure that the principles of due process are followed. To arrive at the “truth” in law, conflicting points of view are presented within the strict parameters of a trial or appellate hearing, with each side putting its best case forward. Bias, self-interest, and advocacy are not only permitted but heralded as one of the strengths of the process. Indeed, what is of immediate concern and the driving force for the opposing lawyers is victory. Psychology, in contrast to law, attempts to arrive at “truth” (i.e., an understanding of some phenomenon) using a variety of diverse data-gathering methods. Common to all of these methods is the systematic collection of data, using procedures that attempt to “reduce bias, error and distortion in observation and inferences” (Haney, 1980, p. 162). Although this is not to say that bias does not enter into the research process, the goal of the psychologist is to attain an “objective” understanding of the phenomena rather than victory over a particular viewpoint.
4. Law is *prescriptive*, whereas psychology is *descriptive*. The law is primarily prescriptive, telling “people how they should behave,” whereas psychology is “essentially a descriptive discipline, seeking to describe behavior as it actually occurs” (Haney, 1980, p. 163). This dimension captures a difference in the values espoused in the disciplines, with law outlining how one ought to behave and psychology adopting a more nonjudgmental orientation of how people do behave.
5. Law is *idiographic*, whereas psychology is *nomothetic*. Law operates on a case-by-case basis, with each case decided on the basis of its specific facts. In contrast, psychology is interested in uncovering the general principles, relationships, and patterns that govern human behavior. The focus in psychology is not on a particular instance but rather on what transcends the particular instance.
6. Decision making in law is based on the *appearance of certainty*, whereas decision making in psychology is based on *probabilistic evidence*. The decisions made in the law typically take on a dichotomous, all-or-nothing quality—the accused in a criminal trial is deemed either guilty or not guilty, the defendant in a civil case is found liable or not liable. Psychologists, in contrast, operate in terms of probabilities; for example, claims are asserted on the basis of evidence associated with a probability level (i.e., level of statistical significance). As a result, conclusions drawn by psychologists typically are qualified and not categorical by nature.

7. Law is *reactive*, whereas psychology is *proactive*. The issues that arise in the law originate from outside the system, namely, cases are brought to the attention of lawyers. In contrast, psychologists, notwithstanding the presence of various external pressures (e.g., funding availability and the pressure to publish), have considerable control over the issues they study.
8. Law is *operational*, whereas psychology is *academic*. Law is an applied discipline, and it is designed to deal with real-world problems. The players within the system (e.g., lawyers, parole officers, etc.) have clearly defined roles that prescribe the issues on which they will concentrate. In contrast, similar to the distinction noted previously, psychologists have considerable say over the issues they pursue. The driving force tends to be more of a quest for knowledge for its own sake (i.e., for academic reasons rather than for purely pragmatic reasons).

With these eight tensions in mind, included in what follows is an expansive survey of what can be characterized as forensic psychology. At a macro level, forensic psychologists can assist the legal system in four ways: (1) providing legal decision makers with information about complicated matters that they would not otherwise have, (2) assisting specific legal actors, (3) researching the legal system and its operation, and (4) researching psychological phenomena that are of particular interest or relevance to the legal system.

PSYCHOLOGISTS ASSISTING THE LEGAL SYSTEM

Psychologists assist the legal system in a number of ways including providing expert testimony in legal, administrative, and legislative proceedings; conducting and testifying about research conducted in anticipation of litigation; testifying about research not conducted in connection with litigation but that is nonetheless relevant; and researching the legal system's operation.

INFORMING DECISION MAKERS IN LEGAL, ADMINISTRATIVE, AND LEGISLATIVE PROCEEDINGS—EXPERT TESTIMONY

Every day in the United States and other countries, thousands of psychologists appear before and provide expert opinions to courts, administrative proceedings (e.g., parole boards), attorneys, and legislative hearings via reports and/or sworn testimony. Psychologists' involvement in these matters is predicated on the assumption that their observations and opinions will educate the recipient (e.g., attorney, judge, jury) about some psychological phenomena that are relevant to the legal matters in dispute, and, as a result, a more accurate and presumably better decision will be made.

Psychologists' participation in this way is based on the premise that, in some cases, legal decision makers must consider complicated psychological matters that

are beyond their understanding, and their judgments would benefit from the insights and opinions of someone with specialized knowledge about the matter, such as a psychologist. Indeed, Federal Rule of Evidence 702 makes clear under what circumstances psychologists (and any other proffered experts) are permitted to testify:

A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if: (a) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue; (b) the testimony is based on sufficient facts or data; (c) the testimony is the product of reliable principles and methods; and (d) the expert has reliably applied the principles and methods to the facts of the case.

CONDUCTING AND TESTIFYING ABOUT FORENSIC PSYCHOLOGICAL EVALUATIONS

Most frequently, psychologists provide assistance to the legal system by offering observations and opinions about the emotional, behavioral, and/or cognitive functioning of someone whose mental state is at issue in a legal proceeding. These activities are quite varied and range from involvement in criminal proceedings in which the competence of a defendant to stand trial is in dispute to civil proceedings in which the court is faced with making decisions about what type of living and legal arrangements would be in a child's best interests in the case of divorcing parents who cannot reach an agreement about these matters. In these cases, psychologists are offering observations and opinions about a specific person (or persons) who has been evaluated, and these observations and opinions are considered helpful to the court.

Because a defendant's, litigant's, or other person's emotional, behavioral, or cognitive functioning can be at issue in a variety of legal proceedings, psychologists find themselves evaluating and testifying about the functioning of persons in many different legal matters (Melton et al., 2007). In criminal proceedings, psychologists may evaluate defendants when there are questions about their competence to proceed (stand trial), criminal responsibility (sanity), and/or treatment needs for consideration at the time of sentencing. In civil proceedings, psychologists can evaluate litigants and others when there are questions about their capacity to parent, manage their personal and financial affairs, execute a will, work, or make health-care decisions, or when there are disputes about their emotional, behavioral, or cognitive functioning as it relates to their risk for harming themselves or others or an alleged wrong committed by a third party. In all of these cases, psychologists assess individuals in light of the parameters the law has established for the particular question being addressed (e.g., standards of the insanity defense) in order to gather data and offer opinions that will be presented to the decision makers to assist them in reaching a more informed decision in the matter (Grisso, 1986, 2003).

CONDUCTING AND TESTIFYING ABOUT LITIGATION—SPECIFIC RESEARCH

In some cases, psychologists can be of assistance to attorneys and the court by conducting and presenting the results of research that addresses important points of contention in litigation. Although the kind of questions that are asked and the type of research that is conducted can vary dramatically, common to this work is that the research was conducted in the context of the litigation at hand and designed to answer some case-specific questions.

In their influential treatise, Walker and Monahan (1987) described the products of this research as “social fact” testimony. For example, often in dispute in patent and trademark litigation is the issue of consumer confusion about competing products. In these cases, such as recent legal disputes regarding the Apple and Samsung wireless phones, one company alleges that another is unfairly encroaching on a unique aspect of its product that (1) results in consumer confusion and/or (2) harms the company economically. In *Zippo Manufacturing v. Rogers Imports* (1963), the Zippo lighter company argued that a competitor—Rogers Imports—copied the design of its cigarette lighter in such a way that it violated its trademark, caused consumers to confuse the two brands, and resulted in decreased Zippo sales and income. In support of its claim, Zippo presented results of research conducted expressly for purposes of the litigation, in which participants who were presented with both lighters demonstrated confusion about the brands. The results and interpretation of these data were offered by Zippo Manufacturing in support of its ultimately successful claim that consumer confusion resulted from the trademarked design similarities.

Psychologists also conduct research to inform legal decision makers about the knowledge and attitudes members of a jury pool have about a case that is about to go to trial, so as to inform decisions about whether an unbiased jury can be empaneled or whether a change in venue is necessary (i.e., if the location of the trial must be moved to another community). Such jury venire research, which is conducted in both criminal and civil proceedings, involves developing and administering surveys (typically by way of telephone calls) that query potential members of the jury pool about their knowledge and attitudes concerning the case at hand (Posey & Dahl, 2002). Psychologists who conduct this research may be called to testify about their findings, thereby providing the court with information about the jury pool that it would not otherwise have, to allow it to make a more informed decision about potential jurors’ knowledge of case matters and the potential need for a change in venue.

TESTIFYING ABOUT PSYCHOLOGICAL RESEARCH NOT CONDUCTED
IN ANTICIPATION OF LITIGATION

Psychologists can also assist legal decision makers and legislative bodies by providing observations and expert opinions about more general matters that are nonetheless of interest to the court or legislature. Walker and Monahan (1987) referred to this as “social authority” testimony.

In litigation contexts, psychologists typically testify about research that sheds light on some matter or an assumption that is relevant to the case at hand, but they typically do not offer opinions about specific case matters. For example, a psychologist knowledgeable about research regarding eyewitnesses might be called to testify and educate a jury in a criminal proceeding about the poor relationship between eyewitness confidence and eyewitness accuracy, or how crime witnesses tend to focus on weapons that are brandished and pay less attention to the perpetrator and his or her appearance (Wells & Loftus, 2012; Ross, Tredoux, & Malpass, Chapter 17 this volume). Or, in a child abuse prosecution, a psychologist knowledgeable about sexual victimization of children might educate the jury about why child victims of sexual abuse do not always come forward immediately to report the abuse or identify the perpetrator (Bussey, Lee, & Grimbeck, 1993; Kuehnle & Connell, 2009; 2012).

Social authority testimony is also introduced in legislative hearings, to inform lawmakers about psychological phenomena that are relevant to pending legislation. Thus, a psychologist knowledgeable about limitations in how adolescents understand and exercise their constitutional rights while in custody might testify before a legislative body that is considering a law mandating that minors be appointed counsel in delinquency proceedings, whereas a psychologist knowledgeable about the relationship between watching violence on television and aggressive behavior of children might offer expert testimony to a body considering legislation limiting what is broadcast on television during daytime hours.

RESEARCHING THE LEGAL SYSTEM

Psychologists conduct research that examines the legal system and its operation, and their findings can provide direction that is of considerable value. This research is quite varied in nature and focuses on phenomena as disparate as the effectiveness of various legal interventions or programs (e.g., drug courts, mental health courts, juvenile courts, or boot camps; divorce education classes; offender rehabilitation programs; crisis intervention teams), the prevalence and characteristics of various phenomena (e.g., criminal offenses, the nature of criminal offenders or victims, vicarious trauma experienced by jurors), and the operation of the legal system more generally (e.g., court efficiency, behavior of legal decision makers). What unifies this type of research is its focus on understanding and improving the legal system and its potential to provide important information to those who fashion policy and make laws that shape the legal system and its operation.

As an example, in their program of research, Redlich and her colleagues (Redlich, Steadman, Monahan, Petrila, & Griffin, 2005; Redlich, Steadman, Monahan, Robbins, & Petrila, 2005) have examined the outcomes associated with special “mental health courts” designed to respond to criminal defendants whose involvement with the criminal justice system is related to their problems with severe and persistent mental illness. Peters and his colleagues, in a similar line of research, examined

the efficacy of courts devoted to managing the special challenges of offenders with substance abuse problems (Hiller, Belenko, Welsh, Zajac, & Peters, 2011; Peters, 2011; Peters & Belenko, 2011; Peters, Haas, & Hunt, 2002; Peters & Murrin, 2000; Peters & Young, 2011). Finally, in a very different line of research, Kovera and her colleagues examined the impact of expert testimony on the legal decision making of judges and jurors (e.g., Kovera, Levy, Borgida, & Penrod, 1994; Kovera & McAuliff, 2000; Kovera, McAuliff, & Hebert, 1999; McAuliff & Kovera, 2007, 2008).

As the information just presented demonstrates, the roles of forensic psychologists in these contexts can be broad and varied. At a micro level, their work can involve conducting and reporting the results of psychological assessments or conducting case-specific research, both with the intent of providing case-relevant knowledge to the legal decision maker that it would not otherwise have, so that more informed and better decisions are made. At the other end of the spectrum—the macro level—psychologists can help policy makers, legislators, and decision makers better understand the need for or potential impact of proposed legislation or the legal system and its operation. Common to their involvement in all matters, however, is the fact that psychologists are relying on their expertise to provide helpful information that would otherwise not be available to legal decision makers. We turn next to a discussion of the role that forensic psychologists play in assisting specific legal actors.

PSYCHOLOGISTS ASSISTING LEGAL ACTORS

Given that forensic psychology involves the application of psychology to the legal system, it is not surprising that much of the work of forensic psychologists involves assisting specific legal actors. In this section, we discuss the different ways in which forensic psychologists may assist law enforcement agencies, attorneys, litigants, and others.

ASSISTING LAW ENFORCEMENT

Within the area of policing and law enforcement, psychologists may play a variety of roles (Scrivner, Corey, & Greene, Chapter 15 this volume). A large body of research exists that establishes psychologists' potential to assess law enforcement officers in matters of investigation and interrogation (Bartol, 1996). Once a crime is reported, law enforcement officials conduct an investigation to establish whether a crime has in fact been committed, whether it can be solved, and whether they can obtain evidence to facilitate a prosecution. At the level of investigation, a number of popular books, television shows, and movies depict criminal profilers. James Brussel (1968), a psychiatrist who began consulting to the New York City Police Department in the 1950s, described the first case in which he was asked to assist the police. The "Mad Bomber of New York" detonated more than 20 bombs in theaters,

transportation terminals, libraries, and offices around New York City for 16 years during the 1940s and 1950s. Despite notes and letters mailed to them by the bomber, the police were at a loss to identify a suspect and eventually consulted Brussel, who examined the evidence the police had collected, including the notes, letters, and photographs and details of the crime scenes. Brussel developed a precise “criminal profile” of the bomber, which turned out to closely match the characteristics of the man the police eventually apprehended and prosecuted. Since that time, the field of criminal profiling has developed. Of course, in many cases, the efforts of psychologists and psychiatrists have not been so successful (Holloway, 2003; Porter, 1983), and there remains concern that criminal profiling is nothing more than so-called smoke and mirrors (Hicks & Sales, 2006; Snook, Cullen, Bennell, Taylor, & Gendreau, 2008).

In reality, most criminal profilers are police officers, but forensic psychologists sometimes are called on to assist with investigations (Douglas, Ressler, Burgess, & Hartman, 1986). Over time, investigative psychology and offender profiling have developed into an area of forensic psychology with an empirical base, and modern approaches to offender profiling are far removed from the early speculative approaches that are still so often depicted in television and film (Alison & Rainbow, 2011; Canter & Youngs, 2009).

Psychologists have also conducted research to investigate the efficacy of police interviews (McLean, 1995) and assist police with interviewing witnesses and suspects, including child victims and witnesses (Cronch, Viljoen, & Hansen, 2006; Wilson & Powell, 2001). This work assists police in developing interviewing skills for use with witnesses and suspects that will maximize the amount of accurate information that is obtained and minimize bias and error.

In addition to direct involvement with police with respect to conducting investigations and questioning witnesses and suspects, psychologists are involved in a range of other activities. Psychologists may be called on to assist the police in their interactions with persons with mental disorders (International Association of Chiefs of Police, 2010; Kesic, Thomas, & Ogloff, 2013; Ogloff et al., 2013). A great deal of work has also been done to assist law enforcement agencies with respect to screening, selection, and recruitment of police candidates (e.g., Corey & Borum, 2013; Craig, 2005) and providing critical mental health services to sworn officers and their families (e.g., see Scrivner, Corey, & Greene, Chapter 15 this volume). We now turn to a brief review of the roles psychologists play in assisting attorneys.

ASSISTING ATTORNEYS

Psychologists frequently provide consultation to attorneys with respect to case formulation and jury matters (Posey & Wrightsman, 2005). To this end, some psychologists assist attorneys in conceptualizing and presenting cases in a way that will be most advantageous to their clients. Moreover, a growing area of study focuses on the psychology of the jury, in which psychologists assist attorneys by

developing strategies for selecting and working with juries. Indeed, the area of psychological trial consulting and scientific jury selection has grown significantly over the past two decades.

Psychologists can assist attorneys by helping them conceptualize their case in a way that will be most compelling for the jury. Trial consultants argue that, because attorneys develop specialized legal knowledge, they may not be able to conceptualize cases or present them in a way that will be best understood by the jury. Research shows that jurors use a so-called story model to assist them to make sense of the facts presented at trial (Pennington & Hastie, 1986). According to this model, after hearing the evidence at trial and being provided the legal instructions by the judge at the end of the trial, jurors attempt to find the best match between the arguments made by the competing attorneys and the verdict options. To this end, it is important that attorneys conceptualize and explain the case (i.e., “tell the story”) in a way that the jury understands and that will best fit the verdict option that suits their clients. Relying on general decision-making research and surveys or questionnaires that may be developed for the case at hand, psychologists can assist attorneys by helping them understand how jurors may make sense of and consider the evidence and crafting their arguments accordingly (Brodsky, 2009). More recently, psychologists acting as trial consultants have begun to assist attorneys in presenting the information to the jury using modern technology to maximize the effectiveness of their arguments (e.g., PowerPoint presentations, computer simulations).

Beyond assisting attorneys in conceptualizing the case and presenting information to jurors in the most compelling manner, psychologists may assist attorneys with jury selection. In the United States, the jury is selected in a process known as the *voir dire*, during which potential jurors are questioned by the judge or attorneys in order to ensure that a fair and impartial jury is empaneled. As such, jurors may be “challenged for cause” in cases where it is determined that a juror has preexisting beliefs that would prevent him or her from making a decision in the case based solely on the evidence presented. In addition to challenges for cause, attorneys may challenge a designated number of prospective jurors in each case without providing a justification. (These are known as peremptory challenges.) With peremptory challenges, prospective jurors may be excluded “without a reason stated, without inquiry, and without being subject to the Court’s control” (*Swain v. Alabama*, 1965, p. 219).

When empaneling a jury, attorneys may rely on forensic psychologists to assist them in selecting the most desirable jurors (i.e., jurors most likely to be sympathetic to their claims/arguments and reach a verdict in their favor) and deselecting the least desirable jurors (i.e., jurors least likely to be sympathetic to their claims/arguments and reach a verdict in their favor). (For summaries, see Kovera, 2012, or Robbenolt, Groscup, & Penrod, Chapter 16 this volume.) Typically, psychologists survey members of the community about matters pertaining to a case in order to identify those characteristics that relate attitudes about the case and case outcomes

(e.g., older people may be less accepting of illegally downloading media from the Internet than younger people). Consultants may also use focus groups to obtain additional information about the views of people regarding matters at issue in the particular case and how prospective jurors may respond to different arguments. Based on the results of the surveys, psychologists can help identify the characteristics of people who would be more or less likely to make a decision in favor of their client. Then, during *voir dire*, potential jurors would be asked a series of questions designed to identify sympathetic and unsympathetic jurors. Research shows that, without assistance, attorneys may not be very skilled at identifying jurors who might be biased against their clients (Olczak, Kaplan, & Penrod, 1991). Although the empirical evidence shows an increased likelihood that jurors will find in favor of the side that employs jury consultants, the results vary across studies and in actual cases—particularly since both sides may use consultants (Kressel & Kressel, 2004; Posey & Wrightsman, 2005).

In addition to the strategies just outlined, psychologists who work as jury consultants sometimes employ mock or shadow juries (Brodsky, 2009). The use of mock juries involves bringing a group of jury eligible people together, presenting them with case information in order to determine how various arguments and presentations affect their thinking and decision making, and shaping the case presentation accordingly. The complexity and sophistication of mock juries can vary from providing the participants with a summary of information about the case to actually having attorneys present their arguments in a mock trial format. Finally, during the course of the trial, jury consultants sometimes employ “shadow jurors” who sit in the courtroom throughout the trial, listen to the arguments and evidence, and provide the consultant with their perceptions and opinions as the trial proceeds. The attorneys then use this information to shape their subsequent presentations and strategies (Brodsky, 2009).

ASSISTING LITIGANTS AND OTHERS

With respect to assisting litigants and others in the legal system, psychologists act in quasi-judicial capacities and also provide therapeutic services.

Quasi-Judicial Roles. Over the past quarter of a century, psychologists have become increasingly involved in a number of activities in which they serve as decision makers for persons involved in the legal process. Psychologists’ involvement in such activities is presumably predicated upon assumptions that their interpersonal skills provide them with abilities that will facilitate examination and settling of disputes. A number of these activities are presented next.

Mediation. In some jurisdictions, psychologists can serve as legally recognized mediators, in which they function in a quasi-judicial role. Mediation is a dispute resolution process that helps persons involved in legal disputes avoid the adversarial

process and courtroom litigation. Although there is considerable variability across jurisdictions and contexts, mediation at the most general level involves a neutral person (the mediator) who works with parties to a dispute in order to craft an agreement that is acceptable to them, with the understanding that a return to traditional litigation channels will occur if such an agreement cannot be reached.

Unlike many if not most litigants, parents in custody disputes must have continued contact with each other involving matters of their minor children after the court hearing their dispute has rendered a judgment. Thus, mediation proponents argue that it can be of particular value in divorce and custody proceedings because of its potential to diminish some of the acrimony and emotion that is associated with the adversarial process. Proponents of using mediation in cases of contested custody argue that it has the potential to facilitate settlement of a large number of cases headed for court, speed litigation times, decrease litigation costs, increase compliance with custody agreements, and improve family relationships, including the relationship that the divorcing or separating parents have with each other and their children (Association of Family and Conciliation Courts [AFCC], 2000; Emery, Sbarra, & Grover, 2005). Some, however, have questioned the value of mediation in matters of divorce and custody (see, e.g., Beck & Sales, 2001). In 2005, the AFCC—an interdisciplinary organization of attorneys, mental health professionals, social service professionals, and accountants—published the *Model Standards of Practice for Family and Divorce Mediation*, which serve as a guide for the conduct of family mediators, educate service recipients about the mediation process and what to expect, and promote public confidence in mediation as a family dispute resolution process.

Parent Coordination. Over the past 15 to 20 years, psychologists and other mental health professionals have taken on a new role in family court proceedings of parent coordinator or special master. According to the APA (2012), parent coordination is

a nonadversarial dispute resolution process that is court ordered or agreed on by divorced and separated parents who have an ongoing pattern of high conflict and/or litigation about their children...[and] is designed to help parents implement and comply with court orders or parenting plans, to make timely decisions in a manner consistent with children's developmental and psychological needs, to reduce the amount of damaging conflict between caretaking adults to which children are exposed, and to diminish the pattern of unnecessary relitigation about child-related issues. (p. 63)

Parent coordinators typically are appointed only in the most challenging cases involving divorced parents who experience enduring high conflict surrounding the caretaking of their children (AFCC, 2002; Johnston, Roseby, & Kuehnle, 2009). Parent coordinators generally have responsibility for resolving ongoing and day-to-day disputes that may develop (e.g., decision making and conflicts regarding education, health care, visitation, and social matters), while the court retains the right to rule on more significant matters (e.g., changes in parenting time, visitation and legal

decision-making authority, relocation issues). All commentators agree that serving as a parent coordinator is particularly challenging, given the level of conflict that is inherent to all cases, the hybrid role that is assumed by the professional, and the myriad regulatory and professional bodies that might consider the work of the professional (see, e.g., Coates, Deutsch, Starnes, Sullivan, & Sydlik, 2004; Kirkland & Sullivan, 2008; Sullivan, 2008).

Not surprisingly, the legal authority for parent coordinators, their rights, and their responsibilities varies across jurisdictions (Sullivan, 2008). Recently, however, at least two organizations have provided important direction to psychologists serving in this role by publishing practice guidelines. In 2005, the AFCC published the Guidelines for Parenting Coordination, the purpose of which is to provide direction to professionals, jurisdictions, and educational institutions regarding (1) appropriate parent coordinator practice; (2) the parent coordinator's ethical obligations; and (3) educational, training, and experience qualifications for parent coordinators. Similarly, in 2012, the APA published the Guidelines for the Practice of Parenting Coordination, the purpose of which is to "describe best practices for ethical and competent functioning in this unique role" (p. 64).

TREATMENT AND INTERVENTION

Much of the work of psychologists in the legal system involves treating those within it, including victims and offenders.

Crime Victims. By definition, virtually every crime has a victim. Research and clinical experience show that crime victims can experience a range of physical and psychological responses to the event, ranging from transient distress and discomfort to more enduring mental disorders, such as posttraumatic stress disorder (Karmen, 2010). Working with victims is a growth area within forensic psychology. Indeed, all states in the United States have enacted crime victim legislation, most of which provides for funding of mental health treatment services.

Despite the stark reality of the large number of crime victims at any point in time, with the exception of work focusing on victims of interpersonal violence, rape and sexual assault, and child abuse (Briere & Jordan, 2004; Cutajar et al., 2010; Jewkes, 2002; Kilpatrick, Resick, & Veronen, 1981), surprisingly little psychological research exists regarding the impact that offending has on victims more generally. That we know so little about the efficacy of interventions designed to assist victims' responses to and manage adverse psychological outcomes is particularly surprising given the ubiquitousness of criminal victimization in our society. Accordingly, it is important that greater attention be paid to evidence-based approaches that aim to assist victims of crime.

Offenders. Much of the work of forensic psychologists involves assessment and treatment of offenders. An expanding body of empirical literature demonstrates

that offender rehabilitation can significantly reduce recidivism (see Gendreau, Goggin, & Smith, Chapter 23 of this volume; Gendreau, Little, & Goggin, 1996; Gendreau & Ross, 1979; Losel, 1995; McGuire, 2002; Morgan, Kroner, Mills, & Batastini, Chapter 24 this volume). Contemporary approaches to offender rehabilitation have been drawn from the Psychology of Criminal Conduct (PCC) and the Risk-Needs-Responsivity (RNR) principles that are derived from the model (Andrews & Bonta, 2003). The PCC, which was developed by Andrews and Bonta in the 1980s and refined over time, is a theory concerned with individual differences in criminal behavior, making it a particularly useful guide both for assessing the risk of recidivism and for planning rehabilitation attempts. The PCC provides directions for the assessment and treatment of offenders that are embodied in the principles of RNR. The risk principle directs that the degree of intensity of treatment programs for offenders must be matched to an offender's level of risk (Simourd & Hoge, 2000). Therefore, more intensive intervention is provided to those assessed as being a high risk for reoffending. Conversely, lower-risk offenders have been shown to derive better outcomes from a less intensive level of service and intervention. The needs principle posits that, to reduce recidivism, treatment must focus on the offender's "criminogenic needs" (i.e., the characteristics that contribute to the individual's offending). The responsivity principle considers factors that may affect or even impede an offender's response to interventions. Two general types of factors affect responsivity. One involves factors internal to the individual including, for example, intellectual functioning, self-esteem, and motivation level (i.e., idiographic components). A second type involves external factors such as staff characteristics, therapeutic relationships, environmental support, and program content and delivery (i.e., nomothetic components). Taken together, offender rehabilitation programs that are based on the RNR principles have been found to significantly reduce reoffense rates among offenders (Andrews & Bonta, 2003).

The PCC and the development of the RNR model have formed the basis for many of the gains made in offender rehabilitation (Morgan, Kroner, Mills, & Batastini, Chapter 24 this volume; Ogloff & Davis, 2004). Using these principles, rates of reoffending can be reduced by as much as 30% across different types of offenders (i.e., sexual offenders, violent offenders, and those who perpetrate interpersonal violence).

PSYCHOLOGISTS RESEARCHING PSYCHOLOGICAL MATTERS OF PARTICULAR INTEREST TO THE LEGAL SYSTEM

In addition to researching the legal system and its functioning, psychologists conduct research on a multitude of psychological factors or phenomena that are of particular interest to the legal system. Some of these areas of inquiry are discussed next.

RESEARCHING PSYCHOLOGICAL PHENOMENA

Some psychological phenomena are of particular interest to the legal system given their nature, and research psychologists have made considerable contributions in these areas. Because identification of the accused by eyewitnesses is integral to many convictions (and responsible for a significant number of wrongful convictions), understanding how (in)accurate eyewitnesses are and what factors may affect their recollections is highly important. Not surprisingly, a voluminous literature examines this issue, most of which has been produced by experimental and social psychologists (see Wells & Loftus, 2012, for a concise summary). Similarly, factors associated with increased suggestibility of child witnesses and strategies that can be employed to increase accuracy of their accounts are of great import to the legal system. In response, research psychologists (typically developmental psychologists) have made great contributions in this area as well (see, e.g., Bottoms, Najdowski, & Goodman, 2009; Ceci & Bruck, 1999; Hobbs et al., Chapter 18 this volume; Kuehnle & Connell, 2009).

As a final example, psychopathy—defined as a “constellation of affective, interpersonal and behavioral characteristics, including egocentricity, impulsivity, irresponsibility, shallow emotions, lack of empathy, guilt, or remorse; pathological lying; manipulateness; and the persistent violation of social norms and expectations”—is of particular interest to the legal system, given the high rates of persistent offending displayed by persons with this disorder (Cleckley, 1941; Hare, 1998, p. 188). A great deal of research literature has been compiled by clinical and experimental psychologists examining the causes, correlates, manifestations, assessment, and treatment of psychopathy. (For summaries of work in this area, see Millon, Simonsen, Birket-Smith, & Davis, 1998; Patrick, 2007.)

RESEARCHING PSYCHOLEGAL CONSTRUCTS AND THEIR ASSESSMENT

Many legal issues involve matters of psychology (e.g., whether a defendant’s intellectual limitations affect his ability to understand and participate in the legal process; if and how a person’s cognitive functioning affects her ability to execute a will; whether the psychiatric symptoms experienced by a person limit his ability to make a decision about consenting to or refusing treatment; whether psychiatric symptoms a defendant was experiencing at the time of the offense impaired her in such a way that she should not be held criminally responsible). These are typically referred to as psycholegal capacities (Grisso, 1986, 2003). Understanding the nature of these psycholegal capacities is important to both the legal system and the mental health professionals who may be called on to assess persons when such capacities are at issue. In addition, psychologists researching these matters (most of whom are clinical psychologists) have devoted considerable efforts to defining the nature of these capacities (see, e.g., the work of Grisso & Appelbaum, 1998, examining the psycholegal construct of capacity to consent to treatment consent or the work

of Poythress, Monahan, Bonnie, Otto, & Hoge, 2002, examining the psycholegal construct of trial competence).

When evaluating psycholegal capacities, psychologists employ a variety of tools and techniques, some of which are used in more traditional settings that psychologists work in and some of which have been developed for use in forensic settings. These assessment tasks are either descriptive (insofar as they require the psychologist to describe the examinee's emotional, behavioral, or cognitive functioning at some point in time as it relates to some issue before the court, such as how psychiatric symptoms experienced by the defendant affect his or her understanding of and ability to participate in the legal process) or predictive (insofar as they require that the psychologist assess the likelihood that the examinee will engage in some behavior in the future, such as the risk that the inmate will reoffend violently if paroled to the community).

Integral to understanding the potential value and accuracy of such evaluations is research that operationalizes these psycholegal capacities and examines the utility of various techniques and instruments that are designed to assess them (Douglas, Otto, Desmarais, & Borum, 2012). Only if research findings indicating that psychologists have some special abilities in understanding these issues are (1) psychologists justified as entering the courts as "experts" and (2) the courts justified in hearing psychologists' testimony. Thus, psychologists conduct research examining the utility of various assessment tools and approaches that have been developed to describe psycholegal constructs or inform predictions of behaviors of interest. (For summaries of research examining assessment of defendants' competence to proceed with the legal process, see, e.g., Stafford & Sellbom, 2012; Zapf & Roesch, 2010; Zapf, Roesch, & Pirelli, Chapter 11 this volume.)

Another common forensic evaluation task involves prediction. Predictive evaluations require the psychologist to comment on the likelihood of some future event (e.g., the risk that the examinee will engage in violent or criminal behavior in the near future). In an attempt to identify the causes of violent behavior (with an eye toward preventing such behavior), a psychologist might investigate its emotional, behavioral, and cognitive correlates and the efficacy of various treatments and interventions. Alternatively, psychologists might conduct research examining the accuracy of a test or tool that has been designed to inform mental health professionals' judgments about such risk. (For summaries of research examining violence risk assessment, see Douglas, Hart, Groscup, & Litwack, Chapter 14 this volume; Monahan, 2012; and Otto & Douglas, 2010.) In both cases, the psychologist is conducting research relevant to an important issue of concern to courts: competence to stand trial or violence risk assessment and intervention.

SUMMARY

There is no uniform or consensual definition of forensic psychology, and it is clear that psychologists make contributions to the legal system in a multitude of

ways. Hugo Münsterberg, the first and perhaps most ardent proponent of what psychology had to offer the legal system, recognized this more than a century ago. In his 1908 treatise, *On the Witness Stand*, Münsterberg discussed a variety of ways in which psychologists could contribute to legal proceedings, with chapters devoted to phenomena as varied as lie detection, eyewitness memory, false confessions, and crime prevention. Münsterberg's early work was widely castigated, and he could not have fathomed the success and growth the field has seen over the past century (Bartol & Bartol, Chapter 1 this volume; Ogloff, 2011).

In the century that has passed since Münsterberg's book, psychologists have become involved with the legal process and provided assistance to the legal system a multitude of ways, providing assistance to the courts in decision making, researching matters of interest to the legal system, and offering services to persons involved with the legal system. As the information in this chapter shows, the roles that forensic psychologists play are broad and varied, with some areas being far more developed than others (e.g., the treatment of criminal offenders versus the treatment of victims of crime). The reason that forensic psychology, as a field, is broad and diverse is that it reflects the breadth and diversity of the law. As such, we can expect an expanding array of topics in law with which psychology can contribute on both the micro (i.e., individual) and macro (i.e., systemic) levels.

REFERENCES

- Alison, L., & Rainbow, L. (Eds.). (2011). *Professionalizing offender profiling: Forensic investigative psychology in practice*. Oxon, England: Routledge.
- American Psychological Association. (2012). Guidelines for the practice of parenting coordination. *American Psychologist*, 67(1), 63–71. doi:10.1037/a0024646
- American Psychological Association. (2013). Specialty guidelines for forensic psychology. *American Psychologist*, 68(1), 7–19. doi:10.1037/a0029889
- Andrews, D. A., & Bonta, J. (2003). *The psychology of criminal conduct* (3rd ed.). Cincinnati, OH: Anderson.
- Association of Family and Conciliation Courts. (2000). *Model standards of practice for family and divorce mediation*. Madison, WI: Author.
- Association of Family and Conciliation Courts. (2005). *Guidelines for parenting coordination*. Madison, WI: Author.
- Bartol, C. (1996). Police psychology: Then, now and beyond. *Criminal Justice and Behavior*, 23, 70–89.
- Beck, C., & Sales, B. D. (2001). *Family mediation: Facts, myths and future prospects*. Washington, DC: American Psychological Association.
- Bottoms, B. L., Najdowski, C. J., & Goodman, G. S. (2009). *Children as victims, witnesses, and offenders: Psychological science and the law*. New York, NY: Guilford Press.
- Briere, J., & Jordan, C. E. (2004). Violence against women: Outcome complexity and implications for assessment and treatment. *Journal of Interpersonal Violence*, 19, 1277–1282.
- Brigham, J. (1999). What is forensic psychology anyway? *Law and Human Behavior*, 23, 273–298.

- Brockman, J., & Rose, V. R. (2011). *An introduction to Canadian criminal procedure and evidence* (4th ed.). Toronto, Canada: Nelson Thomson Learning.
- Brodsky, S. L. (2009). *Principles and practices of trial consulting*. New York, NY: Guilford Press.
- Brussel, J. A. (1968). *Casebook of a crime psychiatrist*. New York, NY: Bernard Geis Associates.
- Bussey, K., Lee, K., & Grimbeck, E. J. (1993). Lies and secrets: Implications for children's reporting of sexual abuse. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses: Understanding and improving testimony* (pp. 147–168). New York, NY: Guilford Press.
- Canter, J., & Youngs, D. (2009). *Investigative psychology: Offender profiling and the analysis of criminal action*. Chichester, England: Wiley.
- Ceci, S., & Bruck, M. (1999). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. Washington, DC: American Psychological Association.
- Cleckley, H. (1941). *The mask of sanity*. St. Louis, MO: Mosby.
- Coates, C. A., Deutsch, R., Starnes, H., Sullivan, M. J., & Sydlík, B. L. (2004). Parenting coordination for high-conflict families. *Family Court Review*, 42, 246–262.
- Corey, D. M., & Borum, R. (2013). Forensic assessment for high-risk occupations. In R. K. Otto (Ed.), *Forensic psychology* (2nd ed., pp. 246–270). Hoboken, NJ: Wiley.
- Craig, R. J. (2005). *Personality-guided forensic psychology*. Washington, DC: American Psychological Association.
- Cronch, L. E., Viljoen, J. L., & Hansen, D. J. (2006). Forensic interviewing in child sexual abuse cases: Current techniques and future directions. *Aggression and Violent Behavior*, 11, 195–207.
- Cutajar, M., Mullen, P. E., Ogloff, J. R. P., Thomas, S. D., Wells, D. L., & Spataro, J. (2010). Psychopathology in a large cohort of sexually abused children followed up to 43 years. *Child Abuse and Neglect*, 34, 813–822.
- Douglas, J. E., Ressler, R. K., Burgess, A. W., & Hartman, C. R. (1986). Criminal profiling and crime scene analysis. *Behavioral Sciences and the Law*, 4, 401–421.
- Douglas, K. S., Otto, R. K., Desmarais, S. L., & Borum, R. (2012). Clinical forensic psychology. In J. A. Schinka & W. F. Velicer (Eds.), *Research methods in psychology* (2nd ed.), pp. 213–244. Hoboken, NJ: Wiley.
- Emery, R. E., Sbarra, D., & Grover, T. (2005). Divorce mediation: Research and reflections. *Family Court Review*, 43, 22–37.
- Federal Rules of Evidence. (2011). Washington, DC: Lexis-Nexis Group.
- Forensic Specialty Council. (2008). *Petition for the recognition of a specialty in professional psychology: Forensic psychology* (renewal). Washington, DC: Author.
- Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works. *Criminology*, 34, 575–607.
- Gendreau, P., & Ross, R. (1979). Effective correctional treatment: Bibliotherapy for cynics. *Crime and Delinquency*, 25, 463–489.
- Grisso, T. (1986). *Evaluating competencies: Forensic assessments and instruments*. New York, NY: Plenum Press.
- Grisso, T. (2003). *Evaluating competencies: Forensic assessments and instruments* (2nd ed.). New York, NY: Kluwer Academic/Plenum Press.
- Grisso, T., & Appelbaum, P. (1998). *Assessing competence to consent to treatment. A guide for physicians and other health care professionals*. New York, NY: Oxford University Press.

- Haney, C. (1980). Psychology and legal change: On the limits of a factual jurisprudence. *Law and Human Behavior*, 17, 371–398.
- Hare, R. (1998). Psychopaths and their nature: Implications for the mental health and criminal justice systems. In T., Millon, E. Simonsen, M. Birket-Smith, & R. D. Davis (Eds.), *Psychopathy: Antisocial, criminal, and violent behavior* (pp. 188–212). New York: Guilford Press.
- Hicks, S. J., & Sales, B. D. (2006). *Criminal profiling: Developing an effective science and practice*. Washington, DC: American Psychological Association.
- Hiller, M. L., Belenko, S., Welsh, W., Zajac, G., & Peters, R. H. (2011). Screening and assessment: An evidence-based process for the management and care of adult drug-involved offenders. In C. G. Leukefeld, J. Gregrich, & T. Gullotta (Eds.), *Handbook on evidence-based substance abuse treatment practice in criminal justice settings* (pp. 45–62). New York, NY: Springer.
- Holloway, J. (2003, January). The perils of profiling. *Monitor on Psychology*, 34, 1.
- Huss, M. (2009). *Forensic psychology: Research, clinical practice and applications*. West Sussex, England: Wiley Blackwell.
- International Association of Chiefs of Police. (2010). *Improving police response to persons with mental illness*. Alexandria, VA: Author.
- Jewkes, R. (2002). Intimate partner violence: Causes and prevalence. *Lancet*, 359, 1423–1429.
- Johnston, J. R., Roseby, V., & Kuehnle, K. (2009). *In the name of the child: A developmental approach to understanding and helping children of conflicted and violent divorce* (2nd ed.). New York, NY: Springer.
- Karmen, A. (2010). *Crime victims: An introduction to criminology*. Belmont, CA: Wadsworth, Cengage Learning.
- Kesic, D., Thomas, S. D., & Ogloff, J. R. P. (2013). Estimated rates of mental disorders in, and situational characteristics of, incidents of nonfatal use of force by police. *Social Psychiatry and Psychiatric Epidemiology*, 48(2): 225–232. doi10.1007/s00127-012-0543-4
- Kilpatrick, D. G., Resick, P. A., & Veronen, L. J. (1981). Effects of a rape experience: A longitudinal study. *Journal of Social Issues*, 37, 105–112.
- Kirkland, K., & Sullivan, M. (2008). Parenting coordination (PC) practice: A survey of experienced professionals. *Family Court Review*, 46, 622–636.
- Kovera, M. B. (2012). Voir dire and jury selection. In R. K. Otto (Ed.), *Forensic psychology* (2nd ed., pp. 630–647). Hoboken, NJ: Wiley.
- Kovera, M. B., Levy, R. J., Borgida, E., & Penrod, S. D. (1994). Expert testimony in child sexual abuse cases: Effects of expert evidence type and cross examination. *Law and Human Behavior*, 18, 653–674.
- Kovera, M. B., & McAuliff, B. D. (2000). The effects of peer review and evidence quality on judge evaluations of psychological science: Are judges effective gatekeepers? *Journal of Applied Psychology*, 85, 574–586.
- Kovera, M. B., McAuliff, B. D., & Hebert, K. S. (1999). Reasoning about scientific evidence: Effects of juror gender and evidence quality on juror decisions in a hostile work environment case. *Journal of Applied Psychology*, 84, 362–375.
- Kressel, N. J., & Kressel, D. F. (2004). *Stack and sway: The new science of jury consulting*. Boulder, CO: Westview Press.
- Kuehnle, K., & Connell, M. (2009). *The evaluation of child sexual abuse allegations: A comprehensive guide to assessment and testimony*. Hoboken, NJ: Wiley.

- Kuehnle, K., & Connell, M. (2012). Child sexual abuse evaluations. In R. K. Otto (Ed.), *Forensic psychology* (2nd ed., pp. 579–614). Hoboken, NJ: Wiley.
- Lösel, F. (1995). The efficacy of correctional treatment: A review and synthesis of meta-evaluations. In J. McGuire (Ed.), *What works: Reducing re-offending: Guidelines from research and practice*, (pp. 70–111). Chichester, England: Wiley.
- McAuliff, B. D., & Kovera, M. B. (2007). Estimating the effects of misleading information on witness accuracy: Can experts tell jurors something they don't already know? *Applied Cognitive Psychology*, *21*, 849–870.
- McAuliff, B. D., & Kovera, M. B. (2008). Juror need for cognition and sensitivity to methodological flaws in expert evidence. *Journal of Applied Social Psychology*, *38*, 385–408.
- McGuire, J. (2002). Criminal sanctions versus psychologically-based interventions with offenders: A comparative empirical analysis. *Psychology, Crime, and Law*, *8*, 183–208.
- McLean, M. (1995). Quality investigation? Police interviewing of witnesses. *Medicine, Science and Law*, *35*, 116–122.
- Melton, G. B., Petrila, J., Poythress, N. G., Slobogin, C., Lyons, P., & Otto, R. K. (2007). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (3rd ed.). New York, NY: Guilford Press.
- Millon, T., Simonsen, E., Birket-Smith, M., & Davis, R. D. (Eds.). (1998). *Psychopathy: Antisocial, criminal, and violent behavior*. New York, NY: Guilford Press.
- Monahan, J. (2012). Violence risk assessment. In R. K. Otto (Ed.), *Forensic psychology* (2nd ed.) (pp. 541–555). Hoboken, NJ: Wiley.
- Münsterberg, H. (1908). *On the witness stand*. Garden City, NY: Doubleday.
- Ogloff, J. R. P. (2011). A century of psychology and law: Successes, challenges, and future opportunities. In P. Martin et al. (Eds.), *The IAAP handbook of applied psychology* (pp. 362–385). London, England: Wiley-Blackwell.
- Ogloff, J. R. P., & Davis, M. R. (2004). Advances in offender assessment and rehabilitation: Contributions of the risk-needs-responsivity approach. *Psychology, Crime and Law*, *10*, 229–242.
- Ogloff, J. R. P., Thomas, S. D. M., Luebbers, S., Baksheev, G., Elliott, I., et al. (2013). Policing services with mentally ill people: Developing greater understanding and best practice. *Australian Psychologist*, *48*, 57–68.
- Olczak, P. V., Kaplan, M. F., & Penrod, S. (1991). Attorneys' lay psychology and its effectiveness in selecting jurors: Three empirical studies. *Journal of Social Behavior and Personality*, *6*, 431–452.
- Otto, R. K., & Douglas, K. S. (Eds.). (2010). *Handbook of violence risk assessment*. New York: Routledge, Taylor and Francis Group.
- Otto, R. K., & Heilbrun, K. (2002). The practice of forensic psychology: A look toward the future in light of the past. *American Psychologist*, *57*, 5–18.
- Patrick, C. J. (Ed.). (2007). *Handbook of psychopathy*. New York, NY: Guilford Press.
- Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision-making. *Journal of Personality and Social Psychology*, *51*, 242–258.
- Peters, R. H. (2011). *Drug court treatment services: Applying research findings to practice—an issues commentary and resource brief*. Williamsburg, VA: National Center for State Courts.
- Peters, R. H., & Belenko, S. (2011). Diversion programs. In M. Kleiman, J. Hawdon, & G. Golson (Eds.), *Encyclopedia of drug policy* (pp. 201–203). Thousand Oaks, CA: Sage.

- Peters, R. H., Haas, A., & Hunt, M. (2002). Treatment “dosage” effects in drug court programs. *Journal of Offender Rehabilitation*, 33, 63–72.
- Peters, R. H., & Murrin, M. (2000). Effectiveness of treatment-based drug courts in reducing criminal recidivism. *Criminal Justice and Behavior*, 27, 72–96. doi:10.1177/0093854800027001005C
- Peters, R. H., & Young, S. (2011). Coerced drug treatment. In M. Kleiman, J. Hawdon, & G. Golson (Eds.), *Encyclopedia of drug policy* (pp. 142–145). Thousand Oaks, CA: Sage.
- Porter, B. (1983). Mind hunters. *Psychology Today*, 17, 44–52.
- Posey, A. J., & Dahl, L. M. (2002). Beyond pretrial publicity: Legal and ethical issues associated with change of venue surveys. *Law and Human Behavior*, 26, 107–125.
- Posey, A. J., & Wrightsman, L. S. (2005). *Trial consulting*. New York, NY: Oxford University Press.
- Poythress, N., Monahan, J., Bonnie, R., Otto, R. K., & Hoge, S. (2002). *Adjudicative competence: The MacArthur studies*. New York, NY: Kluwer Academic/Plenum Press.
- Redlich, A. D., Steadman, H. D., Monahan, J., Robbins, P. C., & Petrila, J. (2005). Patterns of practice in mental health courts. *Law and Human Behavior*, 30, 347–362.
- Redlich, A. D., Steadman, H. J., Monahan, J., Petrila, J., & Griffin, P. A. (2005). The second generation of mental health courts. *Psychology, Public Policy, and Law*, 11, 527–538.
- Simourd, D., & Hoge, R. (2000). Criminal psychopathy: A risk-and-need perspective. *Criminal Justice and Behavior*, 27, 256–272.
- Snook, B., Cullen, R. M., Bennell, C., Taylor, P., & Gendreau, P. (2008). The criminal profiling illusion: What’s behind the smoke and mirrors? *Criminal Justice and Behavior*, 35, 1257–1276.
- Stafford, K. P., & Sellbom, M. O. (2012). Assessment of competence to stand trial. In R. K. Otto (Ed.), *Forensic psychology* (2nd ed., pp. 412–439). Hoboken, NJ: Wiley.
- Sullivan, M. (2008). Coparenting and the parenting coordination process. *Journal of Child Custody*, 5, 4–24.
- Swain v. Alabama, 380 U.S. 202 (1965).
- Walker, L., & Monahan, J. (1987). Social frameworks: A new use of social science in law. *Virginia Law Review*, 73, 559–792.
- Wells, G., & Loftus, E. (2012). Eyewitness memory for people and events. In R. K. Otto (Ed.), *Forensic psychology* (2nd ed., pp. 617–629). Hoboken, NJ: Wiley.
- Wilson, C., & Powell, M. (2001). *A guide to interviewing children: Essential skills for counsellors, police, lawyers and social workers*. Sydney, New South Wales: Allen & Unwin.
- Zapf, P., & Roesch, R. (2010). Evaluation of competence to stand trial. New York, NY: Oxford University Press.
- Zippo Manufacturing Company v. Rogers Imports, 216 F.Supp 670 (Southern District of New York 1963).

CHAPTER 3

Accessing the Law and Legal Literature

DAVID DEMATTEO, MICHAEL E. KEESLER, AND HEIDI STROHMAIER

THE term *forensic mental health assessment* (FMHA) refers to the process by which mental health professionals—typically psychologists, psychiatrists, and social workers—conduct evaluations for the court and/or at the request of attorneys (Heilbrun, 2001). The goal of these evaluations is to facilitate better-informed legal decision making by a court or to assist attorneys in their representation of a client. In contrast to general clinical assessments in which the referral question is necessarily clinical in nature, the referral question in FMHAs is guided by a specific legal issue defined by a law within the jurisdiction in which the assessment occurs.

Given the nature of FMHAs, it is incumbent on the mental health professional conducting such an assessment to be familiar with the applicable law in the jurisdiction and the particular legal question at issue in the case. Indeed, Section 2.04 of the Specialty Guidelines for Forensic Psychology underscores the necessity for clinicians conducting FMHAs to possess knowledge and understanding of the legal and professional standards, laws, rules, and precedents that govern participation in legal proceedings (American Psychological Association [APA], 2013; the Specialty Guidelines are reprinted as the appendix to this volume with permission of the APA). Moreover, a large portion of the questions on the written examination for board certification in forensic psychology by the American Board of Forensic Psychology (n.d.) assess clinicians' knowledge of laws, precedents, court rules, civil/criminal procedures, and judicial practices. Similarly, board certification for forensic psychiatrists includes, among other requirements, passing an examination that focuses on criminal law, civil law, legal regulation of psychiatry, and other legal topics (American Board of Psychiatry and Neurology, n.d.).

In this chapter, we present an overview of conducting legal research, with a specific eye toward mental health professionals who do not have any formal legal training. We begin by discussing the importance of understanding the law, particularly for those mental health professionals who conduct FMHAs. Then we

provide an overview of the law in the United States, which includes a discussion of various sources of law and the structure of the state and federal court systems. Next, we describe specific methods for conducting legal research—highlighting both formal and informal approaches—using a variety of databases and research tools. We then discuss how to read and interpret the law, which is not an entirely obvious process. Finally, after discussing how to ensure that the law found through legal research is “good law,” we conclude this chapter by describing effective ways of synthesizing the accumulated legal research.

IMPORTANCE OF UNDERSTANDING THE LAW

Because law is established through multiple sources in the United States, acquiring legal knowledge can be a complicated, tedious, and admittedly intimidating process. Fortunately, finding and understanding the law may be accomplished in a variety of formal and informal ways. In addition to examining primary and secondary sources of law—such as case law, statutes, administrative codes, and a variety of legally relevant publications—it is possible to gain proficiency with legal issues through discussions with attorneys and consultation with colleagues and by staying abreast of the relevant behavioral science literature. By possessing a strong legal background, a forensic mental health professional is better able to identify the legal question and relevant forensic issue in a particular case. In turn, this legal knowledge guides the structure of the evaluation, specifies domains upon which the evaluation should concentrate, informs appropriateness of assessment tools, and helps structure the communication of one’s findings.

Prior to developing a methodological approach to conducting an FMHA, the clinician should possess a thorough understanding of the law applicable to the case. From that law flows the questions that will be addressed by the mental health expert. However, because legal definitions are not tantamount to psychological concepts, an important element in conducting an FMHA entails translating legal language about functional capacities and behavior into relevant measurable constructs. For example, no equivalent solitary constructs exist in the mental health sciences for legal terms such as *knowing*, *intelligent*, and *voluntary*. Experts conducting FMHAs must therefore operationalize these legal concepts in a way that permits them to be meaningfully measured. As another example, a forensic practitioner must translate “competence to stand trial” (which is a legal matter or construct) into elements that can be assessed (such as “understand and assist”). Additionally, although it is generally true that “the law does not presume that any psychiatric diagnostic condition is synonymous with any legal incompetency” (Grisso, 1986, p. 8), in some jurisdictions, the law specifies a direct link between a diagnostic condition and a legal outcome (Heilbrun, Grisso, & Goldstein, 2009). This distinction has clear implications for the structure of an FMHA, and it requires forensic mental health professionals to be familiar with the relevant law.

In most forensic contexts, a clinical diagnosis is not sufficient to answer the legal question, so an individual's "functional legal capacities" must be analyzed to relate his or her observable abilities to the relevant legal concept (Grisso, 1986, 2003). These functional abilities must be relevant to the legal question and causally related to a clinical disorder (Morse, 1978). A mental health professional must therefore understand the legal question to be able to operationalize and identify the relevant forensic issues and functional legal capacities and to develop an appropriate structure for the assessment. There is, however, at least one circumstance in which a clinical diagnosis constitutes the ultimate legal issue. Specifically, in *Atkins v. Virginia* (2002), the U.S. Supreme Court categorically excluded offenders with mental retardation from capital punishment on the grounds that executing such offenders would violate the Eighth Amendment's prohibition of cruel and unusual punishment. In cases such as this in which a bright-line rule has been established, the structure of the FMHA is relatively straightforward. However, in the majority of forensic evaluation contexts, no bright-line rule applies, and the evaluator must devise a more complex methodology that incorporates relevant legal questions, an examination of the evaluatee's underlying psychopathology, and an assessment of the evaluatee's relevant functional legal capacities (see Morse, 1978).

A forensic mental health professional who knows relevant laws and who can identify the forensic issues is in a good position to focus an FMHA on specific content domains. For example, by specifying the forensic issues and relevant functional abilities, a forensic mental health professional can narrow the scope of an evaluation by asking questions and administering tests that yield data directly relevant to the legal issue. Similarly, the evaluator will be better equipped to focus on important domains during collateral interviews and while reviewing legal documents. Although many types of FMHA require a circumscribed assessment of a small number of clinically relevant domains, other types of FMHAs—such as child custody and capital mitigation—require a much broader assessment that focuses on a number of areas. Here again, appropriately appraising the legal question and relevant forensic issues drives the focus of the evaluation.

Once the relevant forensic issues are identified, the clinician must select a specific assessment approach to examine the evaluatee's functional legal capacities. A combination of nomothetic and idiographic data should be collected to assess clinical condition, functional abilities, and a causal connection between the two. The clinician's formulation of the relevant forensic issues should inform his or her selection of appropriate forensic assessment instruments. Clinical assessment instruments that address constructs relevant to criminal and civil cases may also be used in the evaluation (Heilbrun et al., 2009).

The forensic mental health professional must also strive to ensure that his or her testimony will likely be deemed admissible in court under applicable evidentiary standards. These evidentiary standards compel forensic evaluators both to select well-validated psychological assessment instruments and be familiar with the psychometric properties of the tools they use. Rule 702 of the Federal Rules of

Evidence (FRE) and the U.S. Supreme Court's decision in *Daubert v. Merrell Dow Pharmaceuticals* (1993) stress relevance and reliability as the critical elements for the acceptance of scientific evidence in court. Although the federal court system and the majority of states have adopted the evidentiary admissibility standard outlined in *Daubert* (1993), the less stringent *Frye* (1923) admissibility standard is still used in several jurisdictions. The *Frye* (1923) "general acceptance" standard states that for scientific evidence to be admissible in court, the procedure from which the evidence is deduced must be sufficiently established to have gained general acceptance within the relevant scientific community; of note, it is the procedures used to obtain the scientific evidence, not the evidence itself, that must be generally accepted. Importantly, a forensic mental health professional conducting an FMHA may be held to different criteria depending on which evidentiary standard applies in the jurisdiction in which an evaluation occurs. As such, it is crucial that forensic mental health professionals consider the relevant evidentiary standards when selecting measures and scientific approaches for conducting the evaluation to ensure that their testimony will be admissible in court.

In addition to guiding evaluation procedures, knowledge of relevant laws plays a key role in shaping conclusions and communication of findings in FMHAs. With respect to forming opinions, some jurisdictions prohibit testimony on the ultimate legal issue, reserving this role exclusively for the legal decision maker. For example, although testimony on the ultimate legal issue is permitted under the FRE (see Rule 704(a)), expert witnesses in federal criminal trials are not permitted to state an opinion about whether the defendant did or did not have a mental state or condition that constitutes an element of the crime charged or an element of a defense to the crime (see Rule 704(b)). Because jurisdictional rules regarding ultimate issue testimony may differ, it is crucial that forensic mental health professionals are well versed in the laws of the jurisdiction so they avoid overstepping their boundaries. Moreover, even in jurisdictions that allow a forensic mental health professional to answer the ultimate legal issue, current best-practice standards suggest that practitioners nonetheless refrain from offering ultimate issue testimony (Heilbrun et al., 2009).

As discussed later in this chapter, various laws are relevant to FMHAs, which underscores the need for forensic mental health professionals to possess legal knowledge relevant to their practice. Unfortunately, however, as Heilbrun (2001) notes, many clinicians falsely espouse the belief "that a good 'clinical' evaluation will serve in forensic cases," while others fail "to distinguish properly between different forensic issues" (p. 26). Such oversight has led some courts to become disenchanted with the field of forensic mental health and for some judges to express concern that forensic mental health professionals have failed to deliver what the legal system needs (Grisso, 2003). This has led to a growing literature regarding various concerns about FMHAs, many of which appear to be directly related to forensic mental health professionals' lack of sufficient legal knowledge. Prominent criticisms highlighted by Grisso (1986, 2003) over 25 years ago include ignorance

and irrelevance, intrusion into matters of law, and insufficiency and incredibility of information. Each of these concerns is briefly discussed.

Clinician ignorance of the laws driving a particular evaluation and the subsequent irrelevance of their data are of clear relevance to this discussion. As previously described, some degree of legal knowledge is a necessary precursor to devising and conducting a high-quality FMHA. Lacking a sufficient understanding of pertinent laws may preclude a forensic clinician from identifying the relevant forensic issues and functional legal capacities in a particular case. Clinicians who do not possess adequate legal knowledge may present findings that are irrelevant to the legal question. For example, a common mistake is to substitute a clinical diagnosis for a calculated analysis of a functional ability (Simon & Gold, 2004). Notwithstanding the previously discussed *Atkins* exception wherein a clinical diagnosis constitutes the ultimate legal issue, normative “diagnostic testimony” exclusively concerning a mental health disorder highlights Grisso’s (2003) concerns relating to ignorance and irrelevance (see Slobogin, Rai, & Reisner, 2009). Organizations such as the American Psychology–Law Society (AP-LS; Division 41 of the American Psychological Association) and the American Academy of Forensic Psychology have worked to correct this type of oversight by publishing guidelines that emphasize the responsibility of forensic mental health professionals to understand the legal standards pertaining to an evaluation and how these standards relate to clinical practice (APA, 2013).

The second concern highlighted by Grisso (2003)—intrusion into matters of law—is driven primarily by lack of sufficient legal knowledge. This particular mistake generally occurs when forensic mental health professionals inappropriately provide testimony on the ultimate legal issue. There is a vigorous debate within the forensic psychology field regarding whether it is appropriate for forensic experts to answer the ultimate legal question in a particular case. Some scholars have argued that providing ultimate issue testimony is not problematic because many judges and attorneys request forensic clinicians’ opinions on the ultimate legal question (see, e.g., Rogers & Ewing, 1989; Rogers & Shuman, 2000), whereas others counsel against answering the ultimate legal question because it represents an inappropriate intrusion into the domain of the legal fact finder (see, e.g., Grisso, 1986, 2003; Heilbrun et al., 2009; Tillbrook, Mumley, & Grisso, 2003). These scholars argue that because an opinion on an ultimate legal issue cannot be rendered without consideration of morals and values, experts conducting FMHAs should not stray from their role in offering psychological testimony and opinions related to clinical condition, functional legal capacities, and the relation between the two.

The third concern highlighted by Grisso (2003) concerns the insufficiency and lack of credibility of some clinicians’ reports, which is likely caused at least in part by an insufficient grasp of the relevant legal issues. Specifically, clinicians conducting FMHAs have been criticized for drawing conclusions that have insufficient supporting evidence and rendering opinions that lack credibility (Grisso, 2003). It would seem that a clinician well versed in the evidentiary standards and other laws applicable in the evaluation’s jurisdiction would be much less likely to offer

conclusions based on insufficient scientific evidence. Fortunately, recent developments in the scientific approaches used to conduct FMHAs and improvements in the law theoretically should reduce the incidence of insufficient and incredible testimony. However, testimony that relies on speculation and ill-supported or novel theories is unfortunately not uncommon and is often presented in court as fact, which leads to confusion or mistaken beliefs among legal decision makers.

All of these issues and concerns underscore the need for clinicians interested in conducting FMHAs to possess an understanding of the law and the ability to appropriately apply such knowledge in forensic contexts. By enhancing their knowledge of relevant laws and becoming more aware of how to conduct legal research, forensic mental health professionals can take an important step toward becoming more competent providers of FMHAs, which will in turn provide attorneys and courts with more useful information. The net result will be a legal system that makes better-informed decisions. Fortunately, efforts undertaken by experts and organizations in the field have helped mitigate problems associated with FMHAs. Nevertheless, there remains a need to continue working to improve the competence of forensic mental health professionals and the overall quality of FMHAs, and a good starting point in that undertaking is having a solid understanding of the law.

STRUCTURE OF THE LAW

Laws come from a variety of sources and appear in various formats. For this and other reasons, looking up “the law” sometimes can be a surprisingly convoluted process. It is not uncommon for novice researchers or those without formal legal training to struggle when first conducting legal research. Fortunately, as any first-year law student would likely attest, there is a profound learning curve when it comes to conducting legal research. We have written this section, and the entire chapter for that matter, to help bend that learning curve in a favorable direction.

PRIMARY SOURCES OF LAW

In the United States, multiple entities provide the laws that govern society, but two main forces are typically at play: the federal government and state governments. Due to the Tenth Amendment, the federal government and state governments have separate powers, and states are sovereign in many respects. Both the federal government and state governments can be subdivided into their respective legislative, judicial, and executive branches. As with the overarching governmental structures, a separation of powers also exists among the three branches of the government. Put simply, the legislative branch creates statutes, the judicial branch interprets those statutes via case law, and the executive branch creates administrative agencies that carry out the law and enact regulations. The law created by the executive branch, which is often in the form of an executive order issued by the president or a state governor, is of limited utility to forensic mental health professionals, so here we focus primarily on statutes, case law, and constitutional law.

Statutory Law. When many people envision what a law looks like, they may imagine some code or subsection with a potentially long string of numbers. For example, a law might be referenced as 18 USC § 2340A (United States Code, Title 18, Part 1, Chapter 113 C, Section 2340). (For those interested, this law criminalizes the act of torture committed inside the geographic jurisdiction of the United States or committed by a U.S. national regardless of where the act occurs.) This type of law—the law “on the books” or “black letter law”—is referred to as statutory law. Statutory law flows from statutes enacted by legislatures, at either the federal level (through Congress) or the state level (through state legislatures). These statutes are then compiled into federal or state codes, which we discuss in more detail later in the chapter. Many statutes are quite old, having been enacted decades ago and having never been revised. Others are comparatively young—for example, statutes enacted by legislative authority to address emerging concerns or new issues not previously covered under statute (e.g., privacy of electronic information). Statutes may also be revised by the issuing authority to address concerns relating to content (e.g., to delete provisions that no longer apply) or scope (e.g., ensuring the statute is neither too narrow nor too broad). Legislatures may also revise statutes in reaction to recent judicial decisions about how a statute is to be interpreted (e.g., to make the language more precise so courts can better interpret the statute in line with the legislature’s intended purpose).

Case Law. After a statute is enacted, the courts are responsible for interpreting and applying that statutory law in the context of specific disputes, which means that courts are also a source of law. In the process of applying the law, the courts determine where and when the statute applies, whether there are any exceptions to its application, and how vague statutory language should be interpreted. If a court goes so far as to declare a statute invalid, then the statute must be revised by the legislature before it can legally be applied. Assuming, however, that the statute is not invalidated, courts will apply the statute in the context of legal disputes. As the courts interpret and apply a statute one case at a time, collections of judicial decisions amass. These decisions are referred to as case law or common law. Judges may choose to follow this “legal precedent,” interpreting the law similarly to judges who have already reviewed the issue. However, a judge may choose to depart from precedent altogether or identify the case before the court as factually distinct and therefore not guided by precedent. In so doing, the courts create new case law. New case law may simply answer previously unanswered legal questions, or it may effectively change the law in the jurisdiction, even if no new statute has been enacted.

Constitutions. The third major source of law is constitutions. Constitutions provide specific rights and protections and are often the authority on which courts rely when invalidating a statute. In the United States, individuals are protected by both the

U.S. Constitution (or Federal Constitution) and their individual state constitution. Individuals are *protected* by constitutions, and the legislatures, courts, and executive agencies are *bound* by constitutions. In that sense, constitutions largely protect individuals by limiting governmental power. Many state constitutions closely parallel the U.S. Constitution, but that pattern is more convention or convenience than requirement. Rather, state constitutions may differ from the U.S. Constitution, with one very important caveat: State constitutions may not provide *less* protection than the U.S. Constitution, although they may provide *more* protection.

SECONDARY SOURCES OF LAW

The three major sources of law discussed thus far are all considered to be primary sources because they are the sources that effectively make the law. In addition to these primary sources, there are secondary sources of law. Broadly speaking, secondary sources can be anything describing or summarizing the law that do not fall into one of the aforementioned primary sources. Thus, secondary sources may include books purchased at a local bookstore, Internet articles or forums discussing a topic of law, or even video or audio recordings. Speaking more narrowly, though, there are a few specific secondary legal sources to which legal researchers often turn.

Legal dictionaries provide the meaning and typical use of legal terms and phrases. Legal encyclopedias provide broad and general information on a specific topic of law, and they are much like a traditional encyclopedia in both breadth and depth. Moving from the general to the specific, legal treatises are entire books devoted to one area of law, which means they have much less breadth than legal encyclopedias but considerably more depth. Hornbooks are a specific type of treatise. Although hornbooks are traditionally written to help law students navigate through an unfamiliar area of the law, they are helpful for anyone new to a specific area. Nutshells are another type of treatise, similarly devoted to one area of law, but with an emphasis on conciseness. Last, the Restatements of Law are an effort by the American Law Institute (ALI) to organize, summarize, and “codify” the common law of the United States. Covering 15 separate areas, the Restatements sometimes straddle the line between primary and secondary sources. Whereas they are in themselves a secondary source because they synthesize the law in a given area, it is not unusual for judges to rely on the Restatements as persuasive authority in deciding a novel legal issue.

Secondary sources of law can be tremendously helpful to both novice and expert legal researchers. An individual who must investigate an unfamiliar area of law might begin with a secondary source to get a broad overview of the law, which will provide insight as to how he or she should go about researching primary source material. Someone who feels confident in his or her primary source research may turn later to secondary sources of law as a final step to ensure that nothing has been overlooked. Of course, there is no reason why the two sources of law cannot be used simultaneously.

Having reviewed primary and secondary sources of law, an important note is worth highlighting for the forensic mental health professional. Because there are multiple sources of law, comprehensive legal research involves more than simply looking up the relevant guiding statute. Researching beyond the statute—for example, by looking at case law to see how courts have applied and interpreted the statute—is particularly important if the statute has not been revised in many years or if there is reason to believe that it is no longer “good law,” which is a concept we discuss later in this chapter. A forensic mental health professional engaging in best practices should be familiar with relevant statutory law and case law on a given legal question, and that task can be accomplished more easily by utilizing the secondary legal sources just described.

STRUCTURE OF THE UNITED STATES COURTS

Having explored where law comes from, it is also necessary to understand which laws are applied by which courts. Just as the United States has two major government systems that make the law (i.e., Congress and state legislatures), so too does it have two major court systems to apply those laws: federal courts and state courts. Each court system can be generally divided again into criminal and civil courts. State courts hear state matters: crimes violating state law and civil disputes arising between residents of the same state. Federal courts hear federal matters: crimes violating federal law and civil disputes arising between residents of different states (or on certain issues reserved for the federal courts).

Criminal courts handle trials brought by the government against an individual (e.g., *State v. Jones* or *United States v. Jones*). Because a crime is viewed as an offense against society, the state or federal prosecutor is responsible for bringing charges against a defendant, and he or she reserves the right to drop or amend those charges; there are no private prosecutions in the United States. As civil matters arise between two parties, it is the responsibility of those parties to mount both the offense (plaintiff) and the defense (defendant). Although the typical parties in a civil case are individuals, companies, or corporations, the state can also be party to certain civil proceedings that have relevance to forensic practice. For example, the state is party to a civil commitment proceeding in which an individual may be involuntarily committed to a psychiatric facility if he or she is a danger to self or others. The state is also responsible for challenging parental rights in cases of child abuse or neglect, which are also handled in civil court. Because so many legal questions are answered in the civil courtroom, civil courts are subdivided into various levels, and the civil court system contains a variety of specialized courts, such as probate courts, family courts, and landlord–tenant courts.

Although this organization may seem fairly clear-cut, various complications can make jurisdiction confusing in both criminal and civil matters. Indeed, there are entire courses in many law schools dedicated exclusively to the topic of jurisdiction. A few examples regarding criminal jurisdiction may help illuminate some of the complexities relating to it. Murder, for example, is typically characterized as a

crime against the state (e.g., *State v. Jones*). However, if the murder victim was a federal employee—for example, a postal worker or undercover Federal Bureau of Investigation agent—then that murder would be tried in federal court (e.g., *United States v. Jones*). In addition to the victim, the location of the crime can also play a role in determining jurisdiction. Using illegal drugs is typically a state offense (e.g., *People v. Brown*). However, if someone commits a drug offense in Yosemite National Park (which is federal property within California’s borders), he or she would be prosecuted for that offense in federal court (e.g., *United States v. Brown*). As one last example, stealing a car is typically a state offense (e.g., *Commonwealth v. White*). If, however, that stolen car is subsequently driven across state lines, it would become a federal offense prosecuted in federal court (e.g., *United States v. White*).

Within the federal and state court systems, the courts are further subdivided into hierarchies or tiers. Typically a case is heard at the lowest tier in its relevant jurisdiction, which is called the court of first instance because it is the first court to hear the case. However, through a process of appeals, cases may move up the hierarchy to be reviewed by courts of greater authority. Because both the state and federal court systems have respective tiers and exist parallel to one another, one can conceptualize the U.S. court system as a wide staircase with an iron handrail running up the middle. On one side of the rail lie the federal courts, and the state court system lies on the other side of the rail; the lowest step is the court of first instance, and the top step is the highest court that has the final say on an issue. Typically, cases remain on one side of the rail or the other, but there are instances in which they may jump over the rail from one system to the other.

In the state court system, the first level of courts contains the trial courts of limited jurisdiction. These courts, whose names differ among jurisdictions, handle relatively minor judicial matters, such as parking tickets, vandalism, and accrued fines. Also on the first level are the trial courts of general jurisdiction. Trial courts of general jurisdiction handle the vast majority of cases, both criminal and civil. A trial court’s purpose is to determine the facts of the case and settle the dispute. Criminal trial courts apply penalties such as fines, judicially supervised release, incarceration, and (in some jurisdictions) death. Civil trial courts apply penalties that are typically limited to monetary damages and injunctions.

Moving one step up from the trial courts but remaining in the state court system are the intermediate appellate courts. Appellate courts have appellate jurisdiction, which means that they hear cases that have been appealed. If a party is dissatisfied with the trial court’s ruling, the party may attempt to appeal the case up to the appellate court. However, not all cases are eligible for appeal, and being eligible for an appeal does not guarantee that the appeal will be granted. Grounds for appeal vary by case and jurisdiction, but they often involve procedural error (e.g., a questionable ruling by the judge, prosecutorial misconduct, failure of one party to turn over all of the evidence under the rules of discovery). Simply being unhappy with having lost the case is not sufficient grounds to be granted an appeal. Rather, the case must have an “appealable issue.” (An exception to this rule is death penalty

cases, which are automatically appealed in almost every jurisdiction based on a public policy interest in making sure that all death penalties are scrutinized by a number of individuals before the sentence is affirmed and ultimately administered.) Toward that end, appellate courts rarely reexamine the facts of the case (e.g., who did what to whom, when, and how). Rather, an appellate court, having granted an appeal, will examine and address whether the trial court correctly applied the law in arriving at a verdict and penalty. In this sense, appellate courts are said to resolve issues of law, not issues of fact.

At the highest step in the state court system, each state has its high court, which most states call the state supreme court. The state supreme court is also an appellate court inasmuch as it only hears cases for which it has granted an appeal. In fact, several states (all with smaller populations) lack the aforementioned intermediate appellate courts (e.g., Montana, Rhode Island, West Virginia). In those states, an appeal granted after a trial court case would be heard directly by the state's high court. In states with an intermediate appellate court, the high court would only hear a case that has been adjudicated at the trial level, answered on appeal at the intermediate level, and then appealed again to the state high court. In most states, the state high court is a discretionary court, meaning it is not required to hear any cases on appeal and only decides cases that it accepts. Typically, once the state high court rules on an issue, its decision is final. However, it is possible that a case resolved at the state high court can be reheard in the federal court system. We address this topic after covering the federal court system.

It is worth noting that states often have different names for the trial, appellate, and high courts in their jurisdictions. For example, in New Jersey, the trial court is the superior court, the intermediate appellate court is the superior court appellate division, and the high court is the New Jersey Supreme Court. In Pennsylvania, the trial court is the court of common pleas, the intermediate appellate court is the superior court, and the high court is the Pennsylvania Supreme Court. Perhaps most confusing, in New York, the trial courts are the state supreme courts, the intermediate appellate courts are the appellate courts, and the high court is the New York Court of Appeals.

The federal court system also arranges courts in hierarchical fashion. At the first step, the lowest federal courts are the district courts, which function much like state trial courts. One major difference, though, is that whereas a state trial court has only one judge presiding, a federal district court is presided over by a three-judge panel. There are 94 district courts spread geographically across the continental United States and outlying states and territories. All states fall under the geographic jurisdiction of at least one district court, with many states falling under the jurisdiction of multiple district courts. The federal district courts have general jurisdiction, and they hear both civil and criminal matters. Like state trial courts, federal district courts determine the facts of a case and resolve the dispute.

As in the state system, cases move up the federal court system through an appeals process. The intermediate federal courts are known as the United States Circuit

Courts of Appeal. Fewer in number than the district courts, these 13 circuit courts are similarly spread out across the United States. The 13 circuit courts include 11 regional courts numbered 1 through 11 (e.g., U.S. Court of Appeals for the Third Circuit), the 12th Circuit for the District of Columbia, and the Federal Circuit Court. All of the circuit courts, except for the Federal Circuit Court, have jurisdiction over the courts in their region; the Federal Circuit Court has nationwide jurisdiction, but it hears only certain types of cases (e.g., patent, international trade). Each circuit court has jurisdiction over a specified number of the 94 district courts. For example, the United States Court of Appeals for the Third Circuit has appellate jurisdiction over all of the district courts in Delaware, New Jersey, Pennsylvania, and the U.S. Virgin Islands. As with the federal district courts, circuit court cases are presided over by a three-judge panel.

At the top tier of the federal court system is the U.S. Supreme Court. The U.S. Supreme Court is the highest court in the country, and being a frequent topic of discussion, it is often abbreviated as USSC (United States Supreme Court) or SCOTUS (Supreme Court of the United States). As an appellate court, the Supreme Court only hears cases for which the justices have granted an appeal. Unique to the Supreme Court is its composition of eight Associate Justices and one Chief Justice. There is, however, no basis in the Constitution for setting the number of justices at nine. Article III of the U.S. Constitution authorizes Congress to establish the number of Supreme Court Justices, and the size of the Supreme Court has been changed a few times over the years. For example, the Judiciary Act of 1789 called for the appointment of six justices, but the Supreme Court was expanded to seven members in 1807, nine members in 1837, and 10 members in 1863. The Judiciary Act of 1869 set the number of justices at nine, and it has remained at this size since that time. Each justice is appointed by the president, confirmed by the Senate, and serves for life (until death, retirement, or impeachment). Typically, the Supreme Court is charged with interpreting the U.S. Constitution, meaning that it reviews whether lower court decisions comported or conflicted with constitutional rights and protections. However, the Supreme Court also addresses other matters relating to federal law.

Like most state high courts, the Supreme Court has discretionary jurisdiction. Each year, thousands of cases are appealed to the U.S. Supreme Court, but it hears only a small fraction of those cases (Thompson & Wachtell, 2009). At least four of the nine justices must agree to hear a case (the “rule of four”), and at least five of the nine justices must agree on the outcome for there to be a majority opinion when deciding the case (although four votes can be sufficient if at least two of the justices abstain from voting). The Supreme Court has jurisdiction over all cases decided by federal circuit courts and all cases from state courts that implicate the U.S. Constitution or federal law. As the highest court in the United States, all decisions of the U.S. Supreme Court are final.

As mentioned earlier, a criminal case that has made or is making its way up the steps of the state court system can potentially jump over to the federal court system; the reverse situation—that is, going from federal court to state court—is

not permissible in the American legal system. One common method of jumping from state court to federal court is through a *writ of habeas corpus* (or *habeas corpus ad subjiciendum*), which is an appeal to a federal district court to hear the case at hand. *Habeas corpus*, roughly translated as “you have the body,” is technically a challenge to the legitimacy of a prisoner’s detainment. Thus, the typical habeas petitioner is a prisoner who has been convicted of a crime. In practice, habeas corpus is an assertion that a prisoner’s federal constitutional rights were violated along the way toward his or her current detainment (e.g., Fourth Amendment right against unreasonable search and seizure, Sixth Amendment right to effective assistance of counsel, Fourteenth Amendment right to due process of law). Once a case has jumped over the rail into the federal court system, the case remains in the federal system. Thus, any appeals flowing from the federal district court’s findings must go through the federal appellate process.

Another way for a case to move from the state court system to the federal court system is by a *writ of certiorari*. If issued, a writ of certiorari provides that a case will be heard by the U.S. Supreme Court. Toward that end, a writ of certiorari can bring a case before the U.S. Supreme Court after the case has been settled by either a federal court of appeals or a state high court. In contrast with habeas corpus, writs of certiorari must be filed within 90 days of a judgment being entered. Thus, whereas a writ of habeas corpus may arise after conviction, sentencing, and lengthy detainment, a writ of certiorari requests review of a decision recently rendered. Also, whereas a writ of habeas corpus is applicable only in criminal contexts, a writ of certiorari applies in either civil or criminal contexts. One common misunderstanding is that individuals file writs of certiorari to the Supreme Court requesting that the court hear a particular case. However, it is the Supreme Court that announces whether it will grant an appeal by issuing a writ of certiorari to the appellant (who has appealed the case). This makes sense given the meaning of writ of certiorari, which can be translated as “to be more fully informed.” Recent statistics suggest that the Supreme Court receives roughly 7,500 petitions for certiorari each year but typically only grants 1% to 2% of the petitions (Thompson & Wachtell, 2009).

BINDING VERSUS PERSUASIVE AUTHORITY

Having discussed the sources of law and how cases move through the different courts systems, it is important to understand the reach or influence of a court’s ruling. United States courts follow a system of *stare decisis*, short for the Latin phrase *Stare decisis et non quieta movere*, roughly translated to mean “maintain what has been decided and do not alter that which has been established.” In practice, when a judge faces an issue that has previously been addressed by a court, *stare decisis* dictates that the judge follow the previous court’s decision. That said, the judge is not strictly required to follow the previous court’s decision; it is just an expectation. But the weight of that expectation can vary depending on which court previously addressed the issue at hand. Therein lies the nature of binding authority versus persuasive authority.

If the earlier ruling came from a higher court in the same jurisdiction, then a judge is bound to follow that court's ruling. Because the judge is bound, it is referred to as binding authority. For example, if the Supreme Court of California rules on an issue, the lower courts within California are expected to follow the state Supreme Court's ruling. Strictly speaking, the judge in a lower court is not mandated to follow the higher court's ruling; should the judge deviate, he or she will not be fined, removed, or otherwise penalized, although reelection (in those states in which judges are elected) may be more difficult if the issue is polarizing. But such a deviation would certainly provide grounds for appeal.

By contrast, a court reviewing decisions from a lower court or a court in another jurisdiction is not bound by those decisions. In that context, how other courts have resolved a legal issue is viewed merely as persuasive authority; the deciding court is not bound by those decisions, but it may be guided by those other decisions. Attorneys arguing before the court may cite these other rulings and cases as examples of how they would like the court to rule, attempting to persuade the court to follow the lead set by other courts. Thus, persuasive authority may help guide a court toward a decision, but there should be no expectation that the court will follow it.

One final caveat relates to the nature of binding authority and the interplay between federal courts and state courts. All decisions by the U.S. Supreme Court are binding on federal courts, but not all U.S. Supreme Court decisions are binding on state courts. The determining factor is whether the legal issue implicates state courts or only federal courts. For example, if the U.S. Supreme Court determines that execution is unconstitutional because it constitutes cruel and unusual punishment under the Eighth Amendment, then no states may execute offenders. However, if the Supreme Court's decision is limited to federal courts—for example, how the FRE are to be interpreted—then that decision does not affect state courts. State courts are free to adopt the Supreme Court's position, but they are not required to do so.

CONDUCTING LEGAL RESEARCH

A variety of options are available for those interested in researching the law, and selecting from among the multitude of available resources and various approaches to conducting legal research certainly can be challenging for a mental health professional. Although there are many research options available, forensic mental health professionals must be knowledgeable about which option is best suited for their particular needs. Some options for researching the law are more formal and presume some degree of legal knowledge, whereas other options are quite informal and require no up-front legal knowledge; some options for conducting legal research are expensive and perhaps cost prohibitive to some mental health professionals, whereas other options are free and available to anyone with an Internet connection; and some resources provide comprehensive coverage of federal and state case law, federal and state statutes, and secondary legal literature, whereas other resources

primarily provide coverage of recent decisions from the U.S. Supreme Court. In this section, we review the many options that are available for conducting legal research.

FORMAL LEGAL RESEARCH: ELECTRONIC LEGAL DATABASES

Perhaps the best-known option for conducting legal research is commercially available electronic legal databases. The two undisputed leaders in this industry—Westlaw and LexisNexis—are likely familiar to most readers, even those without formal legal training who have never conducted legal research. Both of these electronic legal databases provide comprehensive coverage of primary and secondary legal materials, and numerous courts, law firms, attorneys, and other legal professionals throughout the United States and many other countries have a long history of relying on these databases. Westlaw and LexisNexis are available only by paid subscription and therefore may not be “available” for those with limited economic resources, although many university libraries and some public libraries subscribe to these databases.

Both Westlaw and LexisNexis have been providing commercially available legal research services since the early 1970s. When these companies were first started, which was largely pre-Internet and certainly before the days of widespread personal computers, these services were available as dial-up services via dedicated terminals. It was not until the late 1980s that both companies began offering Internet-based programs for use on personal computers, and that marked a watershed moment in legal research. Westlaw and LexisNexis offer most of the same services, so choosing between the two is often a matter of personal preference, although the different subscription plans are also likely a factor.

Westlaw was started by West Publishing, which has been headquartered in Eagan, Minnesota, since 1992, and it was acquired by Thomson Corporation in 1996. Westlaw provides comprehensive coverage of state and federal case law, state and federal statutes, administrative regulations, and secondary legal resources (e.g., newspapers, magazines, law journals, law reviews, treatises). Importantly, though, coverage of trial court decisions is largely limited to federal courts; the state cases included in Westlaw typically are limited to appellate decisions. Westlaw has also been providing increasing coverage of international law, and a variety of Westlaw services are currently available in nearly 70 countries worldwide. Besides its comprehensive coverage of legal materials, Westlaw has several features that make it user friendly. For example, West developed a proprietary classification system—West Key Number System (or West American Digest System)—that provides a standardized taxonomy for all U.S. legal materials. Using this system, numerous points of law are organized by topic and key number, which makes it easier for users to identify materials that are relevant to their needs. Westlaw offers several search features, including natural language and Boolean, and it is possible (depending on the nature of the subscription) to search, for example, through all state and federal cases or to limit a search to cases from one jurisdiction. Another

notable feature of Westlaw is KeyCite, which is a citation checking service that allows users to determine whether cases and statutes are still “good law” (which we discuss in more detail later in this chapter).

As mentioned, Westlaw’s primary competitor is LexisNexis. LexisNexis has undergone several name changes since the early 1970s, and it is currently headquartered in Dayton, Ohio. Like Westlaw, LexisNexis provides comprehensive coverage of state and federal case law, state and federal statutes, administrative regulations, and secondary legal resources. As with Westlaw, there is limited coverage of trial court opinions at the state level. LexisNexis provides extensive coverage of published U.S. cases dating back to the late 1700s, and it also provides wide-ranging coverage of other primary and secondary legal resources. The Nexis aspect of LexisNexis is dedicated to providing access to content from thousands of global news sources, public records, legislative and regulatory filings, and many other types of legal materials. Like Westlaw, LexisNexis offers a variety of services and features that enhance its usability and popularity, including services that allow those conducting legal research to determine if a particular law (court decision, statute, etc.) is still good law.

Westlaw and LexisNexis have set the standard when it comes to electronic legal databases, but the services they provide can be costly. Both companies offer a variety of subscription plans tailored to consumer needs; for example, a small law firm that restricts its business to one state does not have the same needs for electronic legal databases as a large law firm that has offices in several states or even several countries. Some of the available pricing plans are flat rate, which means that users have unlimited access to the service for a set time, while other pricing plans are based on the number of transactions conducted (either within an hour or in total over a set period of time). The range for these pricing plans is quite large. Some services may be available for a (relatively) few dollars, while other services can cost thousands of dollars in a given billing period.

Not too long ago, commercial databases such as LexisNexis and Westlaw were the only options available for conducting legal research. Fortunately for those who find Westlaw and LexisNexis to be cost prohibitive or unavailable, several electronic databases that are considerably less costly—sometimes even free—have been developed in recent years. A few examples are presented here. TheLaw (www.thelaw.net) provides coverage of all federal cases (district courts, circuit courts, and U.S. Supreme Court), all federal bankruptcy cases, and most state court cases from 1950 (or earlier in some states); all state and federal statutes and regulations; judicial and administrative rules and forms; and several other legal resources. As of this writing, annual subscriptions range from \$575 for unlimited use by one lawyer and one paralegal to \$1,995 for unlimited use by five lawyers and five paralegals. Another affordable electronic legal database is NationalLawLibrary (www.itislaw.com). The federal law collection includes coverage of case law from the U.S. Supreme Court and all federal circuit courts and coverage of federal rules, and the state law collection provides 50 years of case law

coverage. NationalLawLibrary offers a wide variety of subscription plans, including a federal plan (\$43/month), a single state plan (\$42/month), an all states plus federal plan (\$86/month), and a plan that charges by the transaction (\$3–\$5/search). VersusLaw (www.versuslaw.com) provides coverage of all federal court cases, all state appellate court cases, state and federal statutes, and some additional legal resources. Subscription plans for VersusLaw range from a standard plan (with basic case law coverage) for \$13.95/month (or \$167.40/year), to a professional plan (most comprehensive coverage) for \$39.95/month (or \$479.40/year).

There are also several free electronic resources that provide a wealth of legally relevant information. For example, FindLaw (www.findlaw.com) is a popular option that contains state and federal case law and state and federal statutes, and it also offers access to recent legal news, an attorney directory, education about various legal topics, and a free legal magazine. Cornell University's Legal Information Institute (www.law.cornell.edu) is a useful resource for mental health professionals interested in conducting legal research. It offers access to numerous legal resources, such as the U.S. Constitution, state and federal statutes, and U.S. Supreme Court decisions. It also includes user-friendly resources, such as a legal dictionary and encyclopedia, to facilitate interpretation and comprehension of various laws. Many law schools and law libraries maintain Web sites that provide useful information for laypeople interested in conducting legal research. For example, the Law Library of Congress (www.loc.gov/law) has a number of helpful tutorials and direct links to primary and secondary sources of law. The Southern Illinois University School of Law Library (www.law.siu.edu/lawlib/) maintains a useful Web site, with a variety of free and helpful tips, articles, and blogs on conducting research. The Georgetown Law Library (www.law.georgetown.edu/library/) maintains a similar Web site that includes self-guided tutorials for learning about legal research. Finally, Google Scholar (<http://scholar.google.com/>), which has offered a Legal Opinions and Journals database since 2009, exemplifies recent improvements in public accessibility to legal resources. To reduce costs associated with using subscription-based commercial databases, one option is to begin researching a legal issue on Google Scholar and then move to a commercial database once the research has been sufficiently narrowed.

INFORMAL LEGAL RESEARCH

Although formal legal research using primary sources of law and commercial electronic legal databases, such as Westlaw and LexisNexis, provide accurate and detailed information relevant to forensic mental health professionals, these options can be costly and difficult to access for some mental health practitioners. Fortunately, there are a variety of less formal approaches to conducting legal research. These informal approaches are often as helpful as formal approaches and may be more desirable and practical for forensic clinicians in many situations. For example, forensic mental health professionals may benefit from utilizing libraries and various

print materials and from networking and consulting with other professionals and organizations within the forensic psychology and legal fields.

Although the Internet provides an abundance of information about various legal topics and sources of law, print sources are a viable alternative to electronic sources of legal material. These print sources are available in law school libraries and occasionally even public libraries, and many law schools permit visitors to use their library at certain times. Legal digests, which are series of books that provide brief overviews of the points of law contained in important cases, often serve as a particularly useful starting point when conducting an informal legal search. Digests are organized alphabetically by subject, which generally benefits the forensic practitioner interested in researching a circumscribed number of specific legal issues, and they are available in versions that include state, federal, and U.S. Supreme Court cases.

Legal encyclopedias are similarly beneficial in providing an introduction to a particular area of law. Legal encyclopedias are accessible and user friendly, and they provide citations to primary sources of law. The two primary legal encyclopedias, *American Jurisprudence* (*Am. Jur.*) and *Corpus Juris Secundum* (*C.J.S.*), are excellent resources for both novice and experienced forensic mental health professionals. *Am. Jur.*, which is one of the most frequently referenced legal encyclopedias, is relatively selective in its coverage of topics, which enhances its readability. *C.J.S.* is more exhaustive, containing lengthier and more detailed entries than those found in *Am. Jur.* The legal treatises and hornbooks discussed previously are also useful sources of secondary legal authority.

Finally, a number of books have been written over the past few decades specifically targeted at mental health professionals interested in FMHAs. The *Law and Mental Health Professionals* series, published by the APA, provides a comprehensive overview of state-specific information about many legal topics relevant to forensic practice. A more recent contribution to the forensic literature is the multivolume *Best Practices for Forensic Mental Health Assessments* series published by Oxford University Press. The Oxford series includes books dedicated to specific forensic topics, such as violence risk assessment, competence to stand trial, civil commitment, and personal injury, and each book covers the legal standards applicable to the type of assessment it is covering. Numerous other books written for and by forensic mental health professionals (including the present volume) are available as useful sources of legal information. Some of these books focus on specific types of FMHAs (e.g., criminal responsibility, capital sentencing), and others focus more broadly on the process of conducting FMHAs.

Legal periodicals and other scholarly journals may also assist the forensic mental health professional in acquiring valuable legal information that will facilitate his or her understanding of relevant law. Although many legal periodicals emphasize legal topics that are not directly relevant to FMHAs, other publications specialize in psycholegal topics that are likely to be particularly useful to clinicians who conduct FMHAs. Law reviews are a valuable type of legal periodical. The University of

Alabama's *Law and Psychology Review* is one such journal likely to assist forensic clinicians interested in gleaning information about applicable law. Another relevant law review is the *Mental Health Law and Policy Journal*, published by the Cecil Humphreys School of Law at the University of Memphis. In addition to law reviews, other types of legal periodicals may be useful. The *Journal of Psychiatry and Law* and the *American Journal of Law and Medicine* are peer-reviewed journals that include articles with clear relevance to forensic practice. A number of other scholarly peer-reviewed journals focus on interdisciplinary topics related to psychology and law. *Psychology, Public Policy, and Law*, which is a hybrid law–psychology journal, and *Law and Human Behavior* are two examples of journals published by the APA that provide legal information relevant to forensic clinicians. Articles published in such nonlegal periodicals typically are not written using highly technical language and are therefore more likely to be comprehensible to forensic mental health professionals who do not have formal legal training.

Although the printed legal resources just discussed are available for purchase, many of them are also accessible for free at local libraries and university libraries. Forensic mental health professionals are encouraged to obtain particularly helpful legal digests, treatises, or other legal resources to have on hand as convenient references. However, the advantages of utilizing public, university, and law libraries should not be overlooked. Law libraries provide comprehensive and varied source of legal resources, and occasionally they permit visitors to access their collections. Moreover, law librarians can be a valuable resource in guiding a forensic clinician through the legal research process and suggesting relevant legal materials. Forensic mental health professionals who do not reside near law libraries may nevertheless benefit from visiting public, university, or courthouse libraries. Although these libraries generally do not house as many legal resources as law libraries, they typically contain sufficiently large legal collections to be of use to forensic clinicians interested in conducting legal research.

Mental health professionals interested in increasing their legal knowledge will certainly benefit from consulting print and electronic legal resources, but the value of seeking information and consultation from professional organizations and knowledgeable colleagues cannot be overstated. Organizations such as the AP-LS often provide opportunities to practitioners interested in augmenting their legal knowledge. AP-LS, for example, hosts workshops and an annual conference, offers a network of knowledgeable scholars and forensic clinicians, and provides a number of useful resources online. The AP-LS Web site provides links to more than a dozen external sources related to law and mental health. These external sites supply valuable information on topics such as mental health law decisions from the U.S. Circuit Courts of Appeal, landmark cases, and important legal updates germane to forensic mental health professionals. Similarly, the ALI, which is self-described as an organization dedicated to clarifying, modernizing, and improving the law, is another optimal resource for forensic clinicians. Mental health practitioners conducting FMHAs can obtain a wealth of legal knowledge by consulting resources on the

ALI Web site (www.ali.org) or contacting the organization directly. Finally, there are a variety of continuing education (CE) workshops that offer legally relevant information for forensic mental health professionals. For example, the American Academy of Forensic Psychology (www.aafp.ws), which is the training and education affiliate of the American Board of Forensic Psychology, sponsors a variety of CE workshops that focus at least in part on legal issues relevant to forensic practice.

In addition to networking and taking advantage of training opportunities through professional organizations, forensic mental health professionals interested in enhancing their legal knowledge can seek counsel from knowledgeable and experienced colleagues. Referring attorneys are an obvious legal resource, and forensic mental health professionals can gain guidance on relevant law from the referring attorney. Experienced forensic mental health practitioners, other attorneys, and other legal professionals can also serve as valuable resources for clinicians interested in enhancing their legal knowledge. Consulting with experienced colleagues about relevant law (both procedural and substantive), applicable psycholegal standards, proper assessment techniques, and other matters relevant to FMHAs is highly recommended and should be considered a best practice.

READING, INTERPRETING, AND SYNTHESIZING THE LAW

At this point, we hope we have made clear the importance of knowing the law, the structure of the law, and how to find the law. After finding the law, the next step is knowing what to do with it, which is often not as straightforward as it might appear. Reading and interpreting exactly what is being said by a court or legislature in a case or statute can be challenging. In fact, a common criticism levied against attorneys, judges, and lawmakers is that they make the law inaccessible to the lay individual. One such method of keeping the law at arm's reach or "hiding the ball"—a bit of law school jargon—is by the way they write the judicial decisions and statutes. Although a growing movement of lawyers advocate for "plain English" drafting, many persist with tradition, and of course all settled law and cases remain written in traditional style.

It is not uncommon for new students of the law, after first reading through a case, to be left with no idea of what they have just read. A statute may read like a list, within a larger list, containing its own sublists; a judicial opinion might read like one long run-on sentence in outdated prose or Olde English. Making sense of it all can prove to be a daunting task, even for those with formal legal training. However, as with other topics discussed in this chapter, there is a steep learning curve, enhanced by knowing a few tips and tricks about how to digest legal literature. The topic is large enough to fill an entire chapter or book by itself and has done so many times over (see, e.g., Oates & Enquist, 2011; Sloan, 2012). That level of detail cannot be matched in this brief section, but we can offer a few tips and guidelines.

Statutes typically are organized in hierarchical fashion. Thus, someone interested in a particular statute may have to look at the criminal code, divided into articles,

each of which has chapters, with subsections, and so forth. Once the statute being sought is identified, the first step is to read it to confirm that it is the correct statute. Even if some of the statutory details are skipped over, a quick read of the statute should, it is hoped, confirm that it is indeed the correct statute that was being sought. After confirming it is the correct statute, a few rules of interpretation are worth following. The first section of many statutes defines key terms that are used later in the statute, and being familiar with these definitions provides a context for interpreting the statute. Look for whether the elements in the statute are conjunctive (indicated by “and”) or disjunctive (indicated by “or”); whether a legal disposition rests on several elements being met (“and”) versus one of several elements (“or”) can affect the interpretation of a statute. Similarly, look for key terms like “may” and “shall.” Although it is tempting to read them over quickly, these are examples of legalese that confer specific requirements. An individual who “shall” do something violates the statute if he or she fails to do so. Toward that end, resist the temptation to infer a word or phrase’s meaning based on context. Although harmless while reading a novel, misinterpreting a legal term can alter the statute’s meaning. It might be a good idea to have a copy of a legal dictionary handy while reading the statute, just in case such a situation arises. Last, it is common for statutes to reference other statutes, and only by reading those other statutes is it possible to get a full picture of the statute at hand.

Unlike statutes, cases do not fit so neatly into a hierarchical format. Whereas statutes typically are published or revised holistically, such that everything fits together nicely, precedential cases are decided only as novel legal issues come before a court. Thus, it becomes the responsibility of the researcher to determine how a case or cases fit into the overall area of law. This synthesis of case law is one of the most challenging aspects of legal research for the novice and experienced researcher alike. However, convention dictates that cases are typically organized in a similar fashion, or that they at least contain certain elements. Knowing what to look for is an effective way to sift through a lengthy legal decision and take away only the most important information.

The text of major cases (also called opinions) can be divided into five categories: facts, procedural history, issue, legal reasoning, and holding; a possible sixth category is legal precedent (case law or statutory law affecting the decision), but this category can be subsumed under legal reasoning. The five categories also tend to appear in this order, although there may be a summary of information (often called a syllabus) provided by Westlaw or LexisNexis. The facts of the case vary in length but are typically as brief as possible while still conveying all necessary information to provide the relevant context. The procedural history consists of the legal history of the case since its inception, including the holding by the court of original instance, appeals filed, appeals granted, and the like. Although information about such things as pretrial motions, testimony, and evidence may straddle the line between facts and procedural history, in practice the distinction is often insignificant. It is the combined facts and procedural history that can help

inform readers about whether the case is relevant to their needs. Stated another way, the facts tell whether two cases are alike or different in key elements (e.g., crime, offense, charge), and the procedural history determines whether the case is binding or persuasive authority in a given jurisdiction.

The facts and procedural history contextualize a case. However, the issue, the legal reasoning, and the holding are what make each case unique and what the legal researcher is ultimately looking for when reading a case. The issue is a shorthand name for the legal question the court is answering in its review. Sometimes the judicial opinion will clearly state the issue (“The issue before the court is . . .”). In other instances, the court may talk about the case “turning” or “hinging” on, for example, “whether the court decides to interpret the law” in one way or another.

By the end of a case, the court has answered the legal issue in its holding. The explanation connecting the issue to the holding is the court’s reasoning. Depending on whether the researcher is looking at the case as potentially binding or persuasive authority, emphasis may be placed more on the holding or the reasoning. In the case of binding authority, the holding is the most important component of the case. Although even when looking for binding authority, the legal reasoning by which the court arrived at its holding should not be overlooked. To do so could risk misunderstanding or misusing the case or being caught off guard by citing a case that is not fully understood. When relying on a case as persuasive authority, the legal reasoning is as important as the holding. After all, if someone is trying to convince the court to adopt a certain line of reasoning, understanding how the court in a cited case applied its own legal reasoning is a necessary precursor.

DETERMINING IF THE LAW IS “GOOD LAW”

An important step in doing legal research is determining if the law is “good law.” An authority (e.g., case, statute, regulation) that is good law carries legal weight and can be safely relied on. Confirming that something is good law can happen at the front end of research by making sure the research is properly executed, or it can happen at the back end by making sure the law being relied on is still valid authority. In practice, individuals who are savvy at conducting legal research check the status of law on a rolling process throughout the research process, making sure that any time spent reading the legal literature is well invested. We suggest this approach, although we recommend at the very least that the validity check happen as a final step before submitting a report or appearing in court. An attorney or forensic mental health practitioner relying on outdated or “bad law” may be committing procedural errors, making mistakes in practice, or making faulty assumptions.

Although it is somewhat tautological to say as much, a legal authority is good law unless it has otherwise been made bad law. What is important to note is that a legal researcher need not—in fact, may be unable to—find history supporting or confirming that something is in fact good law. The important point here is that something is good law unless something is found to the contrary. Thus, a case

published very recently is likely to be good law, if for no other reason than there has not been time for it to be reversed, overruled, updated, or repealed. On a related note, a decision being old should not automatically raise the inference that it is bad law. Plenty of law remains settled for decades. Rather, an older decision or statute just means there is more time for potential history that needs review.

The process of determining whether a law is good law has been described by many different names, including cite-checking, verifying, confirming, and updating. However, Westlaw and LexisNexis have two proprietary terms for determining whether something is good law: *KeyCiting* (Westlaw) and *Shepardizing* (LexisNexis). The term *Shepardizing* refers to using Shepard's Citations, which was the first system by which cases and their subsequent history were tracked. Shepard's Citations provides all subsequent authority that cite a legal authority and also whether any court has overruled the original opinion, which would make it bad law. For purposes of this chapter, it is important to emphasize that Westlaw and LexisNexis provide all relevant procedural history at the click of a mouse, for any case, statute, or regulation. It should be obvious at that point whether the authority is still good law, but both Westlaw and Lexis provide tutorials and round-the-clock customer support if there are still questions about whether it is good law. Although some university and public libraries offer access to Westlaw and LexisNexis, these services may be unavailable to many people. As such, it is worth discussing how to determine whether something is good law without the aid of Westlaw or LexisNexis.

Earlier in the chapter, we discussed three types of law: statutory law, case law, and constitutional law. Constitutional law does not change unless a constitution is amended, which is a fairly rare occurrence, so determining whether you have good law is most relevant when it comes to statutory law and case law. With respect to statutes, a statute is good law unless it has been repealed or superseded by a new statute or declared unconstitutional by the court. A classic example of a statute being declared unconstitutional is the U.S. Supreme Court's decision in *Dred Scott v. Sandford* (1857), in which the Supreme Court declared a recent act of Congress to be unconstitutional. In this regrettable decision, the Supreme Court concluded that African Americans brought into the United States and held as slaves were not, in fact, citizens of the United States and therefore not entitled to protection under the U.S. Constitution.

To determine whether a statute has been repealed or superseded by a newer statute, a good starting point is determining when the statutory code was most recently updated. Then it is important to check which statutes have been updated or enacted since that update. This information is located in what is called the "pocket part" of the code. Essentially, because statutes change regularly and reprinting codes is expensive, many legal references are purchased as part of a subscription. As part of this subscription, the publisher sends out updates at regular intervals, usually either quarterly or annually. These updates come printed in a small booklet, cut to fit the size of the legal code, and the booklets conveniently fit within the back of the originally printed hardback code. Hence, the pocket part of the statute

is located in the back of the book, and it includes all of the updates since the code was last bound (and since the last pocket part was mailed by the publisher). Pocket parts should also include citations to any cases that have interpreted the statute. By reading those cases it is possible to determine whether any major court cases have addressed the statute.

As with statutory law, case law can become bad law when a new statute supersedes the case law or a new case replaces the old legal rule. However, evaluating whether a case is still good law is a bit more complicated than for a statute. As such, it is necessary to discuss some new terminology. A case is not good law if it has been reversed or overruled or if the statute on which it was based has been updated. The term *reversed* means that a higher court reviewed the specific case and disagreed with the lower court, *reversing* the lower court's decision. *Overruled* means that another case occurred at a later date but involved the same legal issue, and the court decided the issue differently than it had previously. A classic example is the *Brown v. Board of Education* (1954) desegregation decision, which overturned the "separate but equal" doctrine upheld in *Plessy v. Ferguson* (1896). The terms *reversed* and *overruled* sometimes are used interchangeably, but a good way to think of it is that *holdings* get reversed and *legal rules* get overruled.

To determine whether a case is still good law, start by checking the case's procedural history, which should be included at the end of the case. The case's procedural history would show whether the case was appealed and whether that appeal was granted, and, if so, whether a higher court upheld or reversed the lower court's decision. As another general rule, it is common to cite to the highest authority that heard a case. For example, if the results of legal research yield a trial court case but later research reveals that the case was ultimately reviewed by the U.S. Supreme Court, in almost all situations the Supreme Court case would be cited. This illustrates that checking the validity of a case is useful not just to make sure something is good legal authority but also to make sure that the highest legal authority for a given legal issue has been identified.

After investigating a case's procedural history, it is necessary to determine whether any statutes have been enacted that have changed the case from good law to bad law. This is a bit more arduous but basically involves reading the case and looking for any statutes the court references in the opinion. If the court does cite one or more statutes, then it must be determined whether that statute was related to a legal issue in the case. If the statute was related to a legal issue in the case or if the case turned on how the court interpreted the statute, it is necessary to see whether that statute has been updated (see earlier text for determining whether a statute has been updated). If the statute has been updated, it becomes necessary to interpret the change in light of the case, synthesizing the literature that has been amassed. Only at that point is it possible to determine whether the case is good law or has been overruled by the statute.

One complication has to do with the degree to which someone is relying on a given legal authority. As discussed earlier, some authority is binding, and other authority

is persuasive. Whether an authority is good law or bad law partially depends on whether it is being relied upon as binding authority or persuasive authority. Thus, if a court higher than the one that decided a particular case has reversed or overruled the decision, then the case at hand is bad law. If, however, a court in another jurisdiction has simply declined to adopt the approach taken by the case at hand, then the case is still good law, although there is now persuasive authority against the case. Most challenging is when one or more courts in the same jurisdiction and at the same level (e.g., trial level) have treated a case differently. Although courts are free to reach different decisions, there is some pressure and expectation for different courts on the same level to adhere to *stare decisis* and rule similarly. A judge wishing to do otherwise might try to distinguish the facts of the case at hand so the argument can be made that he or she is not answering the same legal question differently but rather answering a related but slightly different question.

SYNTHESIZING THE LITERATURE

Having found, digested, and ascertained the validity of the relevant statutes and cases, the final step involves synthesizing the authorities. Synthesizing may sound easy at first; after all, is it not simply putting together all of the information into a list of holdings? Sometimes synthesis is this simple, and by bringing together all of the relevant holdings and statutes on a legal issue, it is possible to infer the legal rules the courts will likely follow when addressing a given legal issue. However, it is not unusual to find multiple authorities that may seem to contradict one another. Based on the previous discussion about determining whether something is good law, the obvious answer in such a situation would be that something is bad law. But what if nothing has been overruled, superseded, or revised?

In such a situation, a series of questions may alleviate the confusion. First, are the cases and statutes in the same jurisdiction? Second, is everything still good law? If the answer to both of those questions is yes, then the most likely explanation is that there is some minor distinction between the cases or statutes. Stated another way, the cases are not, in fact, addressing the same legal issue. After all, very similar legal issues can be answered in entirely different ways. At the time of this writing, for example, the U.S. Supreme Court has concluded that the death penalty does not violate the Eighth Amendment, unless of course the defendant is intellectually disabled or was under the age of 18 at the time of the offense. This illustrates how what may have looked like one legal issue (i.e., “Is the death penalty constitutional?”) is in fact three separate issues (i.e., “Is the death penalty constitutional for intellectually intact adults?” “Is the death penalty constitutional for juvenile offenders?” “Is the death penalty constitutional for individuals who are intellectually disabled?”). Thus, synthesizing can sometimes feel more like reconciling, but the importance of this step cannot be overstated. After all, if all of the research is not correctly synthesized, resulting in some mistake of law, much of the research process will have been for naught.

CONCLUSION

Understanding how to access, find, interpret, and synthesize the law is a wide-reaching topic that can fill many volumes. The existence of countless books and legal resources devoted entirely to the subject of legal research attests to the complexity and breadth of material that we could have covered in this chapter. In this chapter, we have provided a broad outline designed to help forensic mental health professionals become more comfortable and efficient in conducting legal research, which we believe will enhance the quality of their forensic work.

REFERENCES

- 18 USC § 2340A (2013).
- American Board of Forensic Psychology. (n.d.). *The written examination in forensic psychology*. Retrieved from <http://www.abfp.com/pdfs/certification/WrittenExamination.pdf>
- American Board of Psychiatry and Neurology. (n.d.). *Initial certification in forensic psychiatry*. Retrieved from http://www.abpn.com/cert_fp.html
- American Psychological Association. (2013). Specialty guidelines for forensic psychology. *American Psychologist*, 68(1), 7–19. doi:10.1037/a0029889
- Atkins v. Virginia, 536 U.S. 304 (2002).
- Brown v. Board of Education, 347 U.S. 483 (1954).
- Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993).
- Dred Scott v. Sandford, 60 U.S. 393 (1857).
- Federal Rules of Evidence 702.
- Frye v. United States, 293 F. 1013 (D.C. Cir. 1923).
- Grisso, T. (1986). *Evaluating competencies: Forensic assessments and instruments*. New York, NY: Plenum Press.
- Grisso, T. (2003). *Evaluating competencies: Forensic assessments and instruments* (2nd ed.). New York, NY: Kluwer Academic/Plenum Press.
- Heilbrun, K. (2001). *Principles of forensic mental health assessment*. New York, NY: Kluwer Academic/Plenum Press.
- Heilbrun, K., Grisso, T., & Goldstein, A. M. (2009). *Foundations of forensic mental health assessment*. New York, NY: Oxford University Press.
- Morse, S. J. (1978). Crazy behavior, morals, and science: An analysis of mental health law. *Southern California Law Review*, 51, 527–654.
- Oates, L. C., & Enquist, A. (2011). *Just research* (3rd ed.). New York, NY: Aspen.
- Plessy v. Ferguson, 163 U.S. 537 (1896).
- Rogers, R., & Ewing, C. P. (1989). Ultimate opinion proscriptions: A cosmetic fix and a plea for empiricism. *Law and Human Behavior*, 13, 357–374.
- Rogers, R., & Schuman, D. (2000). *Conducting insanity evaluations* (2nd ed.). New York, NY: Guilford Press.
- Simon, R., & Gold, L. (Eds.) (2004). *Textbook of forensic psychiatry*. Washington, DC: American Psychiatric Publishing.
- Sloan, A. E. (2012). *Basic legal research: Tools and strategies* (5th ed.). New York, NY: Aspen.

- Slobogin, C., Rai, A., & Reisner, R. (2009). *Law and the mental health system: Civil and criminal aspects* (5th ed.). St. Paul, MN: Thomson/West.
- Thompson, D. C., & Wachtell, M. F. (2009). An empirical analysis of Supreme Court certiorari petition procedures: The call for response and the call for the view of the solicitor general. *George Mason Law Review*, 16, 237–302.
- Tillbrook, C., Mumley, D., & Grisso, T. (2003). Avoiding expert opinions on the ultimate legal question: The case for integrity. *Journal of Forensic Psychology Practice*, 3, 77–87.

CHAPTER 4

Practicing Ethical Forensic Psychology

IRVING B. WEINER AND ALLEN K. HESS

PROFESSIONS derive considerable benefit from having a set of principles, standards, and guidelines that constitute their code of ethics. A code of ethics serves important purposes for professionals and the public they serve. For practitioners, an ethics code provides a moral compass to guide their conduct and help shape their decisions in uncertain situations. It lends prestige to their profession, fosters pride in their allegiance to values and skills shared with their colleagues, and enhances their sense of professional identity. For the public, the existence of an ethics code fosters their trust in a profession, informs them about what they can expect from practitioners in the profession, and safeguards their welfare by deterring unethical professional conduct.

In this last regard, unethical conduct, whether intended or not, can result in serious consequences for professional practitioners. Even without violating criminal law, which could lead to being fined or jailed by court order, misbehavior that comes to the attention of an ethics committee or state board of psychology can provoke letters of reprimand, dismissal from membership in professional societies, and suspension or revocation of a license to practice. Having been penalized in these ways can prevent practitioners from securing liability insurance, being listed in directories and on insurance reimbursement panels, and, most critically for forensic psychologists, having credibility as an expert witness whose integrity will not be impugned in the courtroom.

Along with being potential sources of penalties when violated, ethics codes and the committees and boards that enforce them can provide support when practicing psychologists confront problematic expectations or unreasonable demands. In an illustrative case, a patient whose therapist assigned her some reading complained to an ethics committee that he was abandoning her and failing to provide adequate care. Acting on the psychologist's behalf, the committee that heard the patient's complaint informed her that bibliotherapy was a legitimate and acceptable way for

her therapist to augment her sessions in an effort to be helpful, and the complaint was not pursued further.

Unreasonable or intrusive demands may on occasion come from supervisors, employers, or organizations. As an example in a correctional setting, a prison psychologist was instructed to serve on disciplinary committees that were hearing his patients' cases and deciding their punishments. By referring to his ethical obligation to avoid problematic multiple relationships and conflicts of interest, as elaborated in this chapter, the psychologist was able to clarify the impropriety of his serving on these committees and was excused from doing so. A recurring and sometimes problematic demand in forensic practice involves psychological testing protocols that are subpoenaed and deemed discoverable by the court. In these instances, psychologists can refer to their ethical obligation to "make reasonable efforts to maintain the security of test materials" (American Psychological Association [APA], 2002, Standard 9.1) as well as to copyright laws safeguarding published tests, to appeal to the court that the released information should have restricted circulation and consist of test data but not test forms.

To practice principled forensic psychology, psychologists must recognize both the constraining and the supportive features of ethics codes and related guidelines and conduct themselves in a professionally proper manner at all times. The main document providing psychologists direction for ethical conduct is the just mentioned "Ethical Principles of Psychologists and Code of Conduct" promulgated by the APA (2002, 2010a). This document, commonly referred to as the Ethics Code, pertains broadly to psychological practice, teaching, and research but also has numerous implications for forensic practice. The Ethics Code comprises five general principles and 10 specific standards. The general principles are considered "aspirational in nature" and are "intended to guide and inspire psychologists toward the very highest ethical ideals of the profession" (APA, 2002, p. 1062). The specific standards translate these aspirations into mandatory and enforceable rules of conduct that psychologists are obliged to follow.

Two other documents providing guidance in practicing principled forensic psychology are the "Specialty Guidelines for Forensic Psychology" and the "Guidelines for Child Custody Evaluations in Family Law Proceedings," both of which are modeled after the Ethics Code and are endorsed by the APA (2010b, 2013). Of particular relevance for psychologists involved in family law cases are two additional documents promulgated by the APA (2012a, 2012b, 2013): "Guidelines for the Practice of Parenting Coordination" and "Guidelines for Psychological Evaluation in Child Protection Matters." Like the general principles in the Ethics Code, the forensic specialty, child custody, parenting coordination, and child protection guidelines are not mandatory rules of conduct and were not intended to constitute a basis for disciplinary action or legal liability. Instead, these guidelines are aspirational in nature and recommend various procedures for maintaining high-quality forensic services. Forensic practitioners should be as familiar with these recommended procedures as they are with the obligatory rules of conduct in the Ethics Code.

This chapter reviews the five general principles and 10 specific standards in the APA Ethics Code, with attention to their implications for forensic practice and how they may be elaborated in the forensic specialty and child custody guidelines. The chapter then considers some aspects of personal values and professional responsibility that forensic psychologists need to keep in mind, and it concludes with a summary of recommendations for minimizing vulnerability to ethical complaints and legal actions in forensic practice.

ETHICS CODE: GENERAL PRINCIPLES

The section on general principles in the APA Ethics Code delineates five aspirational goals toward which psychologists should strive in their practice, teaching, and research.

1. *Beneficence and malfeasance.* Psychologists should safeguard the rights and welfare of those to whom they provide services and maintain vigilance to ensure that their influence is not misused. They should strive to benefit those with whom they work and avoid doing harm, and they should recognize any adverse effect of their own physical and mental health on the services they provide.
2. *Fidelity and responsibility.* Psychologists should establish trusting relationships with their clients, clarify their professional roles and obligations, and coordinate services with other professionals to each client's benefit. They should in addition attend to the ethical probity of colleagues and provide some measure of pro bono service.
3. *Integrity.* Psychologists should promote truthfulness in research, teaching, and practice and avoid dishonesty, deception, subterfuge, and misrepresentation of fact. Should they deem any deception justifiable, they should consider carefully whether it is necessary, whether the benefits of the deception outweigh any adverse consequences it might have, and what steps should be taken to minimize or repair any resulting harmful effects of the deception.
4. *Justice.* Psychologists should allow equal access to their services by all people, whether advantaged or disadvantaged and whatever their background, and they should provide services of equal quality to all. Psychologists should take reasonable care to prevent any biases or limitations of their competence from leading to improper or inadequate practices on their part.
5. *Respect for people's rights and dignity.* Psychologists should respect the dignity and worth of all people and their rights to privacy and autonomy. This respect should extend to persons with diverse backgrounds, including diversity related to age, gender, gender identity, race, ethnicity, national origin, religion, sexual orientation, disability, and socioeconomic status. Psychologists should neither condone nor participate in discriminatory practices based on such individual differences.

When applied in individual cases, these principles overlap in many respects, and a case illustrating any one of these principles usually illustrates one or more other principles as well. Similar overlap characterizes the 10 standards in the Ethics Code discussed next. For extended discussion of how these principles and standards interrelate and how they should be applied in various contexts, readers are referred to contemporary texts by Bersoff (2008), Kitchener and Anderson (2011), Knapp and VandeCreek (2012), and Koocher and Keith-Spiegel (2008).

ETHICS CODE: SPECIFIC STANDARDS

As previously stated, the APA Ethics Code comprises 10 specific standards intended to serve as enforceable rules of conduct that psychologists are obliged to follow. Unlike the general principles to which psychologists should aspire, these standards constitute requirements they are expected to meet in order to remain in compliance with the Ethics Code. The standards deal broadly with considerations in practice, teaching, and research, and they vary in their specific implications for practicing principled forensic psychology. However, forensic psychologists are psychologists first and forensic specialists second, and they are accordingly responsible for ensuring that all of their professional actions are consistent with the Ethics Code.

STANDARD 1: RESOLVING ETHICAL ISSUES

The first standard in the Ethics Code provides instructions for resolving conflicts between ethical considerations and the expectations or demands of legal, governmental, or organizational entities. A key provision in the 2002 Ethics Code concerning such expectations and demands was amended in 2010. This provision had specified that, when faced with ethical conflicts that could not be resolved, “psychologists may adhere to the requirements of the law.” Many psychologists viewed this provision as appearing to endorse the so-called Nuremberg plea that commission of evil acts was acceptable “because I was following orders.” To protect psychologists from feeling obliged to comply with demands to engage in inhumane conduct and to prevent psychologists from appealing to “orders from above” to justify such conduct, this provision was replaced with “Under no circumstances may this standard be used to justify or defend violating human rights.”

Other parts of the standard on ethical issues address procedures to follow should psychologists suspect ethical violations by another psychologist. Ordinarily in such instances, an informal contact with the suspected violator should precede filing any formal complaint, and filing such complaints should be considered when informal resolution seems inappropriate or would violate confidentiality rights or after collegial discussion has failed to resolve the matter of concern. Even with such clear directives to follow, however, ethical issues arising in forensic psychology practice may evoke respectable differences of opinion that call for some exercise of judgment, as in the next case.

Case 1. A psychologist was asked by the attorney representing a father in a child custody dispute to review a report summarizing an evaluation of him that was prepared by a psychologist retained by the mother's attorney. The reviewing psychologist was initially impressed with the quality of the examining psychologist's report, which included sufficient language to justify casting numerous aspersions on the quality of this father's character. On closer inspection, however, the reviewer became concerned by how much of the examiner's language appeared to be taken directly from a computer-generated test report. The reviewer was aware that narrative statements in computerized test reports warrant forming hypotheses but not drawing definite conclusions, because they describe characteristics of people in general who have certain scores on a test and do not necessarily describe the person who was examined (see Butcher, 2013). Upon printing his own computer-based narrative from the examining psychologist's test data, the reviewer confirmed that this examiner had indeed cribbed long sections of his report from the computer narrative without acknowledging their source. Moreover, he discovered that a *not* in the computer narrative had been omitted from the examining psychologist's written report at a critical point in the text where it would have reversed negative conclusions in the report about the father's suitability as a parent. The reviewing psychologist testified to this effect in court, with the computer-generated report and the examining psychologist's written report in evidence. The court concluded that the examining psychologist had misrepresented his findings and altered his report to please the attorney who had retained him. News of this outcome subsequently circulated among the local community of attorneys and psychologists working in family law cases.

Although the ethical transgression was clear in the preceding case, the sequence of events raises some procedural questions. To begin with, an informal collegial discussion of the matters of concern, prescribed in the Ethics Code as the first step in dealing with a suspected or potential ethical violation, was not an option in this instance, because the violation had already been written into a signed and discoverable document. Nevertheless, before reporting his discovery to the attorney who had retained him and appearing on the witness stand, should the reviewing psychologist have informed the examining psychologist about finding the missing *not* in his report? This consultation would have given the examining psychologist an opportunity, prior to being confronted in the courtroom, to apologize for an unintended oversight (even if this explanation would have rung hollow) or to prepare some defense of his actions (if he could think of one). As another possibility, should the reviewing psychologist, instead of or in addition to giving testimony, have brought the matter to the attention of the state board of psychology or the ethics committees of any professional organizations of which the examining psychologist was a member? Did the damage to the examining psychologist's future credibility as an expert witness constitute a sufficient penalty for his misrepresentation, or should the involved parties have pressed for further sanctions, such as prosecution for perjury or actions against his license? Such matters of judgment commonly

arise in forensic cases, despite the guidance of unambiguous principles, and equally principled psychologists might for various reasons answer the preceding questions in different ways.

STANDARD 2: COMPETENCE

The standard on competence addresses psychologists' obligations to provide services only within the boundaries of their abilities as gleaned from their education, training, study, relevant supervision, and professional experience. They can serve populations or employ methods unfamiliar to them only after sufficient preparation, consultation, and/or supervision to ensure professional competence. Only in emergency situations, when other qualified help is not available, should they act beyond the boundaries of their competence. Psychologists are further instructed by the competence standard to maintain their competence, which speaks to the importance of continuing education; to base their work on established scientific and professional knowledge in their discipline; and to prevent personal problems and conflicts from interfering with adequate performance of their work-related duties.

In addition to the numerous implications of these requirements for forensic practice, the competence standard specifies that psychologists serving in forensic roles should be reasonably familiar with judicial and administrative rules governing these roles. No psychologist can be adequately educated for every exigency, however, and determining whether competent service is being provided may at times involve some degree of judgment. The next case illustrates effective preparation for a forensic consultation and appropriate candor about professional competencies.

Case 2. A psychologist was asked to analyze some insurance policies for their readability levels. He was experienced in assessing intellectual functioning, but his familiarity with readability analysis and insurance policies was limited. He related his background and skills to the lawyer who wanted to retain him and clarified the nature of the services he was being asked to provide. He then consulted with two colleagues who regularly assessed readability level and reviewed some publications recommended by them. He was then able to describe to the lawyer the procedures he would follow in the analysis and to indicate that he had expert consultation available to assist him as needed. On this basis, having initially been candid about his capabilities and then taken steps to become adequately prepared for his task, he was able to provide a helpful analysis of the readability level in the insurance policies.

The competence standard also bears directly on forensic psychology in requiring practitioners to establish firm ties between their practices and the scientific bases for these practices. Chapter 22 in this volume by Otto, Kay, and Hess emphasizes basing court-related activities—including consultations, administrative hearings, and depositions as well as appearances on the witness stand—on as firm an empirical

basis as possible. Tensions nevertheless arise between what forensic psychologists know for sure and about what they are asked to opine, and discrete decisions must often be made in the absence of absolute certainty. When told only that there is a 70% chance of rain, we must make a yes-or-no decision whether to take an umbrella, and forensic psychologists are regularly asked for similarly discrete opinions (e.g., insane or not, competent or not, suitable parent or not) in the absence of 100% conclusive evidence one way or the other. Ordinarily the most principled course of action in such instances is offering a reasonable estimate, as in reporting a slight, moderate, or considerable chance of reoffending by a person being considered for probation or parole and letting the court or administrative body address the dichotomous decision that is properly its to make. In this vein, Monahan and Steadman (1996) advised mental health law specialists to model the way meteorologists announce a 70% change of rain, as by responding to a question about violence risk with a percentage estimate of its likelihood in a particular case.

Because competence is a critical consideration for forensic psychologists and one that often undergirds an ethics complaint or a lawsuit, the forensic specialty, child custody, parenting coordination, and child protection guidelines all pay considerable attention to it. In addition to paralleling many of the Ethics Code specifications concerning competence, the forensic specialty guidelines stress acquiring and maintaining not only knowledge but relevant skills as well and becoming knowledgeable not only about the legal system but also about the legal rights of individuals. These guidelines further admonish forensic psychologists to:

- Give full and accurate accounts of both their skills and their knowledge when presenting themselves to clients.
- Base their opinions and testimony as much as possible on scientific foundations.
- Clarify the limits of their conclusions and recommendations as well as the factual basis for them.
- Keep abreast of developments in the law as well as in psychology.
- Ensure that their own values do not compromise the services they are rendering. (Standards 2.01–2.07)

The child custody, parenting coordination, and child protection guidelines supplement the Ethics Code by defining *competence* in these areas of endeavor to include having a thorough understanding of child and family development and psychopathology, being familiar with what is known about the impact of divorce on children, and keeping current with the child custody literature.

As attested by this specification in the child custody guidelines, the competence of forensic psychologists is unlikely to be uniform across criminal, civil, and family law proceedings. A forensic practitioner who is expert in conducting competency and insanity evaluations in criminal cases may not be familiar with the literature, techniques, and considerations relevant to a fitness-for-duty evaluation, such as confidentiality issues specific to the workplace and prediction questions central to

personnel evaluations. Likewise, a psychologist with expertise in criminal cases may not know enough about the technicalities of probate and estate management to provide useful psychological consultation in a civil case involving alleged malfeasance by an estate administrator or executor.

As a stark example of expert testimony that was incompetent by virtue of inadequate foundation, Golding (1990) described the testimony of Dr. James Grigson, a psychiatrist, whose work formed the basis for the U.S. Supreme Court's analysis in *Barefoot v. Estelle* (1983). Without having evaluated Barefoot, Grigson claimed that he could "predict future dangerousness of an individual within reasonable medical certainty," that Barefoot was in the "most severe category of sociopaths," and that he (Grigson) was "one hundred % [sic] and absolutely [certain that] Barefoot would commit future acts of violence that would constitute a continuing threat to society" (Golding, 1990, p. 291). Cross-examination has been the mechanism traditionally used by the courts to challenge or dismiss such outlandish claims. However, cross-examination is often inadequate in this regard, because it relies on attorneys, judges, and jurors to make determinations concerning technical matters that are known only by other experts in the area. As a partial solution to this problem, Golding suggested placing an affirmative duty on potential experts to clarify the basis of their testimony, to show the evidence of the specific bases of their conclusions, and to describe the limits of their testimony, all of which would be consistent with the Ethics Code and forensic specialty guidelines. Unfortunately, this suggestion would still leave it in the hands of attorneys to conduct telling cross-examinations on technicalities of psychology, which they may not be equipped to do.

As another potential solution to instances of incompetent expert testimony, Shuman and Greenberg (1998) commented that many such problems could be resolved from the bench if judges would take professional ethics codes more seriously than they apparently do. In *Barefoot*, for example, Grigson's testimony passed judicial review but not psychiatric peer review. An American Psychiatric Association amicus brief asserted that Grigson should not have been permitted to testify in *Barefoot v. Estelle* (1983) because it is "unethical for a psychologist [sic; Grigson was a psychiatrist] to offer a professional opinion unless he or she has conducted an examination" (Shuman & Greenberg, 1998, p. 7).

Shuman and Greenberg (1998) also decry instances in which courts accept expert testimony in custody cases from a psychologist who has examined only one parent and offers an opinion about which of two parents would better serve the needs of the child. Aside from being patently unsound—how can a comparative opinion be formed when only one parent has been seen?—this type of practice, like testifying about the psychological characteristics of a person who has not been examined, violates the Ethics Code except in special circumstances, such as providing a record review concerning someone who is not available to be examined (APA, 2002, Standard 9.01 (b) (c); see also APA, 2012a, 9.03).

As a further example of questionable procedure, consider the courtroom performance of the two psychologists testifying in the next case.

Case 3. In a murder trial involving an insanity defense, a psychologist retained by the defense attorney testified that impairments associated with borderline personality disorder had caused the defendant lifelong adjustment problems, including mood fluctuations, excessive anger, and poor emotional control. This disorder, the defense psychologist said, had prevented the defendant from conforming his conduct to the law and distinguishing right from wrong when he stabbed to death a 74-year-old woman and her 4-year-old granddaughter. The psychologist offered this testimony despite the existence of a letter found in the purse of the defendant's ex-wife outlining the steps she was to take in helping him construct a convincing insanity defense. Could the defendant have been legally insane at the time of the stabbings but subsequently been able to become involved in the planning of his defense? This interesting question became moot, however, because the judge ruled the letter inadmissible due to the marital privilege—even though the couple had divorced a year before the murders. (Note: The judge's questionable ruling in this regard aside, borderline personality disorder has no recognized implications for impairing a person's ability to distinguish right from wrong.) The psychologist retained by the prosecutor testified that the defendant knew the nature and quality of the charges against him (Note: This is a competency criterion, not a sanity criterion.) and was able to distinguish right from wrong because voices he claimed to hear were coming from inside and not outside his head. (Note: This may have implications for whether the defendant was hallucinating, but it has no direct relevance to his knowing right from wrong at the time of committing the offense.)

Some questions about the competence of the two expert witnesses in this case come quickly to mind. Were they familiar with the applicable legal criteria for insanity? Did these criteria, like the laws in most states, disallow personality disorders from consideration as a valid insanity defense? Did these psychologists employ adequate methods to assist them in reaching their conclusions? (Both based their testimony solely on a clinical interview with the defendant, without reviewing historical documents, talking with collateral informants, or conducting psychological testing.) Did the defendant's claim of hearing voices raise some doubt about the accuracy of a borderline personality disorder diagnosis? Did one or both of the psychologists have an adequate foundation for their testimony, and did they articulate their findings in a manner consistent with the standard set forth in the Ethics Code? The probable answers to these questions bring into sharp relief the kind of inept courtroom display that can foster public cynicism and mistrust of mental health experts. Such incompetence is by definition unethical (see Weiner, 1989), and it is unethical conduct in the courtroom, not proper forensic practice, that provides critics of forensic mental health practice with grist for their mill.

Finally, with respect to competent practice, forensic psychologists should recognize and express clearly the certainty of their opinions and conclusions. In the insanity case just discussed, the two experts could have drawn on definite and well-established criteria to guide their evaluation of the defendant, these being the two "prongs" of insanity as defined in statutory law. The first of these prongs,

which is universally applicable in U.S. jurisdictions, invokes a *cognitive* incapacity that prevented the accused from recognizing the criminality of his or her actions at the time of the alleged offense or that prevented the defendant from appreciating the legal or moral wrongfulness of this conduct. The second prong, which is an alternative statutory criterion for insanity in some but not all jurisdictions, is *volitional* in nature and speaks to a loss of behavioral control such that the accused was unable to alter or refrain from his or her criminal conduct at the time. (See Zapf, Golding, Roesch, & Pirelli, Chapter 12 this volume.)

In other situations, forensic psychologists may have to form their opinions with less certain criteria to draw on than are available when insanity is the issue. In jury selection, for example, a consulting psychologist who can assess a potential juror's opinions regarding the issue at trial may have little way of knowing how this person might be influenced by the leadership style of the jury foreperson and the group dynamics among the jurors when they assemble in the jury room. In practice, this circumstance means that forensic consultants are likely to be more helpful in *juror* selection than they can be in *jury* selection.

Because of the uncertainty inherent in many of the conclusions they form, forensic psychologists are well advised to work with a graded system of levels of confidence in their conclusions. These levels of certainty should cover a broad range of possibilities that can be expressed in such terms as "reasonably certain," "strongly suggestive," "some indications but not compelling," and "the available data do not warrant an expert opinion on this matter." The level of certainty chosen should be based on the state of knowledge in the areas, the degree to which the particular case touches on definite provisions in the law, and the amount and clarity of available information pertaining to the legal considerations in the case. In the instance of an insanity defense, for example, the law provides clear criteria that can readily be translated into well-researched psychological characteristics for which clinical research and assessment methods are abundant. In custody cases, by contrast, statutory law provides little in the way of specified criteria for being a suitable parent.

Forensic psychologists are further well advised to discuss with attorneys who retain them not only the certainty of the knowledge base relevant to their case but also the certainty of the findings they obtain in the particular case and can testify to. Even the most sensitive measurement methods may not reveal relevant information in a particular instance, and psychologists should say as much in such cases. Competent forensic practitioners take responsibility for describing the weight of the evidence on which their testimony is based, and, even in the absence of a compelling level of certainty, the court may ask for expert guidance, as in the next statement from the bench.

It is, of course, not easy to predict future behavior. The fact that such a determination is difficult, however, does not mean that it cannot be made. Indeed, prediction of future criminal conduct is an essential element in many of the decisions rendered throughout our criminal justice system. . . . What is essential is that the jury have before it all possible relevant information about the individual defendant whose fate it must determine. (*Jurek v. Texas*, 1976, 274–276)

STANDARD 3: HUMAN RELATIONS

The Ethics Code standard on human relations speaks to practicing, teaching, and researching psychology in ways that are considerate of others and avoid causing them harm to the degree possible. Included in the several sections of the standard are proscriptions against discriminating against, harassing, or exploiting others and guidelines for not becoming involved in multiple relationships and conflicts of interest. Forensic activities often put psychologists in situations fraught with risk for being drawn into multiple relationships and conflicts of interest, and the forensic specialty guidelines spell out numerous precautions that practitioners should take in this regard. The guidelines have in turn been supplemented by substantial attention in the literature to the boundaries of and exceptions to multiple relationships (Greenberg & Shuman, 2007; see Knapp & VandeCreek, 2012, chap. 6).

As an important aspect of protecting the well-being of persons they evaluate in forensic cases, psychologists should take care to prevent their findings from becoming available to persons who should not receive and do not need to have this information. The next case illustrates one examiner's consideration in minimizing harm that might have resulted from his providing necessary information in open court.

Case 4. A psychologist conducting a defendant's competence to stand trial evaluation found reasonably clear evidence that the defendant was intellectually disabled. He informed the defendant that he would be saying some things about him in court that might be upsetting to him and make him look bad in the eyes of friends and relatives who were present but that he should not take these comments as a personal criticism. He then told the defendant what the result of intellectual testing indicated, but he added that there were many different ways of defining and measuring intelligence and that these test results did not say anything about whether he was a good or bad person. This sensitivity on the psychologist's part fostered mutual respect between him and the defendant and gave the defendant a frame of reference that eased the distress of hearing his limited tested intelligence reported in the courtroom.

STANDARD 4: PRIVACY AND CONFIDENTIALITY

Standard 4 concerns the responsibility of psychologists to respect and protect the privacy of the people to whom or for whom they provide professional services. The forensic specialty guidelines clarify that this requirement pertains both to clients being evaluated and to retaining parties who request a forensic psychological consultation (APA, 2013, Standard 8). Because of the extent to which forensic psychologists are charged with unearthing and reporting sensitive findings, issues of privacy and confidentiality can become particularly complex and challenging in their practice.

The key questions for forensic psychologists in this regard concern the kinds of information they should release (e.g., summary impressions, full reports, case notes, test results, computer printouts), in what form this information should be released (written documents, electronic transmission, oral communication), to whom the information should be released (the court, a retaining or opposing counsel, the person evaluated, the media), and the circumstances under which information should be released (a court order, a subpoena, an informal request). The answers to these questions vary, as prescribed by the applicable regulations in a state. The overarching consideration established by the Ethics Code is that psychologists should release information only to an individual or entity identified in a written consent form signed by the person evaluated, except when the law requires or permits proceeding otherwise. In the absence of such a signed release or legal requirement or permission, any disclosure of information is a violation of privacy rights (APA, 2002, Section 4).

Issues concerning the preservation of privacy have been complicated in contemporary times by technological advances in modes of information exchange. In a recent survey, McMinn, Bearnse, Heyne, Smithberger, and Erb (2011) found considerable uncertainty among several hundred psychologists regarding the ethicality of using certain communication technologies in professional practice. Over one-third of their respondents reported being uncertain about whether it is ethical to use e-mail to provide professional services, contact clients about payment or insurance issues, and send confidential information to other health-care professionals, and more than one-third were uncertain about the ethicality of using instant messaging to provide consultation and supervision, either by computer or cell phone. These survey findings speak to the point made throughout this chapter that ethical right and wrong sometimes hinges on relative rather than absolute judgments, and indefinite circumstances can allow for respectable differences of opinion concerning proper conduct. The following case, however, illustrates a confidentiality requirement that is unequivocal with respect to what must be done—the presence of a nondisclosure agreement.

Case 5. An attorney asked the psychologist to assess the extent of psychological damage in a 32-year-old woman who had been sexually assaulted by an appliance repairman. Records indicated that the employer had previously received complaints about the repairman's conduct when he was doing clerical work in their central office. As their "solution" to his being intolerably obnoxious in the central office, the company assigned him to go to people's homes by himself, without supervision, and service their malfunctioning appliances. The consulting psychologist found evidence of substantial psychological dysfunction in the plaintiff, probably attributable to the sexual assault, and also some additional psychic harm apparently resulting from harsh depositions conducted by the company's attorneys. The woman's attorney reviewed these findings with the company's ethics officer, following which the psychologist was paid for his efforts but received no further

information about the outcome of the case. He inferred from this sequence of events that the company, faced with the compelling evidence of damage and its proximate causes, had made an out-of-court settlement with the victim and that the settlement included a nondisclosure agreement. The point of this case example with respect to disclosure and privacy is that, with a nondisclosure agreement in place, any use of the case that contained identifying information could void the settlement and make the psychologist vulnerable to legal action by the plaintiff, her attorney, and the company. In such circumstances, everyone working with the psychologist on a case, including secretaries and technicians with access to information about it, is similarly bound by the privilege rights of the attorney's client and civilly liable for any breach of privacy.

STANDARD 5: ADVERTISING AND OTHER PUBLIC STATEMENTS

Standard 5 in the Ethics Code indicates that psychologists are free to make public statements about their professional qualifications and to express their professional opinions on matters related to psychology. In so doing, however, they must avoid making any false or deceptive statements. Although they can advertise their credentials and availability as service providers, they must not hire media personnel to provide publicity for them, nor should they solicit public testimonials about their skills from former therapy patients. When they do advertise, forensic psychologists are likely to benefit most from presenting themselves in media to which attorneys attend, given that attorneys are commonly their major referral source. Similarly, most forensic practitioners will find it to their advantage to present a continuing education workshop for attorneys. Local and state bar associations usually welcome such offerings by psychologists, although psychologists need to be modest in what they promise to provide and not guarantee any beneficial outcomes of participating in their workshop.

The forensic specialty guidelines do not speak to advertising, but they do include a section that elaborates several forensically relevant considerations in making professional and public statements. Among these are being accurate and avoiding deception in the presentation of opinions; providing appropriate disclosure of sources of information and the bases of these opinions; and being careful to differentiate among observations, inferences, and conclusions (APA, 2013, Standard 11.01–11.04). The specialty guidelines also caution forensic practitioners to be fair, impartial, and respectful should they have occasion to comment on or criticize the conduct of other professionals or participants in a legal proceeding. Similar restraint should characterize any out-of-court statements by forensic psychologists about legal proceedings in which they are involved (APA, 2013, Standards 11.05–11.07). These particular specialty guidelines overlap somewhat with the Ethics Code requirements concerning privacy, confidentiality, and privilege, and they also relate closely to the chapters in this volume on writing forensic reports (Weiner, Chapter 21) and serving as an expert witness (Otto, Kay, & Hess, Chapter 22).

STANDARD 6: RECORD KEEPING AND FEES

The text of Standard 6 in the Ethics Code obliges psychologists to facilitate the services they provide by documenting their work with adequate case records and making these records available to other professionals, as appropriate and within the constraints of confidentiality. The Ethics Code does not specify the length of time that case records should be kept, but psychologists should be aware of applicable regulations in this regard in the states in which they practice. In Florida, for example, statutory law requires licensed psychologists to retain complete psychological records for each service user for a minimum of 3 years and either the complete records or a summary of them for an additional 4 years. The APA (2007) provides some additional nonmandatory guidelines to assist practitioners in managing their records in the individual case, and the specialty guidelines provide further specific direction in this regard (APA, 2013, Standard 10.07).

With respect to fees, this section of the Ethics Code includes several restrictive provisions with which forensic practitioners must be sure to comply. Psychologists must not withhold records for nonpayment of services when these records are needed for emergency treatment of an individual who has been evaluated; they must refrain from bartering agreements in which goods or services are accepted as payment for fees, unless such nonmonetary remuneration is not clinically contraindicated and is not in any way exploitative; and they must not kick back to referral sources any portion of fees to which they are entitled on the basis of services they have provided.

The forensic specialty guidelines add to these restrictions a caution against contingency fees, which is an especially important consideration in personal injury litigations that could involve substantial settlements. Forensic psychologists should charge for their services on an hourly basis at a prearranged rate or at a prearranged amount for providing particular services. They should never charge on a contingency basis that depends on the outcome of a case. Charging on a contingency basis in a personal injury litigation when retained by the plaintiff's attorney can mean that the larger the settlement, the more the psychologist gets paid. It can also mean not getting paid at all, should the defendant prevail in the case. Operating in such contingency circumstances makes the psychologist an interested party with a clear conflict of interest who cannot be expected to provide objective and unbiased testimony.

STANDARD 7: EDUCATION AND TRAINING

The standard on education and training holds psychologists responsible for designing appropriate experiences and providing appropriate knowledge to accomplish the goals of any program in which they participate. These goals and the nature of the program should be described clearly and accurately to all interested parties. Teachers and supervisors must provide accurate information to their students and their supervisees, and they are additionally obliged to evaluate the performance

of these students and supervisees and give them timely feedback. These and other sections of the standard address general concerns in psychology and have no direct bearing on or implications for forensic practice. Nevertheless, psychology students can become involved in court proceedings, and their involvement can at times become an ethical minefield, with life-and-death consequences, as in the next case.

Case 6. A university professor offered his clinical psychology graduate students to the district attorney's office to evaluate defendants, which gave the students valuable but unpaid experience and saved the prosecutor's office thousands of dollars. In a case in which a mother was accused of killing her 4-year-old daughter, tests administered by one of these students were reported by the media as likely to help seal a first-degree murder conviction. When deposed, however, the student testified that she had been unsupervised and that the Multiphasic Personality Inventory, Second Edition (MMPI-2) she had administered was the first with which she had any experience. She also testified that she had been told by her professor to change some of the answers on the MMPI-2 based on the fact that the test had been administered in prison. The prosecutor's office claimed ignorance of students having been sent alone to conduct evaluations on first-degree murder cases. Nevertheless, as a result of this incident, eight other first-degree murder cases came under review, with one defense attorney claiming that the psychological testing had been critical in her client's accepting a guilty plea instead of proceeding with a jury trial, and the prosecutor's office discontinued its use of student examiners. As for the professor, he said subsequently that it was all a misunderstanding and that the student had become flustered during her deposition and made some misleading statements. However, he did not dispute the production of an altered MMPI-2 answer sheet. The ethical misconduct in this case requires no further comment.

STANDARD 8: RESEARCH AND PUBLICATION

With regard to research, Standard 8 in the Ethics Code delineates requirements for obtaining informed consent from research participants, protecting participants from any adverse consequences of participating in a study, and sparing members of a targeted research population from disciplinary or punitive action should they decline to participate. Also included in this section are considerations in determining appropriate inducements for research participants, whether and what types of deception are permissible, and the kinds of feedback that participants should be provided. With regard to publishing, the standard calls for giving appropriate credit for authorship, and it specifically proscribes plagiarism and the fabrication of data.

This standard is fully applicable to the scholarly work of forensic psychologists, whose research efforts may also present some specialty-specific ethical issues, particularly if they are collecting data within prison walls. For example, can informed consent be given by incarcerated research participants who as prisoners

may not be operating under their own free will? Might some of the requested research information compromise the rights of participants by being tantamount to their testifying against themselves? Alternatively, if researchers do acquire incriminating information from a participant, whether by intent or inadvertently, can they withhold this information from the authorities without risk of becoming an accessory? Does the researcher feel sufficiently comfortable and have the necessary familiarity with a particular prison's facilities and procedures to function effectively in that setting? Likewise, have all of the participants in the research project whose assigned tasks will bring them into the prison been adequately prepared for being behind bars in the company of accused or convicted criminals? These questions call for careful consideration of the circumstances in each instance, and there is no absolute or automatic answer that will inevitably be correct in every instance.

In professional presentations as well as publications, psychologists are obliged to take reasonable steps to avoid factual errors and misleading or unwarranted statements. The following case illustrates one psychologist's public failure to do the homework for which he was ethically responsible.

Case 7. A prominent forensic psychologist was presenting a paper at the annual APA convention on ethical considerations in serving as an expert witness. He emphasized the importance of practicing at the highest standards of empiricism, and he illustrated shortcomings in this regard by referring to a testing procedure developed by a well-known psychologist. He assailed the procedure's lack of reliability and validity and called its use improper. A few minutes into this harangue, a man in the audience rose and introduced himself as the psychologist being vilified. He then listed studies in several prominent journals that documented the psychometric soundness of his procedure and asked the presenter if he had read these studies. The presenter, sweating profusely, mumbled something inaudible and sat down without finishing his talk.

The lesson in this case is that public presenters should speak as if anyone about whom they will be speaking, especially if in critical fashion, were in the room. Comments that could not in good conscience be voiced in the presence of a professional colleague whose work is being criticized should not be voiced at all. Unlike the presenter in this case, moreover, a principled forensic psychologist would not impugn another psychologist's work without having conducted a thorough search of the literature concerning it.

STANDARD 9: ASSESSMENT

The assessment standard in the Ethics Code has substantial implications for the practice of forensic psychology, because assessment constitutes a considerable part of what forensic psychologists do. Questions of competence, criminal responsibility, and diminished capacity in criminal cases are basically assessment questions (see Zapf, Golding, Roesch, & Pirelli, Chapter 12; Zapf, Roesch, & Pirelli, Chapter 11;

and Clark, Chapter 13, all this volume), as are questions of competence and personal injury in civil cases (see Galietta, Garcia-Mansilla, & Stanley, Chapter 9 this volume, and Piechowski, Chapter 7 this volume), questions of child custody and parental suitability in family law cases (see Stahl, Chapter 6 this volume), and questions about violence risk (see Douglas, Hart, Groscup, & Litwack, Chapter 14 this volume).

The first three sections of the assessment standard define the essence of ethical assessment practice. First, psychologists should offer opinions only about persons they have examined, except when the person is not available for examination or when only a consultation or record review is necessary. Second, assessments should be conducted with reliable and valid methods that are appropriate for the purpose of the examination. Third, informed consent must be obtained from the person to be examined, unless the examination is mandated by law or is a routine part of an activity in which the person is a voluntary participant, such as pre-employment screening. Six other sections of this standard delineate considerations related to test construction, interpreting and explaining test results, releasing test data, maintaining the security of test materials, and using test scoring and interpretation services. Two further sections of the standard caution psychologists against using obsolete tests and outdated test results and against promoting the use of psychological assessment techniques by unqualified persons.

The forensic specialty guidelines echo the Ethics Code with regard to conducting assessments, and they elaborate four additional considerations that should guide forensic practice. These considerations include:

1. Focusing assessments on the legally relevant factors in each individual case.
2. Recognizing individual differences that could result in the same test finding having different implications for people from different ethnic or sociocultural backgrounds (see Weiner & Greene, 2008, chap. 4).
3. Providing the appropriate kind and amount of feedback to examinees in legal proceedings.
4. Retaining adequate documentation of assessment findings and making this documentation available as permitted or required by law (APA, 2013, Standard 10.01–10.06).

It may well be that ethical problems for the practitioner arise more frequently in conducting assessments than in any other forensic activity. A relatively recently emerging ethical problem that merits special mention concerns coaching. Most practitioners would consider it appropriate to accept an attorney's request to help coach a person in giving testimony, as in instructing a defendant to sit up straight and establish eye contact with the judge, prosecutor, and jurors. To the extent that slouching and looking down at the floor might send cues that the defendant was guilty or being untruthful, irrespective of the evidence, a psychologist coaching avoidance of these behaviors might well be serving the purposes of justice. However, what if an attorney representing a client in a personal injury or a disability determination case asks the psychologist to coach this client on how to appear impaired

on measures of neuropsychological functioning? Most psychologists would deem this type of coaching unethical, and coaching a person in how to answer certain kinds of questions and perform certain kinds of tasks has in fact been demonstrated to produce misleading results in neuropsychological and competency examinations (Gorny & Merten, 2007; Springman & Vandenberg, 2009; Suhr & Gunstad, 2007; Victor & Abeles, 2004).

STANDARD 10: THERAPY

The therapy standard deals mainly with restrictions on sexual intimacies between psychologists and their current or former psychotherapy patients and these patients' significant others. Other sections of the standard concern such matters as the necessity of informed consent by persons being provided psychotherapy, considerations in offering or providing therapy for persons already receiving mental health services elsewhere, and issues related to the interruption or termination of an ongoing treatment relationship.

For forensic psychologists providing psychotherapy to litigants, there is an inherent conflict between their obligation to protect a patient's privilege and the customary expectation that they will be disclosing pertinent information to officers of the court and other appropriately designated individuals. Whereas forensic practitioners conducting assessments rarely have difficulty identifying the person or entity requesting the evaluation as their client, and not the person being evaluated, deciding who the client is may not come so easily when they are functioning as psychotherapists, particularly when they are working in correctional settings or with parents involved in a custody dispute (see Monahan, 1980). With respect to working in a correctional setting, consider the dilemma of the psychologist in the following case.

Case 8. The psychologist was employed by the Department of Corrections with duties that included providing psychotherapy as part of a prison-based drug rehabilitation program for addicted inmates. During a psychotherapy session, an inmate mentioned in passing that he had "50 ccs of liquid cocaine stashed on the ward." The psychologist was aware that he was now faced with having to choose between maintaining prison security and preserving patient confidentiality. He could report what the inmate had told him, thereby sacrificing his trustworthiness as someone in whom the inmate could confide, or he could keep the inmate's secret, thereby endangering the drug-free status of the inmate and his ward-mates and colluding with him in a violation of prison rules.

As in this case, ethical dilemmas rarely have a simple and entirely comfortable solution. Possible harm could be done by making either of the choices listed, but perhaps more harm would ensue to more people from allowing the cocaine to remain on the ward than from losing the inmate's trust in his therapist, however much the psychologist would regret having to make this choice. Ofttimes the best solution to

a dilemma is avoiding it in the first place, whenever possible. There would have been a regrettable choice but not a difficult one in Case 8 if the psychologist had initially presented a clear statement of the roles and obligations of both parties to the treatment. Beginning psychotherapy with an explicit treatment contract is standard recommended procedure (see Weiner & Bornstein, 2009, chap. 6), but it can be overlooked when the therapy has been mandated, as in this instance. If the psychologist in this case had stated during the first treatment session that he would be obliged to report any violation of prison rules, there would have been no dilemma. Although he might have regretted having to sacrifice a confiding treatment relationship, his honesty and dependability in reporting the violation, just as he had said he would, might have sustained rather than undermined the inmate's trust in him. The next case illustrates effective implementation in a correctional setting of a pretreatment agreement intended to provide protection for the inmates as well as the institution.

Case 9. A prison psychologist was assigned by the Department of Corrections to conduct sex offender group psychotherapy, the purpose of which was to facilitate the offenders' successful adjustment upon returning to the community. A potential ethical dilemma lurking in this assignment concerned the group members' privacy rights beyond reporting their progress toward being considered for parole. The psychologist clarified to the group members at the outset that their general progress and risk of reoffending, which would include their compliance with prison rules, would be reported to the prison authorities, but other specific information they discussed in the group would be kept confidential. Concurrently, the psychologist made it clear to the authorities that content of the group sessions unrelated to general progress and risk of reoffending would not be forwarded and that privileged communication would remain in effect except for the general progress report.

Psychologists providing treatment for parents involved in custody disputes are also likely to face choices between reporting or withholding certain information, and they may at times have difficulty reconciling the specific ethical standard concerning privacy with the general principle of avoiding doing harm. Some potentially difficult decisions are taken out of the psychologist's hands by the law, as in the case of the requirement in all 50 states for professionals to report suspected child abuse (see Condie, Chapter 10 this volume). Usually, however, the confidentiality issue for therapists treating antagonistic parents is more murky than clear, and sometimes, as in the following case, even appropriate efforts to establish protective agreements can go for naught.

Case 10. A psychologist anticipated possible ethical conflict between his providing marital therapy and being called to testify in a bitter divorce and custody battle. He had been asked by an attorney whose marriage was dissolving to counsel his wife and himself, in hopes that, for the sake of their young son, they could find a way to salvage their marriage. The psychologist agreed to work with them on the

condition that whatever emerged during the marital therapy would not be accessible for inclusion in any subsequent legal proceedings. He explained to both parties that this condition was necessary to avoid either or both of them striving to look good in the psychologist's eyes and make the other one look bad, which would undermine efforts to resolve their marital problems. Both parties said that they understood and agreed with this condition. Unfortunately, the marital problems proved refractory to the treatment, and two months following termination of the therapy, the attorney informed the psychologist that he was going to be subpoenaed as a witness to attest to his wife's emotional instability. The psychologist reminded him of the pretreatment agreement, to which the attorney replied that this agreement carried no legal authority and could be revoked without consequence.

The psychologist dealt with this impasse by letting the attorney's lawyer know that, if called to testify, he would report to the court the husband's violation of the pretreatment condition and appeal to the judge that the wife had not waived privilege. If the wife were to waive privilege, the psychologist continued, neither party would be happy with the testimony he would then give in describing his impressions of both of them. The attorney's lawyer decided that it would be best to forgo calling the psychologist as a witness.

Sensible practice and case law do not preclude attorneys calling a psychotherapist as a fact witness. However, attorneys usually have no difficulty appreciating that privileged communication is essential for psychological treatment to be effective and that privacy calls for separating the therapist role from the expert witness role. Nevertheless, to serve their own purposes, attorneys may press a fact witness psychotherapist for expert testimony. In so doing, they create a role conflict for the therapist, who may be required by the court to disclose information that he or she would ordinarily be ethically obliged to keep confidential. However, the Ethics Code allows disclosure of confidential information without consent of the individual "when mandated or permitted by law for valid purposes" (APA, 2002, 4.05; see also Strasburger, Gutheil, & Brodsky, 1997).

VALUES AND RESPONSIBILITY

Forensic psychologists must deal regularly with two related aspects of principled practice that warrant further elaboration. The first of these concerns the obligation of practitioners to prevent their personal values from affecting their professional conduct. The second aspect concerns the professional responsibility of forensic psychologists to resist expectations or demands of attorneys that, although falling within the law, run counter to psychology's ethical principles and standards.

VALUES

The Ethics Code calls on psychologists to be cognizant of how their values as well as their limitations might affect their work (APA, 2002, Justice Principle &

Section 3.01). Forensic practitioners must be conscientious not only in presenting the factual bases of their opinions and their level of certainty about their inferences but also by being sensitive to the possible impact of their personal needs and beliefs on their interpretation of the data. The nature and extent of this value-based impact may not always be apparent, but being alert in every case to this possible threat to their integrity is no less important for forensic psychologists than being sufficiently competent to deal effectively with the substance of the case.

This is not to say that experts will conduct their professional affairs completely free of prejudices derived from their values. However, it is one thing for psychologists to accept a case involving issues about which they have some opinion but nevertheless strive to remain objective and neutral in their work on the case; it is quite something else for psychologists to seek out forensic opportunities in order to promote certain causes or advance some sociopolitical agenda, in which case their neutrality and objectivity would most certainly be suspect. As suggested by Hess (1998), experts should ask themselves several questions before committing to participate in a case in which their values might cloud their judgment: Would they be using the expert role to provide objective expertise or to exercise moral advocacy? Are they qualified by knowledge, skills, education, experience, or training to provide information that will be helpful to the fact finders in the case? Will their involvement in the case be probative (truly informative) or prejudicial (one-sided argumentative)? To what extent will the adversarial nature of the legal system distort the implications of whatever evidence they provide?

Hence the ethical concern is not whether psychologists serving in expert roles have values but whether these values constitute a dedicated purpose that compromises the objectivity and neutrality they bring to their evaluations and testimony in certain cases. Suppose, for example, that a male psychologist loses custody of his own child and subsequently becomes determined that no other father should lose custody of his child. In pursuit of this mission, he comes to believe that the father in every child custody case in which he consults should be awarded custody, and his reports always conclude that the father is the more suitable parent. This obvious intrusion of personal beliefs on professional conduct would have eroded the psychologist's integrity, damaged the examinees, and undermined the legal process as well.

Forensic psychologists are entitled to their personal beliefs, including a conviction that fathers should always be awarded custody. However, proper platforms for promoting such causes might include a presentation to a legislative body or participation in a sociopolitical action group, but they do not include testifying as an expert witness. Forensic practitioners who pursue a personal agenda or moral cause when objective opinion is expected and neutrality is required transgress ethical standards and invite malpractice claims. Forensic psychologists should limit their expert witness involvement to cases in which they will have no difficulty suspending motivation to impose their own values, and they should neither seek nor accept professional opportunities to espouse partisan positions in the courtroom.

RESPONSIBILITY

Forensic practitioners at times have to take responsibility for resisting attorney expectations and demands that reflect adversarial aspects of the judicial system and fall within the boundaries of the law but are not consistent with practicing principled forensic psychology. Lawyers who try cases are accustomed to opposing each other, with a winner and a loser in each instance, and they are obliged to do whatever they can, within the rules of procedure and evidence, to win cases on behalf of their clients. This adversarial stance may define how some attorneys view their experts as well as their clients, in which case forensic psychologists must take care to avoid such consequent problems as inappropriate agency, multiple relationships, geographical intrusion, and misused expertise.

Inappropriate agency in forensic practice derives from the fact that lawyers are agents for their clients and are committed to working for their benefit. As one problematic consequence of this adversarial posture, attorneys may regard forensic psychologists they retain as likewise being agents for their clients who will advocate along with them on their behalf and share their commitment to winning their cases. Except when forensic psychologists have been retained as consultants to provide behind-the-scenes advice on case management, rather than as expert witnesses, they must resist being inappropriately cast as agents for the attorney's client. When they are retained to evaluate an attorney's client and provide expert testimony, their client is the attorney who retained them, not the person they are evaluating, and they have no responsibility for serving this person's best interests or advocating on his or her behalf (APA, 2012a, 4.01).

Nor are forensic psychologists agents of the attorney who retains them, and they are no more responsible for serving the attorney's best interests than they are for serving the best interests of the attorney's client. Forensic psychologists have no professional obligations to the parties in a case other than to conduct a competent evaluation and report their findings clearly, accurately, and honestly, whatever these findings may be, and independently of any motivation to win or lose. Fisher (2009) argues in this regard that identifying who should and should not be considered the forensic psychologist's client in a case is less important than having a clear sense of the practitioner's ethical responsibilities to each of the parties in the case.

However it might be defined, avoiding inappropriate agency and thus remaining free of an adversarial posture is an essential element of complying with the Ethics Code. If an examining psychologist's findings appear to strengthen a retaining attorney's case, all well and good; should the findings seem likely to undermine the attorney's chances of winning a case, this eventuality should be construed as the lawyer's problem, not the psychologist's problem, as in the following case.

Case 11. The plaintiff in a personal injury case was seeking damages for allegedly experiencing an incapacitating stress disorder subsequent to being falsely accused of shoplifting in a department store. A psychological examination revealed few of

the kinds of test patterns commonly found in persons with a stress disorder. To the contrary, the obtained data painted a picture of a psychologically resourceful woman with many personality strengths and above-average capacity to deal with stressful situations without becoming unduly upset by them. The psychologist informed the retaining attorney that the results of the examination were likely to work against his client's case. Acting in accord with allowable adversarial procedures, the attorney declined to have a report written and did not identify the psychologist as an expert witness in the case, thereby shielding the psychologist's information from discovery (see Weiner, Chapter 21 this volume). To the psychologist's surprise, however, the attorney thanked him for these seemingly unhelpful results. He had suspected some exaggeration on his client's part and now had some dependable basis for attempting to persuade her to drop or at least reduce her claims, which would make it less difficult for him to represent her.

Concerning the potential for multiple relationships, attorneys may at times ask or expect practitioners to include therapeutic services in their work with a client they have referred for a forensic evaluation. At other times, an attorney may ask a client's therapist if he or she would be willing to write an evaluation of the person for use in a forensic matter. The Ethics Code and the Specialty Guidelines urge practitioners to avoid such dual roles, out of concern that providing forensic and therapeutic services to the same or closely related individuals might impair their objectivity and risk doing harm to the recipient of their services. If asked to undertake concurrent or sequential forensic and therapeutic services, "Forensic practitioners are encouraged to disclose the potential risk and make reasonable efforts to refer the request to another qualified provider" (APA, 2002, 3.05; APA 2013, Standard 4.02). Should practitioners be required by law, institutional policy, or external circumstances to serve in more than one role in a judicial or administrative proceeding, the guidelines recommend explaining the potential drawbacks of such multiple relationships and striving to minimize any negative consequences of the particular circumstance.

As for geographical intrusion, lawyers are ordinarily free to bring both fact and expert witnesses into court from wherever they reside. Psychologists are likewise free to give testimony wherever they are called upon to do so. However, should a forensic consultation involve conducting an evaluation or engaging in other activities that constitute practicing psychology, the provision of such services in a state in which one is not licensed is a geographical intrusion that can evoke ethical complaints or legal action. Attorneys may not know or be concerned about this practice constraint when they retain out-of-state experts, but licensing laws can pose serious obstacles to practicing outside of states in which one is licensed. It falls to forensic psychologists to take responsibility for learning about and complying with the local regulations whenever they practice across state lines, and these regulations vary from state to state. As reported by Goodstein (2012), 11 states allow no practice at all by psychologists who are not licensed in their state; 23 states allow some practice by out-of-state psychologists, but only after prior approval by the state licensing board; and most states that allow interstate practice with or without prior

approval limit the allowable number of days that psychologists can engage in this practice.

Misused expertise is a regrettable turn of events that can occur as a consequence of attorneys being accustomed to an adversarial system in which everyone is entitled to a defense. Psychologists may respect this tenet of the law, but they are not professionally bound by it, nor are they obliged to assist in a defense when it seems inadvisable or improper for them to do so. Whereas lawyers may be assigned by the court or their firm to construct the best possible defense of a client, no matter how weak the case, forensic psychologists are ordinarily free to accept or decline requests to serve as an expert witness, as they see fit. In exercising this freedom, practitioners do well to avoid cases in which their expertise might be misused to shore up a foolish, poorly conceived, or contrived defense. Testifying as an expert when psychological knowledge has little bearing on the issues in a case or, worse yet, when a manufactured defense borders on the psychologically absurd, can be damaging to a forensic psychologist's reputation and reflect poorly on the profession. The practitioner in the following situation showed professional responsibility by declining to become involved in two cases that he anticipated could damage his reputation and the respectability of forensic psychology.

Case 12. A psychologist received separate calls from two attorneys in a firm to which he had been recommended as a consultant. One of the attorneys wanted help with defending a student who had been caught cheating by copying answers from other students' examinations. The attorney had theorized that the cheater could not refrain from scanning his environment in search of answers due to an innate tendency that caused his eyes to wander. He sought to entice the psychologist's participation by offering to work with him in developing this theory and going public with a "wandering eyes syndrome" that could bring both of them some degree of fame and fortune. The other attorney wanted to pursue whether an innate sensitivity to threat might justify his client having beaten his brother-in-law to death, before the brother-in-law could beat him to death, without their being any evidence that the murdered man had acted aggressively or in a threatening manner toward the client. The psychologist felt strongly that participating in either a wandering eyes syndrome defense or an innate threat sensitivity defense would constitute misuse of his professional expertise and expose him to ridicule in his professional community. Preferring to preserve his reputation and reserve his testimony for cases in which his expertise would be used responsibly, he passed on the opportunity to participate in these cases.

RECOMMENDATIONS

A few final words summarize the essentials of practicing principled forensic psychology in ways that minimize practitioners' vulnerability to ethical complaints and legal actions.

- Adhere to the Ethics Code in general, attend in particular to ethical principles and standards having specific implications for providing forensic services, and be familiar with the forensic specialty guidelines.
- Offer services that reflect proficiency achieved through knowledge, skills, education, training, and experience, and do not let this high level of competence erode over time.
- Be conversant with the terminology, concepts, practices, and standards common in the legal community, and be acquainted with the regulations, statutes, and required or customary procedures pertaining to work in states in which the practitioner is licensed or permitted to practice.
- Prepare accurate and sufficiently detailed records, store them securely, and retain them for at least the length of time specified by state requirements.
- Be mindful of the forensic psychologist's appropriate roles in a case, which may differ from what attorneys and attorneys' clients expect them to do, and be sensitive to conflicts of interest and other circumstances that could compromise principled practice.
- When confronting unfamiliar matters or an unexpected dilemma, consult with a respected colleague, and keep in mind that consulting relationships with attorneys can also provide helpful advice on the legal implications of certain forensic issues.

REFERENCES

- American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57*, 1060–1073.
- American Psychological Association. (2007). Record keeping guidelines. *American Psychologist, 62*, 993–1004.
- American Psychological Association. (2010a). 2010 amendments to the 2002 "Ethical Principles of Psychologists and Code of Conduct." *American Psychologist, 65*, 493.
- American Psychological Association. (2010b). Guidelines for child custody evaluations in family law proceedings. *American Psychologist, 65*, 863–867.
- American Psychological Association. (2012a). Guidelines for the practice of parenting coordination. *American Psychologist, 67*, 63–71.
- American Psychological Association. (2012b). Guidelines for psychological evaluation in child protection matters. *American Psychologist, 68*, 20–31.
- American Psychological Association. (2013). Specialty guidelines for forensic psychology. *American Psychologist, 68*, 7–19 doi:10.1037/a0029889.
- Barefoot v. Estelle, 463 U.S. 880 (1983).
- Barnett, J. C., Lazarus, A. A., Vasquez, M. T., Moorhead-Slaughter, O., & Johnson, B. (2007). Boundary issues and multiple relationships. *Professional Psychology, 38*, 401–410.
- Bersoff, D. N. (Ed.). (2008). *Ethical conflicts in psychology* (4th ed.). Washington, DC: American Psychological Association.
- Butcher, J. N. (2013). Computerized personality assessment. In J. R. Graham & J. A. Naglieri (Eds.), *Clinical psychology* (pp. 165–191). Volume 10 in I. B. Weiner (Ed.-in-Chief), *Handbook of psychology* (2nd ed.). Hoboken, NJ: Wiley.

- Fisher, M. A. (2009). Replacing "Who is the client?" with a different ethical question. *Professional Psychology, 40*, 1–7.
- Golding, S. L. (1990). Mental health professionals and the courts: The ethics of expertise. *International Journal of Law and Psychiatry, 13*, 281–307.
- Goodstein, L. D. (2012). The interstate delivery of psychological services: Opportunities and obstacles. *Psychological Services, 9* (3), 231–239. doi: 10.1037/a0027821
- Gorny, I., & Merten, T. (2007). Symptom information—warning—coaching: How do they affect successful feigning in neuropsychological assessment? *Journal of Forensic Neuropsychology, 41*, 71–97.
- Greenberg, S. A., & Shuman, D. W. (2007). When worlds collide: Therapeutic and forensic roles. *Professional Psychology, 38*, 129–132.
- Hess, A. K. (1998). Accepting forensic case referrals: Ethical and professional considerations. *Professional Psychology, 29*, 109–114.
- Jurek v. Texas, 428 U.S. 262 (1976).
- Kitchener, K. S., & Anderson, S. K. (2011). *Foundations of ethical practice, research, and teaching in psychology* (3rd ed.). New York, NY: Routledge.
- Knapp, S. J., & VandeCreek, L. (2012). *Practical ethics for psychologists: A positive approach* (2nd ed.). Washington, DC: American Psychological Association.
- Koocher, G. P., & Keith-Spiegel, P. C. (2008). *Ethics in psychology and the mental health professions: Standards and cases* (3rd ed.). New York, NY: Oxford University Press.
- McMinn, M. R., Bearse, J., Heyne, L. K., Smithberger, A., & Erb, A. L. (2011). Technology and independent practice: Survey findings and implications. *Professional Psychology: Research and Practice, 42*, 176–184.
- Monahan, J. (Ed.). (1980). *Who is the client?* Washington, DC: American Psychological Association.
- Monahan, J., & Steadman, H. J. (1996). Violent storms and violent people: How meteorology can inform risk communication in mental health law. *American Psychologist, 51*, 931–938.
- Shuman, D. W., & Greenberg, S. (1998). The role of ethical norms in the admissibility of expert testimony. *The Judges' Journal: A Quarterly of the Judicial Division, 37*, 4–43.
- Springman, R. E., & Vandenberg, B. R. (2009). The effects of test-strategy coaching on measures of competency to stand trial. *Journal of Forensic Psychology Practice, 9*, 179–196.
- Strasburger, L. H., Gutheil, T. G., & Brodsky, S. L. (1997). On wearing two hats: Role conflict in serving as both psychotherapist and expert witness. *American Journal of Psychiatry, 154*, 448–456.
- Suhr, J. A., & Gunstad, J. (2007). Coaching and malingering: A review. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 287–311). New York, NY: Oxford University Press.
- Victor, T. L., & Abeles, N. (2004). Coaching clients to take psychological and neuropsychological tests: A clash of ethical obligations. *Professional Psychology, 35*, 373–379.
- Weiner, I. B. (1989). On competence and ethicality in psychodiagnostic assessment. *Journal of Personality Assessment, 53*, 827–831.
- Weiner, I. B., & Bornstein, R. F. (2009). *Principles of psychotherapy* (3rd ed.). Hoboken, NJ: Wiley.
- Weiner, I. B., & Greene, R. L. (2008). *Handbook of personality assessment*. Hoboken, NJ: Wiley.

CHAPTER 5

Training in Forensic Psychology

DANIEL A. KRAUSS AND BRUCE D. SALES

INITIALY, the term *forensics* referred to presenting in the public forum (Shah & Sales, 1991). Today it refers to the application of knowledge and services to the law by numerous disciplines (e.g., American Board of Forensic Psychology, 2006; American Chemical Society, n.d.). Although in psychology, forensics was historically associated with the provision of services by clinical psychologists within legal settings (e.g., the courts or correctional facilities) or to law-involved clients (e.g., offenders on parole or probation), some scholars and practitioners use the term synonymously with the field of psychology and law (DeMatteo, Marczyk, Krauss, & Burl, 2009). For example, a forensic psychologist has been defined as “any psychologist, experimental or clinical, who specializes in producing or communicating psychological research or assessment information intended for application to legal issues” (Grisso, 1987, p. 831), whereas forensic psychology has been defined as “both (1) the research endeavor that examines aspects of human behavior directly related to the legal process . . . and (2) the professional practice of psychology within or in consultation with a legal system that embraces both civil and criminal law and the numerous areas where they intersect” (Bartol & Bartol, Chapter 1 this volume) and as “the application of psychological research, theory, practice, and traditional and specialized methodology . . . to provide information relevant to a legal question” (Goldstein, 2007, p. 5). In this chapter, we adopt this broad definitional approach to be consistent with the other contributions in this volume.

Although the origins of forensic psychology can be traced to the early-20th-century works of Hugo Münsterberg and Sigmund Freud (Bartol & Bartol, 2006), it was not until 1973 that the first dedicated training program in psychology and law, which incorporated a specific forensic psychology component, was founded at the University of Nebraska under the direction of Bruce Sales. Currently, there are over 40 programs offering some form of graduate training in forensic psychology (Burl, Shah, Filone, Foster, & DeMatteo, 2012), and the membership of the American Psychology–Law Society (AP-LS; Division 41 of the American Psychological

Association [APA]) has increased dramatically with over 3,000 full and student members.

The forensic programs can be further classified into subgroups based on the particular type of training they offer. Some programs include dual JD (juris doctorate)—PhD (doctor of philosophy) degrees or dual JD–PsyD (doctor of psychology) degrees. Within the PhD component, the programs offer the option of clinical or nonclinical training, or both. Some programs offer the opportunity for students to solely pursue a PhD with a clinical specialization or a PsyD, with a programmatic emphasis in clinical forensic psychology or in clinical psychology with a subspecialty in forensic psychology. Where the PhD component offers nonclinical training, some programs allow students to focus on the application of nonclinical psychology areas (e.g., cognitive, social, developmental) to forensic issues. Finally, forensic training is also offered by some schools at the master's level (for a recent listing of forensic psychology programs, see Aderhold, Boulas, & Huss, n.d.; AP-LS, n.d.-a).

To understand the reasons for the existing differences in training approaches, it is first necessary to understand the differing skills of forensic scientists and practitioners. It is the breadth of professional duties and the scientific questions associated with forensic practice that necessitate varied types of training approaches. For example, a common use of expert testimony consists of psychologists testifying about the reliability of eyewitness testimony against a criminal defendant. The expert witness might testify as to the effect of several factors on perception, including stress and weapon focus. The expert might also offer testimony on the effect of certain factors on accurate identification, such as the forgetting curve and suggestive pretrial identification procedures (see, e.g., *United States v. Norwood*, 1996). Although this type of testimony is well known and may be presented by a forensic clinical psychologist (Contreras, 2001), it is more likely to be the province of someone trained in cognitive or social psychology who studies basic perception, memory, and/or identification accuracy issues.

In contrast, clinical psychologists are more likely to testify about an assessment they performed on a litigant (e.g., Melton, Petrila, Poythress, & Slobogin, 2007) or a treatment they provided to a litigant or offender (e.g., Ashford, Sales, & Reid, 2001). For example, they may testify as to the results of their forensic assessment on a variety of topics, such as: What is in the best interests of the child for postdivorce custodial placement (e.g., Benjamin & Gollan, 2003)? Is termination of parental rights in the best interests of the child (e.g., *In re L.A.M.*, 2001)? Did the defendant suffer from a learning disability in a lawsuit alleging that the disability was caused by chemical exposure (e.g., *Mancuso v. Consolidated Edison, Co.*, 2000)? In other cases, both clinical and nonclinical forensic psychologists educate the trier of fact (i.e., the jury when there is one or the judge in bench trials) about the state of psychological knowledge on some topic (e.g., rape trauma syndrome; causes of eyewitness identification errors) rather than directly focusing on a specific factual question in dispute (e.g., *People v. Wheeler*, 1992).

Although expert testimony is an important part of forensic work, it is by no means exhaustive of forensic practice opportunities. For example, some forensic psychologists work in the administration of forensic correctional facilities (see, e.g., Hafemeister, Hall, & Dvoskin, 2001), provide treatment services in detention facilities to juvenile or adult offenders, or develop policy for facilities and governmental bodies. Finally, while the preceding activities are practice related, many psychologists produce the research and scholarship that provide the foundation for forensic psychological practice.

What type of education and training is best suited for this diverse set of activities? There is no simple answer. The training options should be varied, based on both the trainee's career goals and the educational administrative limits of the existing programs. This chapter considers both the more mundane and the more nuanced aspects of training goals, approaches, and issues.

TRAINING GOALS

We train and pursue educational opportunities to achieve specific career goals. The student interested in working with elderly patients logically seeks training relevant to gerontological issues, and programs interested in attracting students who wish to work with persons with dementia of the Alzheimer's type will create didactic, experiential, and research training opportunities specific to the needs of students interested in this area. It is no different for students interested in working in legal and law-related settings. In this section, we describe the most common training goals in forensic psychology.

CLINICAL SCIENTIST-PRACTITIONER

One of the more common training goals for general clinical psychology programs is the development of scientist-practitioners based on the Boulder model (American Psychological Committee on Training in Clinical Psychology, 1947; Peterson & Park, 2005). Under this paradigm, an individual is trained first and foremost as a scientist versed in the critical thinking skills, hypothesis testing, research methodologies, and techniques that are specific to the science of psychology. In addition, these individuals are instructed, trained, and mentored to apply the existing and most current psychological research findings in their assessment and treatment activities. Successful completion of this training is recognized by attainment of a doctor of philosophy degree (PhD) in psychology.

Clinical psychologists trained as scientist-practitioners in forensic psychology should be competent in four areas:

1. Conducting research on forensic topics
2. Identifying, keeping abreast of, and evaluating the scientific research and professional literatures specific to various areas of forensic practice

3. Implementing the most scientifically appropriate assessment or treatment techniques for a particular case while at the same time being aware of the limitations of the chosen techniques
4. Recognizing when no scientifically valid technique exists for a particular issue or question

CLINICAL PRACTITIONER-SCIENTIST

Some PhD and most PsyD programs deemphasize the scientific component of their clinical training and replace it with more practice-focused didactic and experiential training. Such programs are designed primarily to train students to become clinical practitioners and consultants to the courts, and they produce fewer graduates who are likely to seek employment in research and/or academic settings. As such, trainees from these programs often (a) graduate with more hands-on clinical training experiences; (b) have had more classes devoted specifically to clinical practice issues; and (c) have had more clinical opportunities to practice their clinical skills on real-world populations. These students, however, spend less time conducting research and have fewer courses directly addressing the design, methodology, analysis, and evaluation of forensic psychology research.

Arguably, individuals receiving this type of training are better prepared to provide clinical services than students who have received less education and training focused on practice issues. Yet the argument is refutable if the discipline of psychology is to be built on knowledge of the science of human behavior. To respond to this concern, faculty who focus on training practitioners endeavor to teach their students how to be effective consumers of scientific research. To date, little research exists examining how the more practice-oriented training inherent in these programs affects actual practice competence, and no such research exists in the forensic psychological area. As a result, controversy still exists regarding the benefits and limits of professionally focused training, with advocates on both sides suggesting the superiority of their model for training clinical practitioners (Peterson, 2003).

NONCLINICAL AND CLINICAL SCIENTIST-SCHOLAR

Not all psychologists are clinically trained and licensure-eligible for clinical practice. Training programs in cognitive, developmental, and social psychology, to name just a few of the subfields of psychology, are also widely available. Through a series of didactic and intensive research experiences, these programs produce PhD psychologists whose focus is research and/or teaching and typically work in academic settings, research laboratories, and industry settings.

It is no different in forensic psychology. Although most forensic psychologists are interested in the assessment and treatment of forensic-clinical populations, many nonclinical forensic psychologists have been trained to pursue careers as scientist-scholars constantly trying to expand the boundaries of our forensic knowledge.

Given the importance of the scientific foundation for understanding human behavior in legal settings (e.g., jury decision making) and for improving the structure and administration of law (Sales, 1983), this type of forensic training is prominent at many universities.

Similarly, there are clinically trained forensic psychologists who, while able to offer assessment and treatment services, do not focus their careers on these practices. These individuals critically analyze and synthesize existing research and carry out new research on clinical forensic topics, so that their colleagues are aware of the state of the science in clinical forensic practice. For example, the most recent advances in risk assessment central to legal decisions on civil commitment, execution (in states that use dangerousness as a criteria), and postconviction confinement of sexual predators have been completed by such scholars (e.g., Monahan et al., 2002).

The training goals within this category are not uniform or unitary. Students need to be trained in two very different kinds of academic skills to carry out this type of work: scientist and scholar.

Scientist. Training for a career as an empirical researcher requires rigorous, didactic, and experiential training in theory, hypothesis generation, methodology, data analytic techniques, and interpretation of findings. These educational experiences are common to all branches of scientific psychology. Although the goal is to produce graduates who design and carry out their own research in the future, the specific types of research that this training prepares one for differ across programs. For example, not all scientists design or carry out their research for forensic purposes. In *United States v. Virginia* (1996), the U.S. Supreme Court had to decide on the constitutionality of the State of Virginia maintaining an all-male military college—the Virginia Military Institute (VMI). As part of its decision, the Court considered research on the psychological effects of females receiving education in a single-gender versus coeducational environment. Although this research was conducted because of the researchers' interest in social development and gender studies, its application to the legal question at issue in the VMI case makes this research forensically relevant work (English & Sales, 2005).

Not surprisingly, nonforensic psychological researchers can find their work being used in litigation or policy decision making across numerous issues. For example, research on child development can be used in the drafting of educational legislation, while social psychological studies of gender stereotyping can be used in sex harassment lawsuits (e.g., *Harris v. Forklift*, 1993). Training psychologists for a nonforensic scientific career, even though the products of the subsequent research may have unanticipated forensic uses, is not an uncommon aspiration of traditional PhD programs.

But an increasing number of PhD programs are offering students the opportunity to specialize in forensically relevant research. This training can include basic, policy-driven, legally driven, or litigation-driven research. Basic forensic studies seek to

expand our knowledge of forensic psychological phenomena (e.g., cognitive biases in judicial decision making; see, e.g., Krauss, 2004) or techniques (e.g., research on assessment of dangerousness; see, e.g., Quinsey, Harris, Rice, & Cormier, 2006). Policy-, law-, or litigation-driven forensic research is designed and carried out to answer specific questions posed by a specific policy or legal question. For example, in *Lockhart v. McCree* (1986), researchers attempted to demonstrate that the process of death-qualifying juries (jurors who were morally opposed and could not sentence a defendant to the death penalty were dismissed from both phases of the trial) in a capital hearing case created an unconstitutionally biased jury in favor of the prosecution for the guilt/innocence phase of the trial. While the majority of the U.S. Supreme Court was unpersuaded by this research for both methodological and legal reasons, the research presented to the Court was undertaken specifically to address these questions that were left open by the Court's earlier decision in *Witherspoon v. Illinois* (1968).

Training for forensic psychological research competency requires more than training in science, however. In order for forensic psychological research to be useful to the law (e.g., courts) or policy makers (e.g., legislators), it must both be relevant to the legal questions being adjudicated or considered and be sufficiently tied to issues under dispute or consideration in a particular case, statute, or policy. The importance of these considerations has been made explicitly clear with regard to the introduction of scientific evidence in the federal courts and in the majority of state courts (e.g., Sales & Shuman, 2005). The U.S. Supreme Court, in *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993), adopted relevancy and fit requirements for the admissibility of scientific expert testimony at trial. In describing the importance of the fit between the science being offered and the legal question being asked, the Court noted that "scientific validity for one purpose is not necessarily scientific validity for other, unrelated purposes" (p. 2796). *Daubert* and two subsequent cases (*General Electric Co. v. Joiner*, 1997, and *Kumho Tire Company v. Carmichael*, 1999) have held that federal trial court judges must evaluate the reliability, relevancy, and fit of proffered expert testimony and research and reject evidence that does not meet these criteria (Krauss, Cassar, & Strother, 2009).

Because *Daubert* controls only decision making in the federal courts, it technically is not required to be used in considering the admissibility of expert evidence in state courts, before state and federal legislatures, or by administrative agencies. The *Daubert* logic should be applicable, however, to studying any issue or question in the law (e.g., Schopp, 2001). Indeed, the majority of states have adopted the federal rule concerning the admissibility of expert evidence into their state rules of evidence (although a significant minority of states have not adopted *Daubert's* holding, and still others use only parts of *Daubert* and its progeny's [i.e., *Joiner* and *Kumho*] admissibility standards) (Sales & Shuman, 2005). Forensic psychological researchers need to understand the law that provides the issues and questions they wish to study. Not to answer the questions posed by the court will likely result in the findings of forensic psychological scientists having little legal relevance and value

to the court (e.g., for a review of the problems inherent in psychological research and expert testimony on child custody determinations, see Krauss & Sales, 2000).

Scholar. Designing and executing research is only part of the skill set that is required of individuals who are attempting to influence a field. It is also important that scholars be able to critically evaluate existing theory and research, identify where the field is weak, and point where theory and research needs to be directed. For example, divorce mediation, which is prevalent throughout the United States, is touted as an alternative dispute resolution system that offers substantial benefits for avoiding common disputes that arise during marital dissolution through the legal system. After critically reviewing the extensive empirical literature on divorce mediation, however, Beck and Sales (2001) concluded that most of the benefits associated with divorce mediation are not yet supported due to a host of legal, theoretical, methodological, and statistical weaknesses in the research literature. Research evidence also suggests that some of mediation's touted benefits are false. Training in scholarly skills is particularly important in order to produce individuals who can generate a critical understanding of the state of the science on any forensic topic and redirect the field when that science is inadequate to significantly advance its knowledge or practice (e.g., Findley & Sales, 2012).

APPROACHES TO ACHIEVING TRAINING GOALS

Given that the training and career goals of forensic psychologists can differ markedly within and across categories, training programs have developed divergent programmatic approaches to educating their students so that they can attain these disparate goals. These programmatic approaches are not based solely on achieving specific training goals, however. They are also subject to administrative factors that affect the type of training a program can offer (e.g., number of faculty lines in forensics). As a result, the existing subtypes of training programs all possess both benefits and limitations that affect their ability to effectively train their students to attain their career goals. In this section, we describe the most common types of training programs and highlight the areas in which they are likely to benefit and limit their students for different types of forensic careers.

FORENSIC CLINICAL PRACTITIONERS

Although we have no data, our impression is that the vast majority of practitioners who describe themselves as forensic psychologists were not trained in graduate programs specializing in forensic psychology, because few such specialty programs existed when they were in training. As a result, most forensic practitioners received general clinical training in graduate school and later undertook more specialized training through postdoctoral work, continuing education courses, on-the-job training, or some combination of these possibilities.

Individuals trained under such a paradigm are likely to benefit in their forensic work from the generalist knowledge gained during clinical training. In addition, these individuals often have accrued a wide range of clinical experience across a significant range of treatment settings, patient characteristics, and disorders before attempting more forensic-based clinical practice. As these practitioners refine their skills in the forensic arena, this general clinical training can serve as part of the foundation for forensic practice.

Yet in a variety of ways, such general clinical training is also likely to serve as a constraint on the forensic psychological skills of these professionals. General clinical training does not often prepare individuals to understand the clinical and research literature most pertinent to forensic practice. For example, forensic psychological assessment often is predicated on the evaluator responding to specific legal questions (e.g., is the person incompetent to stand trial? Which custodial placement would be in the best interests of the child?). Not understanding the governing law can lead to inappropriate assessments. Because general clinical training is unlikely to offer trainees such specific legal training, even if trainees want to access the forensic literature, they would be unlikely to know under what circumstances they would need to access it and where to find it. Moreover, even if practitioners uncovered the appropriate literature for the legal question at issue, they would be unlikely to understand the legal nuances involved.

Consider the case of a practitioner who wishes to perform an insanity evaluation for the first time. Without explicit knowledge of the legal standard governing such an assessment in the jurisdiction prosecuting the defendant, it would be impossible for a practitioner to do an appropriate job. Insanity standards vary markedly across jurisdictions. For example, the federal system defines an insane individual as a defendant who “at the time of the commission of acts constituting the offense . . . as a result of severe mental disease or defect was unable to appreciate the nature and quality or the wrongfulness of his [or her] acts” (United States Code, 18 USC § 17). Other jurisdictions, however, also include a volitional component in their definition that allows for the acquittal of individuals whose mental illness affects the ability to conform their behaviors to the requirements of the law (Wisconsin Code, § 971.15, 2012). Within these two large subtypes of insanity definitions are several additional minor jurisdictional variations. Further, some states do not even allow for an insanity defense and allow the introduction of psychological evidence only for the much more limited purpose of determining whether the defendant had the mental state required for the crime (i.e., a *mens rea* defense). As a consequence, the assessment techniques utilized by the practitioner must be based on the idiosyncrasies of the controlling legal definition, or the practitioner will end up answering a question that the legal system is not interested in. Without specialized forensic psychological training, the clinical practitioner is unlikely to know the controlling law or realize that readings relevant to the insanity defense in one jurisdiction may not be informative about the insanity defense in the jurisdiction in which the forensic practitioner currently practices.

In regard to forensic treatment, the practitioner trained as a generalist is also likely to experience problems in forensic psychological practice. Clinical training typically encompasses courses on various major therapeutic modalities, such as cognitive-behavioral therapy. But these same programs are unlikely to assign the literature that addresses which type of treatment modality will work best with various types of adult and juvenile offenders (e.g., Andrews & Bonta, 2010; Ashford et al., 2001). Indeed, the research literature on treatment for offenders, or persons otherwise involved with the law, typically is not covered in general courses on clinical treatments and interventions.

A related problem is that forensic psychological services often have different goals from those set for therapy with private clients. Whereas in the latter, the client is seeking to “feel better” mentally and emotionally, the goals for the treatment of forensic patients are often set by the law. For example, the most appropriate treatment or intervention for persons found incompetent to stand trial involves making the person competent to return to court, not necessarily making him or her mentally or emotionally healthy. Similarly, correctional administrators are more concerned about clinical services that reduce inmates’ dangerousness and suicidality and are less concerned with programs designed to produce mentally healthy inmates. In general, clinical program trainees will not receive important information specifically relevant to the needs of and the requirements imposed by various laws and legal systems (e.g., state and federal courts; state or federal departments of corrections) that set the standards for the clinicians hired by the government (see also Gendreau, Goggin, & Smith, Chapter 23; Morgan, Kroner, Mills, & Batastini, Chapter 24; and Marshall, Boer, & Marshall, Chapter 25, all this volume).

The lack of specialized forensic training in general clinical training programs is not limited to forensic treatment outcome research or to legally relevant standards and criteria. Generalized clinical training also suffers because it typically does not include didactic training on the unique ethical problems that forensic practitioners face. Shuman and Greenberg (2003), for example, have written on the unique ethical problems that treating therapists confront when they are retained to evaluate and testify about their clients. These kinds of unique concerns have led to the publication of the Specialty Guidelines for Forensic Psychology (APA, 2013; the Specialty Guidelines are reprinted as the appendix to this volume with permission of the APA). These specific ethical issues are unlikely to be covered adequately in general clinical training.

Graduating students from general clinical training programs have relied on several routes to address their lack of appropriate forensic psychological training. These include attending an internship program that focuses on forensic psychology, receiving postdoctoral supervision from a forensic specialist, attending continuing education programs, and engaging in self-directed readings, all of which are likely to improve a practitioner’s forensic abilities. Unfortunately, anecdotal evidence suggests that some graduates of general clinical training programs do none of these

things, assuming instead that what is good clinical practice in other settings will be sufficient in the legal arena.

As a direct solution to this problem, some general clinical training programs offer an emphasis in clinical forensic practice. Such training can compensate for the limits in general clinical training, with the caveat that how well a program compensates depends on the comprehensiveness of its forensic emphasis and the training and education opportunities the program provides. Students would be well advised to check the specialty courses and practica that programs of interest offers to them. For example, many existing programs still lack adequate legal training for forensic practice in most areas. Miller, Sales, and Delgado (2003), for example, identified more than 75 areas of law that substantially affect the provision of forensic services. A recent survey of forensic psychology graduate programs confirms this omission (Burl et al., 2012), with only one-third of clinical programs with a forensic emphasis offering courses in mental health law.

Logically, the most effective form of training for providing forensic clinical services should be provided by forensic psychology specialty training programs. We say logically because there are no empirical studies of training outcomes in this area. These programs typically offer comprehensive forensic coursework and externship placements to ensure that the graduates are well prepared for forensic practice after licensure.

FORENSIC PSYCHOLOGY NONCLINICAL PRACTITIONERS

Not all forensic practice is related to clinical psychology. For example, training to be a government policy analyst may be best accomplished through focusing on evaluation research and methodology. In contrast, training to provide consultation to child protective service agencies may be better accomplished through applied developmental training than clinical training, while individuals interested in providing trial consultation to lawyers are typically best prepared for their occupation through forensic social psychological training.

These programs, because they are organized in similar ways to forensic clinical programs, suffer from the same benefits and limitations. Some are general programs, others offer opportunities for the acquisition of some forensic skills (e.g., faculty offering training to students in eyewitness identification and false confessions), while others offer forensic nonclinical specialty training. These programs, however, often neglect important areas of forensic psychological knowledge. For example, results of a recent survey indicate that less than one-third of these programs offer classes on juvenile offending, psychology of criminal behavior, mental health law, ethics, victimology, and sociocultural issues in forensic psychology (Burl et al., 2012).

FORENSIC SCIENTISTS

Not all forensic psychology trainees aspire to a practice career. As already noted, some of these students will look to academic careers to pursue their research or

to other venues that will allow them to work as researchers (e.g., research think tanks). Training for these positions is in many ways similar to training scientists in any subfield of psychology, with only the content of the research examined changing. Thus, individuals interested in studying eyewitness identifications often study in a cognitive psychology program, while individuals interested in pursuing child suggestibility in interview situations often enroll in a developmental psychology program. Finally, individuals interested in researching forensic psychological assessment could pursue their interests through one of the scientifically driven clinical programs.

Such training has its benefits. Trainees graduate from respected traditional psychology programs, which often open the door to faculty positions in other respected psychology departments. But there are also costs to attending such programs. Often these programs do not have faculty members who are expert on forensic issues beyond their own research interests. For the individual interested in broader training in forensic psychological science, the solution is to attend a program that focuses more generally on forensic science and offers the necessary concomitant didactic and experiential experiences for more expansive forensic scientific training.

DEGREE AND NONDEGREE TRAINING OPPORTUNITIES

Given the different approaches to achieving training goals, what are the specific degree and nondegree training opportunities in forensics that are available to trainees? As it turns out, there are quite a few.

PHD/PSYD PROGRAMS

As noted in the prior section, few schools that offer forensic training do so through a program devoted to forensic psychology. Typically, a student enters a clinical psychology program that offers a forensic emphasis, but there are schools where training in forensics is accomplished through a specialized forensic program. As noted earlier, the benefit of such focused training is that it offers an intensive program of study in forensics.

INTERNSHIP TRAINING

Students in clinical or other professional training programs (e.g., counseling psychology, school psychology), whether they have a forensic emphasis or not, are required to take a year of internship experience, which can be at sites that offer a forensic focus (AP-LS, n.d.-b). Because these experiences are focused on patient assessment and/or treatment, the opportunity for experiential learning is substantial. When combined with a forensic specialty predoctoral training program, these internships can substantially broaden a trainee's skills. Forensic experiences in

predoctoral internship programs have become more common over time, with over 50% of APA-accredited internships offering clinical training in a forensic rotation (DeMatteo et al., 2009), although they vary considerably in the depth of training and supervision. While forensic clinical experiences in predoctoral programs that focus on training generalists are valuable, these clinical experiences will not necessarily provide trainees with knowledge about pertinent forensic literatures.

JOINT MAJORS IN PSYCHOLOGY AND JOINT DEGREE PROGRAMS

Some schools allow students to pursue concurrent training in two specialties within one PhD program. The benefit of this approach typically is greatest for those seeking academic careers. Training in two majors typically forces students to master a greater array of the psychological literature and allows them to seek employment in departments that have jobs available in either specialty. Clearly, for joint training to benefit trainees substantively, the acquired knowledge should be integrated to enhance forensic research and scholarship. For example, studying social psychology can enhance forensic research on jury decision making.

When joint training is mentioned, however, it is typically across colleges rather than within a department, with a number of schools offering the opportunity for students to pursue the PhD or PsyD in combination with the JD degree, or the PhD and the MLS (Masters of Legal Studies) degree. The MLS degree typically is a 1-year intensive training program for individuals not seeking to be licensed as lawyers, but who still want to learn a sufficient amount of law to be able to enhance their forensic scholarship or practice.

There are four benefits of joint degree training programs:

1. Increased proficiency in psychological science and legal research, writing, and thinking
2. Integration of the two fields of study through the course of graduate education, so that the individual can both think like a lawyer and perform research and practice like a psychological scientist
3. Greater understanding of legal norms, rules, and standards so, at the least, forensic practitioners emerging from such programs know the laws that affect their practice and the specifics of the legal questions that the court might ask them to address
4. Greater understanding of legal norms, rules, and standards so that practitioners' forensic research meaningfully addresses relevant legal topics, thereby increasing the likelihood that the research will be accepted by the legal system and policy makers

Finally, some students who pursue the PhD recognize only after graduation that they have an interest in pursuing legal training. This typically takes one of two forms: entering law school to pursue the JD or entering an MLS program after the

doctoral degree is completed. The obvious benefit of taking this route is that the forensic psychologist will obtain valuable legal training. The less obvious deficiency of such an approach is that taking training sequentially increases the chances that the student will not learn how to integrate the two fields. Learning biology and learning chemistry, for example, does not ensure that one will learn the theory, findings, and methods of biochemistry. It is no different in combining psychology with the law.

RESPECIALIZATION AND POSTDOCTORAL TRAINING

Two other postdoctoral pathways exist for increasing forensic knowledge. First, students trained in nonclinical programs can reapply after graduation to achieve respecialization in clinical psychology with a forensic emphasis. This type of training may take 2 or 3 years and is a substantial investment of time, and where one does respecialization will affect how competent the trainee becomes in forensic clinical psychology.

The other alternative is for doctoral graduates to seek additional postdoctoral research or practice experience for 1 or 2 years. Postdoctoral training allows intensive work in a given area under direct supervision, and thereby provides an excellent opportunity for doctoral graduates to acquire or increase their forensic knowledge (e.g., a forensic psychology fellowship). A recent review reports that there are 16 formal postdoctoral clinical fellowships in forensic psychology providing approximately 30 positions (Malesky & Proctor, 2012; see also AP-LS, n.d.-c, n.d.-d).

CONTINUING EDUCATION PROGRAMS

Perhaps the most common means for many practitioners to increase their forensic knowledge in specified areas is to attend continuing education programs focused on forensic practice issues. For example, the American Academy of Forensic Psychology offers continuing education programs on a wide array of forensic topics that are aimed at practitioners possessing different levels of forensic skills and experience (www.aafp.ws). It is not clear, however, if a 1-, 2-, or even 5-day workshop can adequately prepare practitioners to fully understand the complexities of different forensic practice areas without additional supervision by a qualified practitioner, additional self-directed readings, prior training, or attendance at a substantial number of continuing education programs.

MASTER'S-LEVEL FORENSIC PSYCHOLOGICAL TRAINING PROGRAMS

Although doctoral training is the norm for independent professional practice, some students (e.g., lawyers) who are interested in acquiring entry-level knowledge in forensic psychology or in providing nonclinical forensic services can seek a master's

degree in forensic psychology. Although the AP-LS (n.d.-e) Web site lists over 20 of these programs, such training rarely makes a graduate competent to critically analyze the extant literature or arguably to provide the same quality of service provision that the doctoral-level forensic psychologist would offer.

DISCUSSION

Understanding the training goals, approaches to achieving these goals, and alternate structures used for delivering training—although a good first step for understanding forensic training opportunities—is not sufficient. Myriad challenges that these programs face and that training raises deserve extended discussion. Some of these, like law and ethics, were considered earlier but are worth reconsidering, given their importance. The other topics logically extend our earlier discussion.

FACULTY EXPERTISE AND STUDENT GOALS

In assessing a forensic program, it is reasonable to ask about faculty and adjunct supervisor qualifications and expertise. It is important that programs have faculty members who are well qualified to teach and supervise in the areas to which they are assigned (APA, 2002, Standard 2.01). Although this admonition sounds obvious, as the pressure for forensic psychological training increases and new programs are created, it is important that programs and potential students critically evaluate the expertise of faculty for providing forensic training. We should never confuse competence in providing clinical services with competence in providing forensic clinical services or nonclinical forensic services or research.

The type and quality of training also will be affected by the administrative structure and financing of the training programs. For example, generalist programs are less likely than forensic specialty programs to have faculty members who are broadly trained in forensic psychology. The availability and size of the faculty members who are expert in forensic psychology will also impact the type of training that is available. This is not to argue that having one faculty member is necessarily worse than having five faculty members in a program. The issue is one of fit between the particular career aspirations of the trainee and the skills, abilities, and availability of the faculty at a particular program.

Perhaps the most basic and often overlooked issue in forensic training is that forensic services are broader than simply providing assessment or treatment in legal settings. For example, psychologists serving as policy analysts in a state legislature, serving in an administrative capacity within a governmental agency, or providing expert testimony on a wide variety of psychological topics are all providing forensic psychological services. Trainees need to look critically at what specific training programs offer and how well those offerings match their career goals. In addition, programs that focus on clinical assessment or treatment typically

limit their focus to particular areas of forensic services in these two domains. Once again, trainees need to understand how these limitations will affect their competence when they graduate. Forensic research programs face the same issue. Not all programs provide training in all areas of forensic research.

Moreover, it is unreasonable to expect a limited number of faculty members to competently train students in all areas of forensic psychology. What is reasonable, however, is to expect that training programs will accurately represent what areas of forensic psychological research and services training will be offered to trainees if they choose to attend that program.

FACULTY AS ADVISORS AND MENTORS

No matter what the expertise of the faculty, it is important to consider whether faculty members view themselves as advisors or mentors. The former, which is more typical, occurs when faculty members perceive their role only as imparting information to students and being available to answer student questions. The latter role involves faculty members taking personal interest in the growth of students and in their success after graduation. We have no data to know how individual students fare under each type of training role, but students ought to be aware of the existing differences in supervision styles and seek information about the faculty approaches to training before making a decision about which program to attend.

NUMBER OF TRAINEES

Much like the size and expertise of the faculty, the number of students currently enrolled in a program is likely to affect their training experiences and education. Some forensic training programs admit few students, while others admit many more. Being in a program with a small cohort of students is likely to lead to more individualized attention for the trainee, but it will also limit a student's opportunities to interact with, share experiences with, and learn from other graduate students in the field. Programs with a smaller cohort of students may also (but not necessarily) experience a smaller number of experiential and didactic training options. In contrast, larger forensic training programs, while offering less individualized attention, are more likely to offer a wider array of both training opportunities and coursework across a range of forensic psychology topics. This typically occurs because the larger the student body, the more faculty members are likely to be associated with it. Prospective trainees should obtain a realistic picture of the programs in which they are interested.

DIDACTIC AND EXPERIENTIAL TRAINING OPPORTUNITIES

Forensic training requires both didactic and experiential training components. Didactic courses are necessary because they provide the intensive opportunity to

acquire the scientific and practice knowledge base underlying forensic psychology. Simply apprenticing under a practitioner might lead to a narrow perception of what forensics entails and a skill set that is limited to what that practitioner knows and does in practice. Experiential training through externships, internships, and other supervised practica can augment the learning that has occurred in the classroom and in directed readings courses.

The important point for prospective students is to evaluate how well the available courses and experiential learning opportunities match their training needs. This is not an easy task. Programs are unlikely to provide training in all areas of forensic psychology. Thus, all trainees by definition will be deficient in some areas of forensic psychological knowledge (e.g., learning about guardianship law and forensic practice). Although not being expert in all forensic areas is a foregone conclusion given the existing training programs, it is not necessarily a problem. Graduate training is the beginning, hopefully, of a lifetime of learning and not the end of one's learning about forensic psychological knowledge and skills. To the extent that trainees are taught how to identify specific legal issues, the scientific knowledge base available to address the legal question, and the forensic skills related to those issues, trainees can acquire the substantive knowledge after graduation.

Of particular concern in didactic and experiential training is the quality of that training. It is important for students to be educated regarding the current approaches to forensic psychological practice, but the hallmark of the scientist-practitioner model (i.e., the Boulder model) is that trainees learn how to think scientifically and ask whether what is, is what ought to be. The importance of this level of critical analysis in teaching in the classroom, in research settings, and in experiential settings cannot be overemphasized because it fosters critical questioning of forensic psychological skills and services. Externships and internships are particularly vulnerable to not allowing for this analytic process to be learned. Trainees often begin their practice experiences in work settings with professionals who are overworked and report having limited time to stay abreast of the current scholarly literature. What they share with trainees is what they do rather than presenting the full panoply of approaches that are available for use with different problems and clients. For example, a psychodynamically oriented practitioner may provide excellent training to externs concerning his or her approach to therapy with particular types of clients but provide no information about alternative approaches or how one chooses between them, given a particular problem, client, and setting. The result is that the trainees can confuse information acquired during apprenticeship with the best available scholarly information in the field or even the best current clinical practices.

LEARNING THE RELEVANT LAW AND HOW TO FIND IT

Training programs need to be especially cognizant of the legal research and training skills that are a necessary component of effective training. For example, it is one thing for forensic trainees to know the legal standard that governs the specific forensic

evaluation they are being asked to perform in a jurisdiction. It is quite another thing for students to be trained in the skills that would allow them to identify the relevant law, fully understand it and its implications for forensic psychological practice, and keep abreast of changes in that law. At a minimum, expertise to perform a particular type of legal assessment entails the ability to do five things:

1. Identify and read the central case law in that jurisdiction on a particular evaluation question.
2. Use this case law in a meaningful way so that the evaluation best addresses the legal standard controlling the forensic issue in question.
3. Keep current on changes in that law and related evaluation practices.
4. Be aware of legal standards in other jurisdictions that may have implications for the forensic work on the identified case.
5. Be aware of legal changes not central to the evaluation standard itself but critical to the success of the forensic work (e.g., changes in admissibility standards) because they will affect both the evaluative procedures utilized and whether the eventual psychological conclusions generated will be accepted by the court as expert testimony.

Training programs that do not have extensive legal components risk handicapping their trainees by compromising the quality of their future work (see DeMatteo, Kessler, & Stohmaier, Chapter 3 this volume, for further discussion).

FORENSIC ETHICS TRAINING

It is also important that forensic training offer competent training in forensic psychological ethics. However, even forensic psychology specialty programs largely fail to offer coursework relevant to ethics in the forensic context, with fewer than 15% of doctoral programs offering a specific course. Forensic psychological ethics receives more attention in master's programs, with approximately half of these programs providing a specific ethics class (Burl et al., 2012). Lack of specialty ethics training is particularly problematic because ethical challenges in general practice are not always identical to the challenges faced by forensic practitioners, with forensic practice generally raising more and different ethical issues. For example, who is the client when a psychologist performs services with a prisoner—is it the prisoner, the prison, or the Department of Corrections? As noted earlier, to help answer questions such as these, there are specialty ethical guidelines designed specifically for forensic psychological practice (APA, 2013). Although these guidelines are aspirational, all forensic practitioners should be aware of, understand, and strive to follow them as well as keep abreast of more current forensic ethics scholarship. Due to these unique ethical concerns, not having an expert in ethics teach the ethics sequence and not having that teaching supplemented by an expert in forensic ethics is questionable training practice. In addition, given that in many graduate psychology programs,

general ethics training is provided by practitioners or faculty members who are not expert in ethics, forensic programs need to consider if their approach is also acceptable for creating competent forensic specialists. We think not.

FORENSIC PSYCHOLOGICAL COMPETENCE

Defining forensic psychological competence in one area of practice or research or across several areas is not an easy task. For example, the acceptance of certain types of expert testimony, research, or practice by the legal system is often a poor indicator of the competence of individuals who proffer, engage in, or use such techniques. Consider a jurisdiction that has accepted expert testimony based on pure clinical hunches about a defendant's future dangerousness in death penalty sentencing (*Barefoot v. Estelle*, 1983). A practitioner offering such testimony is not forensically competent. A forensically competent practitioner would be aware of the superiority of actuarially based dangerousness predictions over unstructured clinical judgments (Grove & Meehl, 1996; Krauss & Sales, 2001) and recognize the limitations in using an existing actuarial instrument as the basis for expert testimony on dangerousness (Krauss, McCabe, & McFadden, 2009; Monahan et al., 2002).

Forensic psychological competence is also not necessarily performing a forensic service in the exact same manner as supervisors did during training. It entails recognizing that advances in the field, differences in the applicable law based on the jurisdiction where the evaluation takes place, and changes in the law over time will influence how each forensic service should be performed in the future. The key is for forensic training programs to instill in their graduates the necessary intellectual rigor that they will regularly ask themselves this question: Is what I am about to do what ought to be done or am I simply repeating past practices?

CREDENTIALING IN FORENSIC PSYCHOLOGY

The timing of forensic training raises a fundamental training issue. As previously noted, forensic training can occur predoctorally, during internship, postdoctorally, through continuing education programs, through on-the-job training, or through self-directed reading. All may be perfectly appropriate for providing forensic psychological expertise, but we have no data to know what kind of impact these training methods have on the acquisition of forensic psychological knowledge and skills.

Because the discipline and specialty is uncertain about the merits of the different training approaches, licensure for practice still depends on doctoral training that includes an internship component. But after licensure is attained, decisions about whether to engage in a specialty forensic practice are left to the ethics of individual practitioners. Once again, we have no data to know how well practitioners self-monitor and self-evaluate the competency of their forensic skills. Not surprisingly, some jurisdictions, in response to perceived weaknesses in forensic practitioner competency, now require specialty training for certain types of evaluations (e.g., see

2012 California Court Rules 5.225(d)–(i), which specifies specific training, education, and continuing education requirements for individuals performing child custody evaluations for the California courts).

The timing, extent, and type of training individuals receive will also affect their ability to obtain board certification in forensic psychology. Board certification as endorsed by the American Board of Professional Psychology (n.d.) allows certification in 14 distinct psychological areas, including forensic psychology, and represents one of the two board certifications listed in the APA directory. To become a board-certified specialist in forensic psychology, individuals must do six things:

1. Complete an approved internship.
2. Accrue a significant number of general clinical practice experience.
3. Obtain at least a minimum number of hours of specialized training in forensic psychology.
4. Work within the field of forensic psychology for a prescribed number of years.
5. Submit two work samples in two different areas of forensic psychological practice.
6. Pass oral and written examinations.

(For the specific requirements for board certification in forensic psychology, see American Board of Forensic Psychology, n.d.) If forensic practitioners or students are interested in becoming board certified, they must choose training experiences that will fulfill these requirements. Thus, prospective trainees and practitioners interested in specialty certification must address carefully the timing and adequacy of training needs.

MAINTAINING AND INCREASING FORENSIC PSYCHOLOGICAL COMPETENCE

One measure of the success of forensic training programs is whether graduates are motivated to seek continuing education programs that will maintain and increase their forensic psychological competence. Unfortunately, there are no data available to address this issue. In addition, although continuing education programs exist in the forensic arena, there has been limited analysis of the variety of topics covered by existing programs, the comprehensiveness of the coverage of the presented topics, the availability of programs in various parts of the country, and whether the covered topics match the wide diversity of forensic psychological training needs. Hopefully, research will address these important issues in the future.

ACCREDITATION OF FORENSIC PSYCHOLOGICAL TRAINING

Forensic doctoral training programs are not accredited by the APA. When they advertise that they are accredited, they are referring to the fact that their generalist clinical, counseling, or school psychology program is APA accredited. Whether accreditation is something that will increase the quality of forensic psychological

practice training is an issue that the field needs to consider. If the consensus is yes, representatives of the field will need to petition and work with the APA Accreditation Office to ensure that forensic training programs come under the APA's accreditation umbrella. Similar concerns can be raised about the accreditation of forensic internship and forensic postdoctoral training programs.

CHALLENGES IN TRAINING FORENSIC PSYCHOLOGY SCIENTISTS

Programs that specialize in the training of forensic psychology scientists face specific training challenges. In order to perform useful forensic psychological research on a topic, a forensic training program must first teach its students how to identify the questions that the law needs answered by psychological science. For example, in the case of *McCleskey v. Kemp* (1987), Kemp, an African American, was charged with murder of a White man. The defense introduced the testimony of an expert to support its claim that the Georgia death penalty statute and process discriminated against the defendant. Although the expert had analyzed 2,000 death penalty decisions from Georgia courts, the U.S. Supreme Court was not persuaded, noting, among other things, that the case was brought on a challenge to the Georgia death penalty statute under the U.S. Constitution's 14th Amendment Equal Protection Clause. This clause requires not only a showing of discrimination but also a showing that the discrimination was intentional. Although the social science data demonstrated a relationship between death penalty decisions and defendant and victim race, the data did not address whether the discrimination was intentional. In the end, the Court was unpersuaded by the data presented because they did not specifically address this latter issue.

Training programs must teach their students how to identify issues that are testable through psychological science techniques and are also important to legal proceedings. In order to accomplish this goal, students must be taught to operationalize difficult legal constructs in such a way that meaningful empirical examination can occur. Without the ability to operationalize these concepts effectively, it is impossible to measure and study appropriate legal concepts using available psychological techniques and methodology (Krauss & Sales, 2003).

Some scholars argue that it is not always possible to operationalize legal concepts using psychological science because legal standards often include moral and normative judgments that are impossible to evaluate effectively with psychological methodology (Grisso, 1986, 2003). This does not suggest, however, that operationalization and empirical investigation of legal questions are inappropriate or unimportant. Rather, it suggests that empirical research must be seen as a means to enhance and inform judicial and legal decision making rather than as a substitute for it (Krauss & Sales, 2003). Student awareness and understanding of these issues then is a necessary precondition to useful forensic psychological research being performed; students must not only understand how to test important legal issues effectively but also recognize the limitations of empirical research to answer all legal questions.

For example, the law does not always allow for experimentation in legal settings. As a result, researchers often are left with no choice but to use simulated research in artificial settings. This often is the case, for instance, in jury research. It typically is conducted using college students who are exposed to a brief transcript in a laboratory setting. The result is that the research, although internally valid, lacks ecological validity (i.e., realism) and external validity (i.e., the results lack generalizability to other settings).

It is not only the use of simulations that leads to compromised findings for the implementation of legal policy, however. Even ecologically valid research can lack external validity. The results of studying mediation in one jurisdiction may not be generalizable to other jurisdictions because of differences in mediators and mediation procedures (Beck & Sales, 2001). Thus, it is important that in training forensic psychology scientists and in educating the consumers of their research, training programs include specific educational opportunities that focus on understanding the needs of the law and how research can and cannot address those needs under varying conditions.

EMPIRICAL RESEARCH ON FORENSIC PSYCHOLOGY TRAINING

Ultimately, forensic psychology training, like training in all fields, will need to be scrutinized empirically if we are to discern the best pathways to improving forensic competence in practice. This is not a simple task, however, given that the learned professions, including forensic psychology, have never embraced such a research agenda. The problem for forensic psychology is that such research will be confounded by the validity of the knowledge base that is imparted during training. For example, if a training approach is not shown empirically to produce competence in practice, is it the result of the approach or the lack of validity in the treatment method taught? Although a problematic issue for research design, the importance of the larger issue still stands—the need to study empirically the outcomes of different training approaches—and deserves serious scholarly attention in the future.

REFERENCES

- 2012 California Rules of Court 5.225 (d)–(i). Retrieved from http://www.courtinfo.ca.gov/cms/rules/index.cfm?title=five&linkid=rule5_225
- Aderhold, B., Boulas, J., & Huss, M. T. (n.d.). Guide to graduate programs in forensic psychology. Retrieved August 20, 2012, from <http://www.ap-ls.org/education/GraduatePrograms.php>
- American Board of Forensic Psychology. (2006). ABFP brochure. Retrieved from <http://www.abfp.com/brochure.asp>
- American Board of Forensic Psychology. (n.d.). Apply for ABFP/ABPP forensic board certification. Retrieved from <http://www.abfp.com/certification.asp>
- American Board of Professional Psychology. (n.d.). Home page. Retrieved from <http://www.abpp.org/i4a/pages/index.cfm?pageid=3285>

- American Chemical Society. Forensic chemists. Retrieved from http://portal.acs.org/portal/acs/corg/content?_nfpb=true&_pageLabel=PP_ARTICLEMAIN&node_id=1188&content_id=CTP_003390&use_sec=true&sec_url_var=region1&__uuid=8bf4cab8-fee6-42e8-918b-f10dc7de774b
- American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist*, *57*, 1060–1073.
- American Psychological Association. (2013). Specialty guidelines for forensic psychology. *American Psychologist*, *68* (1), 7–19. doi:10.1037/a0029889
- American Psychological Committee on Training in Clinical Psychology. (1947). Recommended graduate training program in clinical psychology. *American Psychologist*, *2*, 539–558.
- American Psychology–Law Society. (n.d.-a). Graduate training programs. Retrieved from <http://www.ap-ls.org/education/GraduatePrograms.php>
- American Psychology–Law Society. (n.d.-b). Pre-doctoral internship training programs. Retrieved from <http://www.ap-ls.org/education/Internship.php>
- American Psychology–Law Society. (n.d.-c). Forensic postdoctoral fellowship training programs. Retrieved from <http://www.ap-ls.org/education/PostDoc.php>
- American Psychology–Law Society. (n.d.-d). Certificate programs. Retrieved from <http://www.ap-ls.org/education/Certificate.php>
- American Psychology–Law Society. (n.d.-e). Masters programs. Retrieved from <http://www.ap-ls.org/education/Masters.php>
- Andrews, D. A., & Bonta, J. (2010). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law*, *16*, 9–55.
- Ashford, J. B., Sales, B. D., & Reid, W. H. (Eds.). (2001). *Treating adult and juvenile offenders with special needs*. Washington, DC: American Psychological Association.
- Barefoot v. Estelle, 460 U.S. 880 (1983).
- Bartol, C., & Bartol, A. (2006). History of forensic psychology. In B. Weiner & A. Hess (Eds.), *The handbook of forensic psychology* (3rd ed., pp. 3–28). Hoboken, NJ: Wiley.
- Beck, C. J. A., & Sales, B. D. (2001). *Family mediation: Facts, myths and future prospects*. Washington, DC: American Psychological Association.
- Benjamin, G. A. H., & Gollan, J. K. (2003). *Family evaluation in custody litigation: Reducing risks of ethical infractions and malpractice*. Washington, DC: American Psychological Association.
- Burl, J., Shah, S., Filone, S., Foster, E., & DeMatteo, D. (2012). A survey of graduate training programs and coursework in forensic psychology. *Teaching of Psychology*, *39*, 48–53.
- Contreras, R. (2001, August 10). More courts let experts debunk witness accounts. *Wall Street Journal*, pp. B, P1.
- Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993).
- DeMatteo, D., Marczyk, G., Krauss, D., & Burl, J. (2009). Educational and training models in forensic psychology. *Training and Education in Professional Psychology*, *3*, 184–191.
- English, P. W., & Sales, B. D. (2005). *More than the law: Social and behavioral knowledge in legal decision-making*. Washington, DC: American Psychological Association.
- Findley, J. D., & Sales, B. D. (2012). *The science of attorney advocacy: How courtroom behavior affects jury decision making*. Washington, DC: American Psychological Association.
- General Electric Co. v. Joiner, 522 U.S. 136 (1997).
- Goldstein, A. M. (2007). Forensic psychology: Toward a standard of care. In A. M. Goldstein (Ed.), *Forensic psychology: Emerging topics and expanding roles* (pp. 3–41). Hoboken, NJ: Wiley.

- Grisso, T. (1986). *Evaluating competencies: Forensic assessments & instruments*. New York, NY: Plenum Press.
- Grisso, T. (1987). The economic and scientific future of forensic psychological assessment. *American Psychologist*, 9, 831–839.
- Grisso, T. (2003). *Evaluating competencies: Forensic assessments & instruments* (2nd ed.). New York, NY: Plenum Press.
- Grove, W. & Meehl, P. (1996). Comparing the efficiency of informal (impressionistic, subjective) and formal (mechanistic, algorithmic) predictions procedures: The clinical-statistical controversy. *Psychology, Public Policy, & Law*, 2, 293–323.
- Hafemeister, T. L., Hall, S. R., & Dvoskin, J. A. (2001). Administrative concerns associated with the treatment of offenders with mental illness. In J. B. Ashford, B. D. Sales, & W. H. Reid (Eds.), *Treating adult and juvenile offenders with special needs* (pp. 419–444). Washington, DC: American Psychological Association.
- Harris v. Forklift, 114 S. Ct. 367 (1993).
- In re L.A.M., 2001 WL 246371, Iowa.
- Krauss, D. (2004). Adjusting risk of recidivism: Do judicial departures worsen or improve recidivism prediction under the Federal Sentencing Guidelines? *Behavioral Sciences and the Law*, 6, 731–750.
- Krauss, D., Cassar, D., & Strother, A. (2009). The admissibility of expert testimony in the United States, the Commonwealth, and elsewhere. In D. Krauss & J. Lieberman (Eds.). *Psychological expertise in court* (Vol. 2, pp. 1–24). London, England: Ashgate Press.
- Krauss, D., McCabe, J., & McFadden, S. (2009). Limited expertise and experts: Problems with the continued use of future dangerousness in capital sentencing. In R. Schopp, R. Wiener, B. Bornstein, & S. Wilborn (Eds.), *Mental disorder and criminal law: Responsibility, punishment, and competency* (pp. 135–158). New York, NY: Springer.
- Krauss, D., & Sales, B. D. (2000). Legal standards, expertise, and experts in the resolution of contested child custody cases. *Psychology, Public Policy, & Law*, 6, 843–879.
- Krauss, D., & Sales, B. D. (2001). The effects of clinical and scientific expert testimony on juror decision-making in capital sentencing. *Psychology, Public Policy, & Law*, 7, 267–310.
- Krauss, D., & Sales, B. D. (2003). Forensic psychology, public policy, and the law. In A. M. Goldstein (Ed.), *Forensic psychology* (pp. 543–561). Volume 11 in I. B. Weiner (Ed.-in-Chief), *Handbook of psychology*. Hoboken, NJ: Wiley.
- Kumho Tire Company v. Carmichael, 526 U.S. 137 (1999).
- Lockhart v. McCree, 476 U.S. 162 (1986).
- Malesky, A., & Proctor, S. (2012). Training experiences essential for obtaining a forensic psychology postdoctoral fellowship. *Journal of Forensic Psychology Practice*, 12, 163–172.
- Mancuso v. Consolidated Edison, Co., 2000 U.S. App. LEXIS 20534 (2d Cir. 2000).
- McCleskey v. Kemp, 481 U.S. 279 (1987).
- Melton, G., Petrila, J., Poythress, N., & Slobogin, C. (2007). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (3rd ed.). New York, NY: Guilford Press.
- Miller, M. O., Sales, B. D., & Delgado, J. (2003). *Law and mental health professionals: Arizona* (2nd ed.). Washington, DC: American Psychological Association.
- Monahan, J., Steadman, H. J., Silver, E., Appelbaum, P. S., Robbins, P. C., Mulvey, E. P., ... Banks, S. (2002). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York, NY: Oxford University Press.
- People v. Wheeler, 602 N.E.2d 826 (Ill., 1992).

- Peterson, D. (2003). Unintended consequences: Ventures and misadventures in the education of professional psychologists. *American Psychologist, 58*, 791–800.
- Peterson, C., & Park, N. (2005). The enduring value of the Boulder model: “Upon this rock we will build.” *Journal of Clinical Psychology, 61*, 1147–1150.
- Quinsey, V., Harris, G., Rice, M., & Cormier, C. (2006). *Violent offenders: Appraising and managing risk* (2nd ed.). Washington, DC: American Psychological Association.
- Sales, B. D. (1983). The legal regulation of psychology: Professional and scientific interactions. In C. J. Scheirer & B. L. Hammonds (Eds.), *The master lecture series. Vol. 2, Psychology and the law* (pp. 5–36). Washington, DC: American Psychological Association.
- Sales, B. D., & Shuman, D. W. (2005). *Experts in court: Reconciling law, science, and professional knowledge*. Washington, DC: American Psychological Association.
- Schopp, R. F. (2001). *Competence, condemnation, and commitment: An integrated theory of mental health law*. Washington, DC: American Psychological Association.
- Shah, S. A., & Sales, B. D. (1991). Research in law and mental health: Introduction and Overview. In S. A. Shah & B. D. Sales (Eds.), *Law and mental health: Major developments and research needs* (pp. 1–24). Washington, DC: U.S. Government Printing Office (National Institute of Mental Health Monograph).
- Shuman, D., & Greenberg, S. (2003). The expert, the adversary system, and the voice of reason: Reconciling objectivity and advocacy. *Professional Psychology: Research and Practice, 34*, 219–224.
- United States Code, 18 USC §17.
- United States v. Norwood, 939 F. Supp. 1132 (D.N.J. 1996).
- United States v. Virginia, 518 U.S. 515 (1996).
- Wisconsin Code, §971.15. Mental responsibility of defendant. Retrieved from <http://docs.legis.wisconsin.gov/statutes/statutes/971/16/4>
- Witherspoon v. Illinois, 391 U.S. 510 (1968).

PART TWO

APPLYING PSYCHOLOGY
TO CIVIL PROCEEDINGS

CHAPTER 6

Conducting Child Custody and Parenting Evaluations

PHILIP M. STAHL

IN the past 30 years, there has been a steady growth in the use of psychologists and other mental health professionals in child custody matters (Ackerman & Ackerman, 1997). Evaluations conducted by psychologists assist the court in determining custody, decision making, access, and parenting plans when parents separate or divorce. At the same time, there has been an increase in the number of books devoted to child custody evaluations (Ackerman, 2006; Gould, 1998; Gould & Martindale, 2007; Stahl, 1994, 1999, 2010) and broader forensic psychology practice (Melton, Petrila, Poythress, & Slobogin, 2007; Sparta & Koocher, 2006; Weiner & Hess, 2006).

In considering the necessary ingredients of child custody and parenting evaluation practice, there are many areas in which a psychologist (or other mental health professional) must gain proficiency. At a minimum, these include child development; qualities of parenting; divorce and the impact of the separation and divorce on families; psychological assessment; and “special issues,” such as alienation of children, domestic violence, child abuse, relocation law, family dynamics in cases of extreme conflict, and personality dynamics that contribute to that extreme and ongoing conflict. Finally, evaluators need to have a thorough understanding of the ethical issues that surface when undertaking these complex evaluations for families and the courts.

These evaluations are time and cost intensive and potentially intrusive to the family, and they risk putting the children in the middle of their parents’ conflicts. When ordered by the court to participate in an evaluation, parents are subjected to multiple interviews, perhaps psychological testing, and exposure of their conflicts to teachers, therapists, and other professionals. Children are interviewed and observed

The author would like to thank Robert A. Simon, PhD, and Kathleen McNamara, PhD, PLLC, who provided valuable assistance in reviewing earlier drafts of this chapter.

in offices and their homes. This lengthy process typically takes 3 to 4 months to complete (Bow & Quinnell, 2001) and yields a report that is potentially insightful and potentially damaging to the family.

Child custody and parenting evaluations are among the most difficult and challenging of all psychological evaluations (Bow & Quinnell, 2001). Reasons for this include:

- The number of people and relationships in the family to be evaluated.
- The different ages of the children.
- The range of possible psychopathology.
- The presence of significant situational factors affecting psychological functioning.
- The limitations of psychological tests or interview methods designed for this type of assessment.
- The changing nature of a child's developmental or psychological needs relative to future time-sharing plans.
- The expansive nature of individual questions a court may have about a particular family.

In addition to these complexities, child custody evaluators must have knowledge of relevant statutes and case law. The Model Standards of Practice for Child Custody Evaluation (Association of Family and Conciliation Courts [AFCC], 2006; hereinafter Model Standards) state that "evaluators should be knowledgeable about the legal and professional standards, laws, and rules applicable to the jurisdiction in which the evaluation is requested" (p. 9). In addition, the Guidelines for Child Custody Evaluations in Family Law Proceedings of the American Psychological Association (APA; 2009, hereinafter Guidelines for Child Custody Evaluations), designed to provide guidance to those who perform child custody evaluations, direct that psychologists should strive to gain and maintain specialized knowledge, augment their existing skills, acquire sufficient understanding of the specialized child custody literature, and remain familiar with applicable legal standards in the relevant state in which they practice.¹

The Model Standards state that the child custody evaluation process "involves the compilation of information and the formulation of opinions pertaining to the custody or parenting of a child and the dissemination of that information and those opinions to the court, to the litigants, and to the litigants' attorneys" (AFCC, 2006, p. 6). The goal of the Guidelines for Child Custody Evaluations "is to promote proficiency in the conduct of these particular evaluations" (APA, 2009, p. 1), and the purpose of the evaluation "is to assist in determining the psychological best interests

1. An additional resource for understanding the process of custody evaluations and recommended continuing education for custody evaluators can be found in the California Rules of Court 5.220, 5.225, and 5.230. Although mandatory only for custody evaluators in California, they are a useful reference for all child custody evaluators.

of the child” (p. 5). Ultimately, the purpose of custody and parenting evaluations is to do all of these things. The evaluation process is associated with a dramatically increased risk of licensing complaints against the evaluator (Kirkland & Kirkland, 2001) and is often stressful for the examiner. Thus, evaluators must have the temperament to conduct very comprehensive evaluations and recognize that they may be subjected to anger from parents and an adversarial trial experience. Distorted representations or accusations against the evaluator by one or both parents are not uncommon, both in complaints to the court and in complaints to licensing boards (Kirkland & Kirkland, 2001). Because serious allegations are common to the types of cases that fail at mediation and other attempts at settlement, the evaluator’s recommendations can have particularly significant ramifications for the child’s future.

Although this chapter cannot address all of the issues relevant to child custody and parenting evaluations, it focuses on these issues:

- The best interests of the child standard.
- The purpose of custody and parenting evaluations.
- Ethical considerations.
- Basic research the evaluator must know, especially about children and the impact of divorce on children.
- Critical research in special issues, such as conflict between parents, alienated children, domestic violence, sexual abuse, and relocation.
- The process of custody and parenting evaluation.
- Critical issues in report writing.

BEST INTERESTS OF THE CHILD

It is incumbent on the child custody evaluator to be familiar with the law that governs these issues as they pertain to child custody. In nearly all 50 states and in most Western countries, laws related to the best interests of the child guide decisions about child custody and parenting plans (American Law Institute, 2002; Lewis, 2010). Few states define the term *best interests of the child*, although many identify specific factors that judges are to consider when making decisions about a child’s best interests (e.g., Arizona Revised Statute 25–403; Colorado Revised Statutes 14-10-124), leading some commentators to argue that the best interests standard is not defined (Emery, Otto, & O’Donohue, 2005). Indeed, judges are afforded great latitude to order a parenting plan that they decide is in the child’s best interests. Lewis (2010) argued that “[t]he elegance of the [“best interest”] standard is the simultaneous focus on both the needs of the particular child and, with appropriate weight, the normative child development factors” (p. 21). In jurisdictions where the legislature has identified several specific factors that the judge must consider, the weight assigned to each factor is left to the court.

In nearly all states, there is a “friendly parent provision,” in which the court is to show preference to the parent who is more likely to facilitate the child’s relationship

with the other parent. Likewise, many states (e.g., California and Massachusetts) have a rebuttable presumption that it is not in the best interests of children for a person who commits domestic violence to have custody of his or her child. In some states, the judge may have discretion as to how much weight to give each factor, but in other states, the rebuttable presumption regarding domestic violence may trump the friendly parent provision of the law. In recent years, some states (e.g., Florida and Arizona) have also developed a presumption for joint legal custody on the grounds that the child benefits when both parents are actively involved in parenting, but no states have taken Australia's legislated position for a 50–50 presumption in parenting time provisions (Australia, 1975, as amended in 2006).

The evaluator's task is to gather and present those psychological data related to the best interests factors and answer the questions posed by the court. In Arizona, for example, several best interests factors call for psychological data to be gathered. They include:

- The wishes of the child as to custodian.
- The interaction and interrelationship of the child with the child's parent or parents, the child's siblings, and any other person who may significantly affect the child's best interests.
- The child's adjustment to home, school, and community.
- Which parent is more likely to allow the child frequent and meaningful continuing contact with the other parent.
- Whether one parent, both parents, or neither parent has provided primary care of the child.

Although the evaluator gathers and analyzes data related to the best interests factors, the judge reaches the ultimate determination of the child's best interests based on his or her discretion. In an effort to help judges weigh best interests factors in specific cases, a retired California Superior Court judge (Garbolino, 2001) identified certain traits and issues that are common in divorcing families and suggested that these traits be appropriately weighted in specific cases. The traits include those listed next.

- Quality of parenting (strengths and weaknesses of each parent).
- Quality of coparenting (parents are effective at communicating and making decisions on behalf of the child, or they are in high conflict and unable to do so).
- Substance use/abuse issues.
- Mental health issues.
- Each parent's support of the child's relationship with the other parent.
- History of conflict resolution.
- Presence or absence of domestic violence—and if present, whether coercive control dynamics exist or not.
- The child's psychological, developmental, academic, and social functioning.
- The child's special needs, if any.

Garbolino (2001) suggested that the factors likely to be of importance to the well-being of children vary across families and need to be weighed with discretion. Regardless of whether a state has specific factors delineated in its best interests statute or whether there is a presumption associated with a particular factor, the ultimate decision about weighting of these factors is unique to each family and is left to judicial discretion. When completing an evaluation of a given family, the evaluator's task is to provide a rationale as to why different factors might be more or less relevant with that particular family. This rationale, of course, will be reflected in the analysis of the data gathered and in the recommendations provided to the court. This analysis can assist the judge in considering and weighting the issues. This topic is discussed more fully in the section on report writing near the end of this chapter.

PURPOSE OF A CUSTODY AND PARENTING EVALUATION

Custody and parenting evaluations serve important purposes for the court and for the family.

FOR THE COURT

The primary purpose of the evaluation is to assist the court in case a settlement is not reached. Judges order child custody evaluations for a variety of reasons. These can include those circumstances in which there are significant allegations regarding drug and alcohol abuse, family violence, or child abuse, or significant mental health problems. Often a judge is presented with two parents, both of whom appear good enough at parenting but who cannot agree on a parenting plan. At other times, one or both parents appear to have significant problems. Increasingly, judges look to mental health professionals to help them understand complex psychological questions of attachment between the child and his or her parents, sibling relationships, and the developmental needs of children.

In a mobile society, one parent may wish to relocate with the children for a variety of reasons, such as employment, economics, family support, or a new relationship. Judges may order a child custody evaluation to address the relevant psychological factors associated with the relocation question. For example, in California, judges frequently request the assistance of an evaluator in a relocation case in order to provide information to the court about the relevant psychological issues described in the *LaMusga* decision (*In re Marriage of LaMusga*, 2004). Among the many issues identified in the case were various factors the court would likely consider in a relocation matter. This guidance from case law also helps custody evaluators in California focus on those relevant psychological issues. Similar case law decisions in other states would guide custody evaluators as well.

Although judges are guided by the law in making decisions regarding the best interests of children, they may look to the child custody evaluator to assist in

understanding the family dynamics and the relevant psychological factors in order to reach a decision about what is in the child's best interests. In many ways, the neutrally appointed child custody evaluator serves as a consultant to the judge, providing critical data about the family for a better understanding of the family dynamics and the needs of the children.

FOR THE FAMILY

The majority of separating parents reach an agreement about parenting and custody issues on their own or with the assistance of their attorneys or legal processes, such as mediation (Melton et al., 2007). These parents agree on important decisions about their children, such as where they will go to school, what extracurricular activities they will participate in, and what nonemergency medical procedures to consider (e.g., orthodontia). They agree on a parenting plan that will delineate the times their children spend with each parent. Obviously, these families do not need to undergo an evaluation, nor would the court order one. In fact, given the potentially intrusive nature of child custody evaluations, in such situations it is highly likely that an evaluation would be *harmful* to the family.

However, as many as 20% of families at some point in time after separation experience high conflict (Ahrons, 2005). They cannot agree on the parenting plan or how to make decisions for their child, nor are they able to focus on their child's needs because of their mutual differences. These parents may disagree on seemingly simple and minor issues, such as what time a child exchange is to take place or whether it is appropriate for the children to eat certain foods. With many of these families, there are allegations of domestic violence or significant mental health problems, alienation or estrangement of children, or substance abuse. The primary value of an evaluation in these circumstances is that the evaluation provides an opportunity for parents to voice their concerns to a neutral expert. A neutrally appointed child custody evaluator will spend considerable time with both parents trying to understand their concerns and their perceptions of their child's needs. This can be comforting to parents and sometimes serves as a catalyst for them to move toward cooperation.

By listening to children, evaluators can also identify when they are caught in a loyalty conflict between their parents and describe the impact of this conflict to the parents and the court. It is common for children's voices to be absent in the courts in the United States, and participation in a child custody evaluation can help children voice their concerns, share their wishes, and explore their feelings. Although the child custody evaluator is not serving as a therapist, the evaluation process may be therapeutic to children who participate in the evaluation. If the evaluator concludes that the child is experiencing significant problems, he or she can refer the child for therapy and help the parents understand their child's developmental needs. In these ways, the evaluator serves to hear the child's voice and advocate for the child's psychological and developmental needs.

A third potential benefit comes from the fact that the same mental health professional is observing all family members. In the midst of an acrimonious divorce, services are often fragmented, with each parent having his or her own therapist and attorney whose role is to advocate for their respective client. These professionals do not have critical information about the other family members and are not neutral with regard to the family (Greenberg, Doi Fick, & Schnider, 2012). Although their advice might be helpful and appropriate to their clients, their perspectives and recommendations may not be in the child's best interests. Even a child's therapist may not be serving the child's best interests if the therapist is only seeing one parent with the child and becomes an advocate for that parent (Greenberg, Gould, Gould-Saltman, & Stahl, 2003). By having a neutrally appointed child custody evaluator listen to and observe all family members, interview relevant collateral witnesses, and consider everyone's input before reaching conclusions about the children's best interests, fragmentation is reduced.

At the end of the evaluation process, when an evaluator writes a comprehensive report, parents benefit by learning about their child's needs and how they can work together to meet those needs. The report can help parents focus on the child rather than on their conflicts with each other and can help them learn ways to resolve their conflicts and meet their child's needs. Furthermore, the evaluator can help parents understand relevant issues important to parents, such as when and how to incorporate overnight time with each parent for their young child, the impact of their conflict on their children, and the risks and benefits of shared parenting. This understanding will enable them to parent more effectively.

Ultimately, an evaluation is most helpful to the family when the report and conclusions reduce conflict, help parents reach an agreement without going to trial, and keep parents focused on their child's needs and best interests. Well-done evaluations often help parents recognize the need for solution and compromise, and, while mediation may not have been successful prior to the evaluation, settlement may be much more likely after an evaluation (R. K. Kelly & Ramsey, 2009).

ETHICAL CONSIDERATIONS IN CHILD CUSTODY EVALUATIONS

Along with any state or local court rules or statutes, and in addition to the AFCC Model Standards (2006) and the APA Guidelines for Child Custody Evaluations (APA, 2009) described earlier, a number of other advisory documents guide the ethical practice of the child custody evaluator, including:

- *Ethical Principles of Psychologists and Code of Conduct* (EPPCC; APA, 2002)
- *Specialty Guidelines for Forensic Psychology* (hereinafter Specialty Guidelines, APA, 2013; the Specialty Guidelines are reprinted as the appendix to this volume with permission of the APA)
- *Practice Parameters for Child Custody Evaluation* (American Academy of Child and Adolescent Psychiatry [AACAP], 1997)
- *Guidelines for Brief Focused Assessment* (AFCC, 2009)

These documents provide considerable guidance for the evaluator. Child custody evaluators who are not highly familiar with all of these documents are working at a considerable disadvantage and, therefore, are at increased risk for failing to maximally serve the court and family.

Maintaining Specialized Competence. Child custody evaluations are a unique type of evaluation, one that requires specialized competence. EPPCC Standard 2.01a states, “Psychologists provide services . . . with populations and in areas only within the boundaries of their competence, based on their education, training, supervised experience, consultation, study, or professional experience” (APA, 2002, p. 4). The Model Standards list relevant areas for continuing education training for those performing child custody evaluations (see AFCC, 2006, Standard 1.2). The Model Standards list 18 areas of expected training for all child custody evaluators and 5 areas of specialized training for those evaluators performing custody evaluations in those particular areas. These areas of expected training include, among other items:

- The psychological and developmental needs of children.
- The effects of separation, divorce, domestic violence, substance abuse, child alienation, child maltreatment, and interparental conflict on the psychological and developmental needs of children.
- How to assess parenting capacity and coparenting capacity and construct effective parenting and coparenting plans.

Additionally, the Specialty Guidelines state, “Forensic practitioners make ongoing efforts to develop and maintain their competencies . . . [and] keep abreast of developments in the fields of psychology and law” (see APA, 2013, Standard 2.02). These several documents guide the evaluator in developing and maintaining ongoing continuing education in areas relevant to child custody evaluations.

Avoiding Conflicts of Interest. The APA Guidelines for Child Custody Evaluations state, “Psychologists strive to avoid conflicts of interest and multiple relationships in conducting evaluations” (see APA, 2009, Guideline 7, p. 11), as certain prior roles may impair the objectivity of the child custody evaluator. Furthermore, the Guidelines for Child Custody Evaluations advise against performing a child custody evaluation if the psychologist has provided therapeutic services to any of the parties in the past or present. Evaluators may also consider disclosure of other preexisting professional roles with any family member—such as having been the mediator; personal relationships, such as having one’s child on the same Little League team as the children in the family to be evaluated; or even a close relationship with one of the attorneys—before accepting an appointment. The appearance of conflict may be equally important to actual conflict in these cases. Although there may not be

any actual conflict, and although no ethics will have been violated by taking on the evaluation role in these circumstances, evaluators should recognize the risk that a parent who feels wronged by the evaluator's recommendations might allege that the evaluator was biased because of these prior relationships. If such a parent later finds out about these relationships, it may serve as further reinforcement of such bias in the mind of the parent. Advance disclosure of all prior relationships helps reduce the risk of such allegations.

Obtaining Informed Consent. Even though child custody evaluations typically are court-ordered, parents still need to understand the process. Technically, informed consent is not obtained when the court orders an evaluation; instead, custody evaluators are encouraged to obtain informed consent both in writing and orally at the start of the evaluation process (APA, 2009). The document needs to explain critical issues, such as the general procedures that will be used, each parent's role in the evaluation process, fees, and the limits of confidentiality. The evaluator needs to inform parents that a child custody evaluation is not a health-related procedure and that the evaluator will not bill a parent's health insurance. Additionally, because the EPPCC requires psychologists to avoid doing harm when it is foreseeable (APA, 2002), the evaluator should inform parents that one or both of them may be unhappy at the end of the evaluation process. It is recommended that the evaluator provide this document to the parents and their attorneys in *advance* of the start of the evaluation. Finally, it is also important for the evaluator to inform potential collateral sources of the limits of confidentiality and the purpose for which the collateral information is being gathered (AFCC, 2006).

Employing Balanced and Impartial Procedures. According to AFCC Model Standard 5.5, child custody evaluators strive to use a balanced process in order to achieve objectivity, fairness, and independence:

As one element of a balanced process, the evaluative criteria employed shall be the same for each parent-child combination. In the interests of fairness and sound methodology, evaluators shall ensure that any allegation concerning a matter that the evaluator is likely to consider in formulating his/her opinion shall be brought to the attention of the party against whom the allegation is registered so that s/he is afforded an opportunity to respond. (2006, p. 15)

When the evaluator does not act in this way, it almost assuredly leads to a complaint of bias, sometimes made to the parent's attorney, but potentially to the court or practitioner's licensing board.

Using Multiple Sources of Information. According to Guideline 10 of the APA Guidelines for Child Custody Evaluations, "Multiple methods of data gathering enhance

the reliability and validity of psychologists' eventual conclusions, opinions, and recommendations. Unique as well as overlapping aspects of various measures contribute to a fuller picture of each examinee's abilities, challenges, and preferences" (2009, p. 14). These multiple methods (discussed in greater detail later) usually include, at a minimum:

- Multiple interviews with the parents.
- Interviews with children when appropriate.
- Observations of children and parents interacting.
- Administration of psychological testing and parenting questionnaires.
- Review of collateral documents.
- Interviews with relevant professionals, family members, and friends.

Staying Within the Scope of the Evaluation. Guideline 8 of the APA Guidelines for Child Custody Evaluations states:

Before agreeing to conduct a child custody evaluation, psychologists seek when necessary to clarify the referral question and to determine whether they are potentially able to provide opinions or recommendations. It may be helpful to have psychologists' understanding of the scope of the evaluation confirmed in a court order, or by stipulation of all parties and their legal representatives. (2009, p. 12)

Similarly, AFCC Model Standard 5.1 states:

Evaluators shall establish the scope of the evaluation as determined by court order or by a signed stipulation by the parties and their attorneys. If issues not foreseen at the outset of the evaluation arise and if it is the evaluator's professional judgment that the scope of the evaluation must be widened, the evaluator shall seek the approval of the court or all attorneys prior to going beyond the originally designated scope of the evaluation. (2006, p. 13)

Both of these require that the evaluator carefully consider the relevant issues in the case and make recommendations consistent with those issues. Identifying the scope in advance of performing the evaluation also ensures that the evaluator has the necessary specialized training to conduct the evaluation, as noted earlier.

Differentiating Observations, Inferences, and Conclusions. In the Specialty Guidelines, psychologists are reminded: "In their communications, forensic practitioners strive to distinguish observations, inferences, and conclusions. Forensic practitioners are encouraged to explain the relationship between their expert opinions and the legal issues and facts of the case at hand" (APA, 2013, Guideline 11.02, p. 16). Additionally, Specialty Guideline 11.03 states, "Forensic practitioners are encouraged to disclose all sources of information obtained in the course of their professional services, and to identify the source of each piece of information that was considered and relied

upon in formulating a particular conclusion [or] opinion” (p. 17). Ultimately, this guideline is to help parents understand the rationale for recommendations but also to assist the court in understanding the evaluator’s reasoning. Within this context, it is equally important to provide a description of the risks and benefits of different options available to the court. This topic is discussed in greater detail in the report writing section further on.

Record Keeping. Because records are subject to subpoena and full disclosure is important in the interest of transparency and due process, the APA Guidelines for Child Custody Evaluations direct evaluators to keep complete, readable records with the expectation that others will review them in the event of ongoing litigation after the completion of the report:

Legal and ethical standards describe requirements for the appropriate development, maintenance, and disposal of professional records. The court expects psychologists providing child custody evaluations to preserve the data that inform their conclusions. This enables other professionals to analyze, understand, and provide appropriate support for (or challenges to) psychologists’ forensic opinions. (2009, p. 18)

BASIC CRITICAL AND RELEVANT RESEARCH, ESPECIALLY ABOUT CHILDREN

Given the admonition about training, this next section addresses basic research with which all custody evaluators should be familiar.

Divorce Research. The impact of divorce on children has been studied for more than 30 years. Global findings include that there is an increased risk of harm to children when their parents divorce and that most of the harm comes from the children’s exposure to poverty, conflict, limited access to one of their parents, and poor parenting (Hetherington & Kelly, 2002; J. B. Kelly & Emery, 2003; Wallerstein & Kelly, 1980). Research also demonstrates that children are resilient and that the majority suffer no significant negative long-term effects as a result of the separation or divorce of their parents (J. B. Kelly & Emery, 2003). Children benefit when they are exposed to limited parental conflict and have regular and frequent contact with both of their “good-enough” parents, and when both parents can participate in a wide range of their child’s life experiences and activities (J. B. Kelly, 2012; Lamb, Sternberg, & Thompson, 1997).

Parenting Plans for Young Children. The term *parenting plans* refers to both the legal custody and physical custody of children. *Legal custody* refers to who has the authority to make important decisions about the children, whereas *physical custody* refers to where the children reside. Some states, however, have changed statutory language and no longer use terms such as *legal custody*, *physical custody*,

and *visitation*. In those states, statutory language refers to the child's residential schedule and parents' decision making. This section of the chapter focuses on the residential schedule identified in parenting plans.

Before 2000, much of the divorce literature suggested that very young children should not have overnight visits with their "noncustodial parent" until sometime between the ages of 3 and 5 (Hodges, 1991; Stahl, 1994). J. B. Kelly and Lamb (2000) were the first to identify the importance of regular overnight visits between very young children and their parents. This article led to some controversy (Solomon & Biringen, 2001), and in recent years, it has again been suggested that overnight visits may be harmful for some children (George, Solomon, & McIntosh, 2011). Gould and Stahl (2001) provided a framework in which to consider risks and benefits associated with overnight visits. Most recently, Ludolph and Dale (2012) surveyed the attachment, fathering, and divorce research and concluded that, although attachment is one factor to consider when designing parenting plans for young children, other factors, including the history and nature of father involvement, quality of parenting by each parent, extent and type of conflict between the parents, regular day care that might be provided by third parties, and logistics, are relevant as well. If conducting a child custody evaluation where young children are at issue, it is critical to know this research.

Shared, 50–50, or Sole Physical Custody? As noted, many states have statutes that demonstrate a preference for arrangements that allow the child to have frequent and continuing contact with both parents. For some states, this means substantial parenting time in which both parents are actively involved in the child's life experiences and activities (Lamb et al., 1997). McIntosh and Smythe (2012) and Fabricius, Sokol, Diaz, and Braver (2012) described the benefits and potential risks of a 50–50 parenting plan for children in families with different dynamics. Although most research indicates that children clearly benefit from the active involvement of two good parents, a range of logistical and other factors might interfere with equal parenting time plans. Furthermore, there is no research that indicates that a 50–50 parenting plan is, a priori, best for children.

Conflict, Legal Custody, and Decision Making. Less research has focused on legal custody and decision making between parents than on residential schedules. For most parents, conflict decreases over time (Ahrons, 2005; Maccoby & Mnookin, 1992). In most circumstances, parents will share in the legal custody and decision making for their children, but there will be times when this is not feasible. Typically, in families experiencing coercive-controlling violence (discussed later under "Domestic Violence"), in situations where mental illness and/or substance abuse impairs parenting, and in those families in which parents dislike each other more than they love their children, it may be necessary to have decision making vested in one parent. When this is necessary, it is primarily so that decisions can be

made in a timely and conflict-free way. Parenting coordinators (Sullivan, 2004), who work with a family to help resolve conflicts on an ongoing basis, usually after there are court orders in place, may prove of value when certain high-conflict dynamics exist. This process helps many families avoid frequent returns to court and enables decisions to be made for the benefit of children more efficiently (Sullivan, 2004).

CRITICAL RESEARCH IN SPECIAL ISSUES

In addition to the basic research just described, many child custody evaluations involve special issues, including allegations of domestic violence, sexual abuse, children becoming alienated, and relocation. Even more specialized knowledge is critical when performing evaluations in these areas (see AFCC, 2006, Model Standard 1.2 (c)).

ALIENATED CHILDREN

First discussed by Wallerstein and J. B. Kelly (1980) as children who refused visitation with one parent, Gardner (1987) described what he termed parental alienation syndrome, a phenomenon he described as occurring when a mother (he changed this term to gender-neutral wording in the late 1990s) who harbored extremely negative feelings about the father influenced the child to reject the father. J. B. Kelly and Johnston (2001) discussed “the alienated child” and identified a range of factors that result in children becoming alienated from a parent. Drozd and Olesen (2004) took this a step further when identifying the differences between children who are realistically estranged from one parent and those who have developed an unjustified alienation in rejecting that parent (see also Stahl, 2004, 2010; Warshak, 2001). More recently, a special issue of *Family Court Review* (2010) was devoted to the topic of alienation with a special focus on intervention.

DOMESTIC VIOLENCE

With research suggesting that nearly half of all divorcing families experience some type of domestic violence (Dalton, Drozd, & Wong, 2005), it is not surprising that many child custody evaluations involve such allegations. For many years, the domestic violence literature focused primarily on what has been termed intimate partner violence, and the child custody evaluation literature focused on risk assessments with these families (Austin, 2001). However, not all domestic violence involves the use of power and control as a primary dynamic such as is found in coercive controlling violence (J. B. Kelly & Johnson, 2008). In 2007, the AFCC and the National Council of Juvenile and Family Court Judges convened a symposium to discuss the differentiation of domestic violence cases and develop a recognition that one size does not fit all (*Family Court Review*, 2008). This resulted in a more nuanced understanding of family violence and the range of parenting plan options appropriate for these different families.

Differentiation of types of domestic violence (Jaffe, Johnston, Crooks, & Bala, 2008; J. B. Kelly & Johnson, 2008) suggests that, in some families, violence is unique to the separation of the parents (separation-instigated violence). In these cases, after violence occurs, parents are ashamed of their actions, readily admit to what they have done, and vow never again to engage in violent or abusive behavior. In other families, parents experience situational couples violence, in which neither parent is afraid of the other, both initiate the violence, and there is no evidence of power and control being utilized by either parent. In perhaps the smallest subset of families, but in the most serious of situations, the violence appears to be coercive and controlling in nature. In these coercive-controlling violence families, it is most common for males to aggress, the victim and children are fearful in response to physical violence, and there is often co-occurring emotional abuse, sexual abuse, or control of family economics (Jaffe et al., 2008).

SEXUAL ABUSE

Perhaps the most emotionally charged of cases are those in which there are allegations of child sexual abuse. The challenge in these cases is that the allegation usually sets in motion several events, including but not limited to:

- Independent investigations by child protective services and law enforcement authorities.
- Criminal charges.
- A temporary order suspending or supervising contact between the child and the alleged offender.
- Emotionally charged court hearings in which the alleged offender denies the allegations and claims the allegations are made for purposes of custody and the other parent simply claiming protection of the child.
- A child custody evaluation designed to more fully evaluate the allegations and make recommendations for a parenting plan.

Kuehnle and Connell (2009) focus on the range of hypotheses that must be considered in any case when such allegations are raised and the thoroughness of the evaluation process required. They also identify that the primary role of the evaluator is to perform a risk assessment in these cases.

RELOCATION EVALUATIONS

The one area where many judges and evaluators have the most trouble making decisions is in relocation cases, which pit the right of adults to live wherever they want and the right to parent. These cases come to the court when one parent wants to move with the child and the other parent opposes the move and wants the child to remain. Child custody evaluators are at risk of confounding the research when performing these evaluations by recommending against moves because of research

that demonstrates the benefit to children when both parents maintain ongoing and regular access with their children (Austin, 2000). There is limited research on the effect of parental relocation on children. College students reported greater current problems when they lived more than 100 miles from their other parent during their childhood as a result of the relocation of either parent (Braver, Ellman, & Fabricius, 2003). Drawing on other divorce research but in the absence of relocation research, some have suggested courts allow relocation by the primary custodial parent (Wallerstein & Tanke, 1996) whereas others have suggested that courts should almost routinely order a provisional change of custody to the parent who is not relocating when the primary custodial parent wants to move (Braver et al., 2003). Stahl (2006) has suggested that for-the-move or against-the-move biases operate in those circumstances and should not result in either of these presumptions.

The primary focus in recent years has been on the consideration of risk and protective factors in determining the ultimate decision in these cases (Austin, 2008b; Parkinson, Cashmore, & Single, 2010; Stahl, 2010). Some factors are likely to increase risks for children if they move (e.g., interfering with the relationship with the other parent, starting new schools and having to make new friends and activities, and having a residential parent who is a restrictive gatekeeper and who interferes with the child's relationship with the other parent by discouraging access and limiting communication about the child with the other parent). Other factors will be protective and minimize risks for children if they move (e.g., moving to be nearer to extended family and additional relatives who serve as social capital, children who are adaptable and make changes in their life easily, and having a residential parent who is a facilitative gatekeeper and who supports the child's relationship with the other parent by encouraging access and promoting healthy coparenting communication). Some children who move are harmed (Braver et al., 2003; Taylor, Gollop, & Henaghan, 2010) while others are not (Taylor et al., 2010). Parkinson et al. (2010) wrote that, while it is tempting to resolve these difficult cases with the assistance of wishful thinking, research is needed to test that wishful thinking against the realities of experience. They do not believe there is ample research support to conclude that children who relocate with one parent while the other parent is left behind will, by virtue of the relocation, automatically do well or will be harmed. Rather, the only way to understand the optimal relocation decision in a given case is by focusing on the risk and protective factors existing in that case.

PROCESS OF CONDUCTING CUSTODY AND PARENTING EVALUATIONS

Child custody and parenting evaluations are very different from other psychological or forensic evaluations. They are more complex, involve more people, and entail more procedures than most. These evaluations require a forensic mind-set versus a therapeutic mind-set and the exploration of multiple hypotheses. Typically, there will be allegations made by one parent against the other, and it is not unusual for the evaluator to be unable to reach conclusions about the he-said, she-said allegations

in the case. Each step of the evaluation process is designed to help the evaluator gather information critical to understanding the family. This section addresses some of these critical issues.

GETTING STARTED

In most jurisdictions, a custody or parenting evaluation will be ordered by the court or stipulated to by the parties. It results in appointment of one neutral evaluator focused on assessing all relevant issues in dispute. From a risk management perspective, it is important to receive the court order before beginning the evaluation, as the authority to conduct the evaluation comes from the court. In many jurisdictions, the court order provides the evaluator with quasi-judicial immunity and includes information helpful to determining the scope of the evaluation. Additionally, the court order usually provides information about who is responsible for payment and identifies when the report is due, who is entitled to receive a copy, and the limits of confidentiality. After receiving the court order, it is common for the evaluator to have a joint conference call with the attorneys to gather basic information about the family and the reasons for the evaluation. Although some attorneys like to argue their case for the evaluator, it is best to get some basic facts and reasons for the evaluation during this call and lay out the logistics and proposed time frame for the evaluation. During this call, it is helpful for the evaluator to explain procedures and request documents to be reviewed.

The next step is scheduling initial appointments with the parents and sending them necessary paperwork, with copies to their attorneys. This paperwork is likely to include the retainer agreement and an intake form. As noted, the retainer agreement describes the evaluator's and the parents' obligations through the evaluation process, limitations regarding confidentiality, and other critical information about the evaluation process. It serves as a detailed informed consent document, which is recommended even if the parents have been ordered to participate in the evaluation.

INTERVIEWS WITH THE PARENTS

A good way to start the first evaluation interview with each parent is to ask the parent, "Why are we here?" This question allows the parent to explain his or her concerns, observations, beliefs, and allegations in a rather open-ended manner. With limited prompting (e.g., "Tell me more"), the evaluator can spend much of the first appointment trying to understand the parent's issues, concerns, and proposed solutions. Parents often have a need to be heard, and focusing on the matters important to them during the first interview facilitates cooperation and participation.

During the interviews, it is important for the evaluator to focus on each parent's:

- Concerns and allegations.
- Responses to the allegations and concerns raised by the other parent.

- Understanding of the child and his or her psychological, social, academic, and developmental functioning.
- Description of the history of the relationship between the child and each parent.
- Description of his or her own family history, especially focusing on relevant issues that may relate to the current evaluation.
- Beliefs about the strengths and weaknesses of his or her own and the other parent's parenting.
- Description of the coparenting relationship and the ability of each parent to communicate with the other and make day-to-day decisions on behalf of the children.
- History of and ability to support the child's relationship with the other parent and if there are concerns about this moving forward.
- Understanding of the special issues in the case (e.g., relocation) and how it may affect a parenting plan.
- Recommendation for the specific parenting plan.

In most evaluations, this information can be gathered in three to four interviews, each of which might last 2 hours. It is important to gather the information that each parent wants to relay, but the examiner must be more than a stenographer and seek enough depth and breadth associated with these issues while simultaneously having an opportunity to ask each parent about the concerns raised by the other parent.

In evaluations with more complex issues, the evaluator will want to explore those in depth. For example, in a case with allegations of domestic violence, it is important to understand the parents' description of the history of conflict resolution, whether there are issues of power and control, and the parents' description of physical, verbal, and emotional abuse that may have occurred. If a parent describes physical violence, it is important to ask about the first, most recent, and worst incidents of abuse. It is always important to explore for more than what the parent initially describes, since many domestic violence victims are reluctant to share details of the abuse.

In cases with allegations of alienation, it is important to explore each parent's history of involvement with the child, each parent's perception of his or her own and the other parent's contribution to the child being alienated, and the extent to which the child is rigid in his or her rejection of one parent. The evaluator should explore whether the child has a realistic basis for being estranged from one parent or whether other dynamics are contributing to this alienation.

Finally, in relocation cases, in addition to best interests statutes, it is important to understand specific statutory or case law pertaining to relocation. In some jurisdictions, there is a presumptive right to move by a custodial parent, whereas in other jurisdictions, the burden is on the parent requesting to relocate to show that the move is in the child's best interests. In still other jurisdictions, every relocation matter is considered on a *de novo* basis (i.e., a new hearing on the best interests of the child). Evaluators need to understand these state-specific legal issues in relocation matters.

When interviewing parents in relocation cases, evaluators must:

- Ask questions to understand the motives for relocation and the motives for opposing it.
- Understand how each parent perceives the child will be affected by the move, both positively and negatively.
- Collect family information to understand the social capital in each community (Austin, 2008a).
- Gather information from each parent about a proposed parenting plan should the court allow the move, or should the court not allow the move, or should both parents end up in the same location.

INTERVIEWS WITH THE CHILDREN

Interviews with the children are a crucial part of understanding both the family dynamics and the relationship between the child and his or her parents. Evaluators should:

- Start by establishing rapport with the children.
- Begin by discussing the process of the evaluation, the limits of confidentiality, and the structure of the interview process.
- Encourage children to talk openly about their feelings and help them understand that the evaluation is about their interests and not their parents' wishes.
- Inform the children that a report will be submitted to the judge, which the parents will probably read.
- Tell children that they do not have to answer questions they do not want to answer and that their parents or the judge will ultimately decide where and how they will spend time with their parents.

Evaluators must recognize that children's language skills are not the same as adults'. It is important to know that, although children often do not understand their questions, they may respond as if they do. It may be useful to ask children to repeat or to explain the questions to be sure that they understand them.

How questions are asked affects the way answers are given. When interviewing children, particularly in a forensic context, it is vital to ask open-ended questions (Lamb, Hershkowitz, Orbach, & Esplin, 2008). These questions are far more likely to yield useful, accurate, and honest responses. Asking leading or categorical questions limits the way that the child responds and, therefore, limits the usefulness and validity of those responses.

It is important for the evaluation process to be balanced. As such, it is also important for children to be seen with each parent bringing them to the office. Evaluators must keep in mind that one or both parents may influence their children. To reduce the risks associated with this influence, appointments should be scheduled equally with each parent bringing the children to the appointments. Although children's

suggestibility and the potential for being influenced by parents or siblings is a topic that is beyond the scope of this chapter, it is critical for those evaluating custody and parenting plans to understand this research.

Specific data are important to gather during interviews with children. These include the child's:

- Likes and dislikes, interests, friends, chosen activities, and other aspects of the child's day-to-day life.
- Schooling, including information about how each parent participates in helping with homework and other school-related matters.
- Perceptions of his or her relationships with each parent, including things that the child likes and does not like about each parent.
- Perceptions of discipline.
- Routines in each home and how the child deals with any differences in routines between homes.
- Typical mood, and how the child typically expresses his or her feelings and if there is a difference for each parent.
- Perceptions about the need to care for his or her parents emotionally.
- Anything else the child wants the judge to know.

These data provide important information to the court about the child's life. Because the evaluator is the only unbiased person providing information to the judge about the child (it is assumed that both parents will be biased), such information is vital to the court in helping it to make the ultimate decision about custody and parenting plans.

OBSERVING PARENTS AND CHILDREN TOGETHER

A fundamental purpose for observing children is to understand the nature of the bond between a child and the parents. Although there is no reliable and valid way of measuring whether a child is more bonded to one parent or the other, the job of the evaluator is to describe the behavioral dynamics of the bond for the judge. In young children, the evaluator should observe the way children and parents relate with one another. Do they play together, smile and laugh with one another, exchange affection with one another, or stay relatively distant and isolated from one another? Does the child seem attentive to the parent when the parent enters the room, or does the child seem disinterested? When parents are in the room, it is important to listen to what they say. Parents may want to talk about things that are inappropriate to discuss in front of the child, because they have a need to provide more information to the examiner. The observation session is not a good time for this so it is always important for evaluators to understand each parent's ability to utilize adequate boundaries and keep the child free from anxiety. If the parent offers inappropriate comments in front of the child (e.g., something negative about the other parent or something about the litigation), the examiner should try

to understand how the child feels about it, responds to it, and interacts with the parent about it. For example, some children get into arguments with their parents about things that parents say, and this provides valuable information about the interaction between parent and child.

It is often helpful to provide tasks for the parent and child to complete. Encouraging a father and daughter to draw a picture, for example, will provide data about how they work together to complete a task. Are they cooperative, are they playful, do they use each other's assistance, or do they become quite competitive with one another? This can help the evaluator develop hypotheses about the child's relationship with the parent, which will need to be verified in other ways (e.g., with collateral sources or interviews). Unstructured play, in which the child initiates an activity of his or her choosing, provides an opportunity to see how responsive the parent is to the child in *his or her* space. Many parents can interact quite well with their children when they choose the activity, but they may feel awkward and insecure when their children choose the activity. At the same time, the examiner must observe the affect of the parents and children. Are they relaxed and having fun, or is there tension between the parent and child just as there is between the parents?

Finally, with older children and their parents, it is important to talk about the routines, day-to-day life in each parent's home, and how they and their parents deal with conflicts. Examiners should pay particular attention to disparities between what the child says during individual interviews compared with the observation sessions. It is particularly important to explore a range of feelings between the child and parents in those families where alienation or estrangement is alleged.

PSYCHOLOGICAL TESTING AND PARENTING QUESTIONNAIRES

Use of psychological testing in custody evaluations, although common, is not mandatory. No psychological tests measure the quality of parenting or coparenting, which are critical issues in child custody and parenting evaluations. There are also no valid psychological tests designed for use with the specific child custody population. Although there have been efforts to develop some psychological instruments relevant to child custody (e.g., Ackerman & Schoendorf, 1992; Bricklin, 1989, 1990a, 1990b), these instruments are of limited validity (Otto, Edens, & Barcus, 2000). The most recent effort to develop an instrument to understand children's views of their relationships, the Structured Child Assessment of Relationships in Families (SCARF; Strachan, Lund, & Garcia, 2010), does not have sufficient validity and reliability, although research on this instrument is ongoing. Thus, no specific instruments can directly assess the complex issues inherent in these evaluations.

The AFCC Model Standards direct child custody evaluators to "be prepared to articulate the bases for selecting the specific instruments used" (2006, p. 17) and to use assessment instruments "for the purpose for which they have been validated" (p. 18). The Model Standards add that "[c]aution should be exercised...when utilizing computer-generated interpretive reports and/or prescriptive texts"

(Standard 6.6, p. 18). According to the APA Guidelines for Child Custody Evaluations, “Psychologists strive to interpret assessment data in a manner consistent with the context of the evaluation” (2009, p. 15). These Standards and Guidelines suggest that child custody and parenting evaluators must be careful in choosing assessment instruments, understand the research associated with custody litigants and their scores on various measures (Bathurst, Gottfried, & Gottfried, 1997; McCann et al., 2001), and be careful when using computer-generated interpretive reports (Flens, 2005). If a psychologist quotes from a computer-generated interpretive report, he or she should identify it as a quote and provide the citation.

Psychologists regularly administer psychological testing as part of a comprehensive, multimethod process of child custody evaluation (Ackerman & Ackerman, 1997). The Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen & Kaemmer, 1989) is the most commonly utilized psychological test administered to parents. Other commonly administered psychological tests include the Personality Assessment Inventory (PAI; Morey, 2007), the Rorschach Inkblot Test (Rorschach, 1998/1921), and the Millon Clinical Multiaxial Inventory-III (MCMI-III; Millon, Millon, Davis, & Grossman, 2009). There is considerable controversy about using either the MCMI-III or Rorschach in child custody work (Flens, 2005). Examiners should be aware of the controversies and arguments on both sides of the issues when choosing to use those particular instruments, as presented in balanced reviews by Craig (2006), Dyer (2008), Erard (2005), and Evans and Schutz (2008).

Finally, psychological test instruments should be used in a *forensically informed manner*. Unlike the use of psychological tests in therapeutic settings, where the goal is to aid in diagnosis and treatment, tests in child custody and parenting evaluations should be “informed” by the forensic questions that guide the evaluation. Questions pertaining to each parent’s general psychological and behavioral functioning, as it relates to their functional parenting and coparenting capacities, are common reasons for utilizing tests in parenting evaluations. The test data should be used to develop hypotheses about the parent’s psychological and behavioral functioning specifically in terms of how it relates to parenting and abilities to implement a parenting plan. For example, if test data revealed clinical elevations on a scale that measures psychological turmoil, the evaluator might hypothesize that the parent is reacting to a difficult divorce, and if this hypothesis was confirmed by other data gathered during the evaluation, the impact on parenting might be deemed minimal and temporary. Alternatively, it might be hypothesized that the parent has chronic and pervasive deficits in mood regulation, which would pose more serious concerns with respect to parenting and coparenting. The hypotheses generated are evaluated against other data from the evaluation to either confirm or disconfirm those hypotheses.

It is also common to administer some type of parenting questionnaires or instruments to gauge a parent’s stress (Parenting Stress Index-4, Abidin, 2012), measure each parent’s self-report about his or her relationship with the child (Parent

Child Relationship Inventory; Gerard, 1994), and gather structured information about the child (Achenbach Child Behavior Checklist; Achenbach, 1991). It is important to recognize that these instruments are not definitive but also may provide useful hypotheses about the parents and their observations of their child.

REVIEW OF COLLATERAL INFORMATION

Collateral information falls within two major categories. First, the evaluator will review relevant pleadings, declarations, and other court documents that the attorneys submit. Although these documents are not intended to bring “truth” to the case (even though declarations are signed under penalty of perjury), they do provide a framework from which to understand each parent’s perspectives and concerns. Evaluators must review all materials submitted, though the evaluator can set a deadline as to when materials must be submitted so the evaluation can be completed on time as required by the court.

The second type of collateral information comes from third parties who have relevant information about one or more family members. Collateral data can include information gathered from friends, relatives, babysitters, teachers, pediatricians, psychotherapists, and others. The child custody evaluator looks for convergent and divergent data between collateral and other data to help in understanding the various allegations and assertions made by the parties. Collateral information can be gathered verbally (over the telephone or in person) as well as in writing, with the use of questionnaires and letters, or by a review of affidavits or other written statements of the parties.

The benefits of gathering collateral information are listed next.

- Evaluators need to have a mind-set of disconfirmation rather than confirmation. Reviewing collateral information and talking with collateral sources allows for that.
- Parents in the midst of a custody dispute tend to present themselves in the most favorable light and the other parent more negatively. Collateral data can help balance this defensiveness and positive impression management by the parents.
- Collateral data may include information about parents and/or children that cannot be obtained through clinical interview, testing, and observation.
- Collateral data can help verify or refute claims made by the parents or others.

The AFCC Model Standards (2006) provide specific direction for the gathering of collateral data:

Evaluators shall be mindful of the importance of gathering information from multiple sources in order to thoroughly explore alternative hypotheses concerning issues pertinent to the evaluation. Evaluators shall recognize the importance of securing

information from collateral sources who, in the judgment of the evaluators, are likely to have access to salient and critical data. (Standard 11.1, p. 22)

When assessing the reports of participants in the evaluation, evaluators shall seek from other sources information that may serve either to confirm or disconfirm participant reports on any salient issue, unless doing so is not feasible. (Standard 11.2, p. 22)

In utilizing collateral sources, evaluators shall seek information that will facilitate the confirmation or disconfirmation of hypotheses under consideration. (Standard 11.4, p. 23)

All collateral sources contacted shall be disclosed by the child custody evaluator. (Standard 11.5, p. 23)

Austin (2002) recommended that child custody evaluators use a concentric circle approach for gathering collateral data. He identified that there are people who are close to the family (i.e., in the innermost of the concentric circles) who will have the richest and most complete information about family members. At the same time, these persons are likely to be the most biased and to take sides in the dispute. In the second, broader circle are professionals such as teachers, day care professionals, pediatricians, and therapists who may have considerable information about the family but only within the narrow areas of knowledge that they experience family members. Many of these persons may also be biased in favor of one parent. In the outermost circle are those few people who may know very little about the family but may have very specific information about an event or some particular parent-child information. This may include someone at the school who overhears an argument between a parent and the child's teacher. These persons may have very critical information about a specific event that was witnessed but know absolutely nothing else about any of the family members.

Typically, the court order appointing a child custody evaluator allows the evaluator to speak with any third-party collateral sources chosen, even without the expressed permission of either parent. Authorization from parents is required before speaking with professional collateral sources, such as teachers, therapists, and physicians. Evaluators must obtain the consent of the collateral witness to be part of the evaluation process and provide the same information about the limitations to confidentiality to all third-party collaterals, some of whom may not want to speak with an evaluator if they know that their comments are going to be included in a report to the court and read by the parents. As stated in the AFCC Model Standards, "Evaluators shall provide potential collateral informants with written information that shall include an unambiguous statement concerning the lack of confidentiality in a forensic mental health evaluation" (Standard 11.6, 2006, p. 23).

CRITICAL ISSUES IN REPORT WRITING

The various documents just cited all provide guidance on the critical issues involved in report writing. For example, California Rule of Court 5.220 states:

In any presentation of findings, the evaluator must:

- Summarize the data-gathering procedures, information sources, and time spent, and present all relevant information, including information that does not support the conclusions reached;
- Describe any limitations in the evaluation that result from unobtainable information, failure of a party to cooperate, or the circumstances of particular interviews;
- Only make a custody or visitation recommendation for a party who has been evaluated. This requirement does not preclude the evaluator from making an interim recommendation that is in the best interest of the child; and
- Provide clear, detailed recommendations that are consistent with the health, safety, welfare, and best interest of the child if making any recommendations to the court regarding a parenting plan. (p. 4)

Although this Rule is mandatory only for California child custody evaluators, these suggestions are useful for evaluators in any jurisdiction.

The Specialty Guidelines state:

Consistent with relevant law and rules of evidence, when providing professional reports and other sworn statements or testimony, forensic practitioners strive to offer a complete statement of all relevant opinions that they formed within the scope of their work on the case, the basis and reasoning underlying the opinions, the salient data or other information that was considered in forming the opinions, and an indication of any additional evidence that may be used in support of the opinions to be offered. The specific substance of forensic reports is determined by the type of psycholegal issue at hand as well as relevant laws or rules in the jurisdiction in which the work is completed. (APA, 2013, p. 17)

The Specialty Guidelines also instruct the forensic evaluator to disclose data and information that is not supportive of or contrary to the conclusions and recommendations offered by the evaluator.

Every report should have six complete sections, as discussed next:

1. Procedures
2. Each parent
3. Children
4. Collateral information
5. Analysis
6. Recommendations

Procedures. Based on the need to demonstrate balance and neutrality and to help the court understand that the evaluation process was thorough, the evaluator must explain all procedures and describe the length of appointments and the place where interviews and observations occurred. Materials reviewed must be described, and all collaterals who were interviewed must be listed. If certain procedures could not be completed, such as a stepparent who refused to be interviewed or a collateral witness who would not return phone calls, this should be described. If there was some reason that procedures were not balanced, this discrepancy should be explained. This thorough description of procedures helps reduce the risk that someone might perceive the evaluation and report as biased.

Each Parent. The evaluator must provide complete and relevant information about each parent. Among other things, the data included in this section of the report should include each parent's:

- Concerns and allegations.
- Responses to the other parent's stated significant concerns.
- Description of the children, the children's functioning, and needs.
- Relevant history and psychological functioning.
- Parenting style and strengths and weaknesses.
- Description of the coparenting relationship and each parent's contribution to the coparenting difficulties.
- Desired custodial outcome and reasons for that.

In addition, this section should include each parent's relevant details about any special issues. For example, in a relocation evaluation, it will be important for the evaluator to explain each parent's motives for moving or opposing the move, thoughts about how the move might be positive or harmful for the child, and other important relocation-related data.

Children. The report should provide thorough and relevant information about each child. The examiner must keep in mind that this is likely to be the only opportunity the court will have to gain a truly objective perspective of the children and their adjustment. Among other things, the data should include information about each child's:

- Developmental, social, psychological, academic, and social functioning, including interests, friendships, temperament, and typical mood.
- Relationship history with each parent.
- Thoughts about each parent.
- Feelings about a range of things, including the parents' divorce and their behaviors as divorced parents.

- Exposure to parental conflicts, and/or the extent to which the child feels alienated or justifiably estranged from one parent.
- Perspective of each parent's caretaking and how each parent disciplines the child.
- Opinion(s) about the parenting plan, if expressed.

Collateral Information. The evaluator must present information obtained from collateral sources and indicate whether this information was obtained in writing or verbally. Data about why the collateral informant's information was included in the data set, the nature and type of relationship the collateral informant has with the parties and/or children, and, if relevant, the attitude of the collateral informant about participating in the evaluation should be provided. A common complaint of parents is that the evaluator misrepresented what collateral informants offered. For this reason, it may be useful to review with the collateral informant the information to be included in the report and/or include a statement from each collateral witness confirming the information in the report.

Analysis. The analysis section is the most important component of the report. Rather than a review of information already described, the analysis section should focus on those data that lead to the expert opinions. In the analysis section, it is also important to discuss data that may not be consistent with the examiner's expert opinions or recommendations. Just as in math class, here it is important to show your work and explain the bases for all conclusions. It is important to detail the basis for any expert opinions reached. The analysis section should reflect that the evaluator considered each parent's concerns and responses to the other parent's concerns. It is important that those data are integrated with the psycholegal issues of concern to the court.

Given that, in most evaluations, there is a range of custodial options, it is important for the evaluator to provide a thorough risk-benefit analysis of each custodial option and those data that support his or her conclusions. For example, in a typical case, the evaluator should explain the risks and benefits of primary-mother custody, primary-father custody, and shared custody and the risks and benefits associated with substantial versus equal time with each parent (if shared). In a relocation case, the examiner should explain the risks and benefits of primary-mother custody in location X as opposed to the risks and benefits of primary-father custody in location Y. Finally, in all cases, the evaluator should explain the risks and benefits of shared decision making as opposed to some other plan that may give one parent decisions in certain areas of the child's life or perhaps even utilizing a parenting coordinator. In some cases, it might be best to provide the court with detailed parenting plans reflecting different options rather than a single recommended parenting plan. In such a case, the evaluator should detail the risks and benefits of each potential parenting plan in the report. Finally, and most important, it is critical to present

both the data that support the conclusions as well as the data that do not support the conclusions (as described in California Rule of Court 5.220 earlier in the chapter).

Recommendations. In recent years, there has been a renewed debate about whether examiners should make recommendations about the ultimate issue in child custody cases (*Family Court Review*, 2005; Stahl, 2005; Tippins & Wittmann, 2005). Judges typically prefer recommendations, and therefore it remains the custom of evaluators to provide them (M. J. Ackerman, Ackerman, Steffen, & Kelley-Poulos, 2004). Nevertheless, it is clearly the judge's job to make orders based on all of the evidence at trial rather than simply rubber-stamping the recommendations of a child custody evaluator (Schepard, 2004). Family law judges use the evaluator's recommendations as a starting place, not an end point, and they assess the usefulness of the evaluator's recommendations based on the consistency with other evidence presented at trial as well as the forensic integrity and quality of the evaluator's work product. Stahl (2005) suggested that, when one or more best interests or protective factors would suggest in favor of the child's relocation and one or more best interests or risk factors would suggest against the child's relocation, the evaluator should not weigh the various factors; this is the judge's job. Instead, in such situations, the evaluator should provide those conclusions to the court and provide multiple recommendations, with the ultimate decision based on the judge's weighting of the various best interests, risk, and protective factors.

In addition to the ultimate issue of parenting time, it is common in child custody evaluations to make recommendations in these areas:

- Legal custody and/or decision making.
- Interventions, including counseling for either parent and/or the children, identifying the suggested goals for that counseling.
- Substance abuse or domestic violence related interventions, if relevant.
- Alternative dispute resolution for ongoing issues (e.g., mediation or parenting coordinator).
- Any other recommendations relevant to the family that was evaluated.

CONCLUSIONS

Child custody evaluations are complex and require integrating disparate information gathered from a variety of participants and information of various types with disparate characteristics. Like a jigsaw puzzle, child custody evaluations require a persistent attitude of gathering more information, not only to confirm but also to disconfirm various hypotheses, until things fall into place. Evaluators need to avoid acting like stenographers and maintain a style of curiosity, always gathering additional relevant information until complex issues are understood. To be an effective evaluator, one must develop a thick skin, because one or both parents are likely to be upset with the recommendations. In some jurisdictions, it is not

uncommon for parents to file licensing complaints alleging bias or unprofessional behavior regardless of how thorough and professional the evaluation is. Indeed, the child custody evaluator carries more risk for licensing complaints than any other role played by the professional psychologist (Kirkland & Kirkland, 2001). Finally, if a case goes to trial, it is possible that one or both attorneys might hire a consultant or testifying expert who might criticize some of the work.

In spite of the risks and difficulties, conducting child custody evaluations can be professionally rewarding and satisfying. A child custody evaluator provides a beneficial service for family law judges that can help them understand the complexities of the most conflicted families they serve. Child custody evaluations can provide guidance for families who can settle their dispute and move forward following a well-done evaluation. Because child custody evaluators work in an interdisciplinary field, opportunities for ongoing learning and professional development are ever present. Most important, child custody evaluators keep the focus on the best interests of the children and therefore help them and their parents to adjust to the change in their lives as they develop healthier and more adaptive ways of moving forward.

REFERENCES

- Abidin, R. *Parenting stress index* (4th ed.). (2012). Lutz, FL: Psychological Assessment Resources.
- Achenbach, T. M. (1991). *Child Behavior Checklist*. Burlington: University of Vermont, Department of Psychiatry.
- Ackerman, M. J. (2006). *Clinician's guide to child custody evaluations* (3rd ed.). Hoboken, NJ: Wiley.
- Ackerman, M. J., & Ackerman, M. C. (1997). Custody evaluation practices: A survey of experienced professionals (revisited). *Professional Psychology: Research & Practice*, 28(2), 137–145.
- Ackerman, M. J., Ackerman, M. C., Steffen, L. J., & Kelley-Poulos, S. (2004). Psychologists' practices compared to the expectations of family law judges and attorneys in child custody cases. *Journal of Child Custody*, 1(1), 41–60.
- Ackerman, M. J., & Schoendorf, K. (1992). ASPECT: Ackerman-Schoendorf Scales for Parent Evaluation of Custody—manual. Los Angeles, CA: Western Psychological Services.
- Ahrons, C. (2005). *We're still family: What grown children have to say about their parents' divorce*. New York, NY: HarperCollins.
- American Academy of Child and Adolescent Psychiatry. (1997). Practice parameters for child custody evaluation. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(10 Suppl.), 57S–68S.
- American Law Institute. (2002). *Principles of the law of family dissolution: Analysis and recommendations*. Newark, NJ: LexisNexis.
- American Psychological Association. (2002). *Ethical principles for psychologists*. Washington, DC: Author.
- American Psychological Association. (2009). *Guidelines for child custody evaluations in family law proceedings*. Washington, DC: Author.

- American Psychological Association. (2013). Specialty guidelines for forensic psychology. *American Psychologist*, 68(1), 7–19. doi:10.1037/a0029889
- Arizona Revised Statute 25-403. Best Interests of the Child.
- Association of Family and Conciliation Courts. (2006). *Model standards of practice for child custody evaluation*. Madison, WI: Author.
- Association of Family and Conciliation Courts. (2009). *Guidelines for brief focused assessment*. Madison, WI: Author.
- Austin, W. (2000). A forensic psychology model of risk assessment for child custody relocation law. *Family Court Review*, 38(2), 192–207.
- Austin, W. (2001). Partner violence and risk assessment in child custody evaluations. *Family Court Review*, 39(4), 483–496.
- Austin, W. (2002). Guidelines for utilizing collateral sources of information in child custody evaluations. *Family Court Review*, 40(2), 177–184.
- Austin, W. (2008a). Relocation, research, and forensic evaluation, Part I: Effects of residential mobility on children of divorce. *Family Court Review*, 46(1), 137–150.
- Austin, W. (2008b). Relocation, research, and forensic evaluation, Part II: Research in support of the relocation risk assessment model. *Family Court Review*, 46(2), 347–365.
- Australia Family Law Act. (1975), including the amendments made by the “Family Law Amendment (Shared Parental Responsibility) Act 2006.”
- Bathurst, K., Gottfried, A. W., & Gottfried, A. E. (1997). Normative data for the MMPI-2 in child custody litigation. *Psychological Assessment*, 9(3), 205–211.
- Bow, J. N., & Quinnell, F. A. (2001). Psychologists’ current practices and procedures in child custody evaluations: Five years after the American Psychological Association guidelines. *Professional Psychology: Research and Practice*, 32(3), 261–268.
- Braver, S., Ellman, I., & Fabricius, B. (2003). Relocation of children after divorce and children’s best interests: New evidence and legal considerations. *Journal of Family Psychology*, 17(2), 206–219.
- Bricklin, B. (1989). *Perception of Relationships Test manual*. Furlong, PA: Village Publishing.
- Bricklin, B. (1990a). *Bricklin Perceptual Scales manual*. Furlong, PA: Village Publishing.
- Bricklin, B. (1990b). *Parent Awareness Skills Survey manual*. Furlong, PA: Village Publishing.
- Brodzinsky, D. (1993). On the use and misuse of psychological testing in child custody evaluations. *Professional Psychology: Research and Practice*, 24(2), 213–219.
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *The Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis: University of Minnesota Press.
- California Rules of Court, Rule 5.220.
- California Rules of Court, Rule 5.225.
- California Rules of Court, Rule 5.230.
- Colorado Revised Statutes, § 14–10–124. Best interests of child.
- Craig, R. J. (2006). The Millon Clinical Multiaxial Inventory-III. In R. P. Archer (Ed.), *Forensic uses of clinical assessment instruments* (pp. 121–145). Mahwah, NJ: Erlbaum.
- Dalton, C., Drozd, L., & Wong, F. (2005). *Navigating custody and visitation evaluations in cases with domestic violence: A judge’s guide*. Reno, NV: National Council of Juvenile and Family Court Judges.
- Drozd, L., & Olesen, N. (2004). Is it abuse, alienation, and/or estrangement? A decision tree. *Journal of Child Custody*, 1(3), 65–106.

- Dyer, F. J. (2008). Using the Millon inventories in forensic psychology. In T. Millon & C. Bloom (Eds.), *The Millon Inventories: A practitioner's guide to personalized psychological assessment* (2nd ed., pp. 117–195). New York, NY: Guilford Press.
- Eddy, W. (2006). *High conflict people in legal disputes*. Santa Anna, CA: Janis Publications.
- Emery, R. E., Otto, R. K., and O'Donohue, W. T. (2005). A critical assessment of child custody evaluations: Limited science and a flawed system. *Psychological Science in the Public Interest*, 6(1), 1–29.
- Erard, R. E. (2005). What the Rorschach can contribute to child custody and parenting evaluations. *Journal of Child Custody*, 2, 119–142.
- Evans, F. B., & Schutz, B. M. (2008). The Rorschach in child custody and parenting plan evaluations: A new conceptualization. In C. B. Gacono & F. B. Evans (Eds.), *Handbook of forensic Rorschach assessment* (pp. 233–254). New York, NY: Routledge.
- Fabricius, W., Sokol, K., Diaz, P., & Braver, S. (2012). Parenting time, parent conflict, parent-child relationships, and children's physical health. In K. Kuehnle & L. Drozd (Eds.), *Parenting plan evaluations: Applied research for the family court* (pp. 188–213). New York, NY: Oxford University Press.
- Family Court Review*. (2005). Symposium issue, *Child custody evaluations*, 43(2).
- Family Court Review*. (2008). Special issue, *Domestic violence*, 46(3).
- Family Court Review*. (2010). Special issue, *Alienated children in divorce and separation: Emerging approaches for families and courts*, 48(1).
- Family Court Review*. (2011). Special issue, *Attachment theory, separation, and divorce: Forging coherent understandings for family law*, 49(3).
- Flens, J. (2005). The responsible use of psychological testing in child custody evaluations: Selection of tests. *Journal of Child Custody*, 2(1–2), 3–29.
- Garbolino, J. (2001). *Judging in the gray areas*. Presentation at the National Judicial College, Reno, Nevada.
- Gardner, R. (1987). *The parental alienation syndrome and the differentiation between fabricated and genuine child sex abuse*. Cresskill, NJ: Creative Therapeutics.
- George, C., Solomon, J., & McIntosh, J. (2011). Divorce in the nursery: On infants and overnight care. *Family Court Review*, 49(3), 521–528.
- Gerard, A. (1994). *Parent–Child Relationship Inventory*. Los Angeles, CA: Western Psychological Services.
- Gould, J. (1998). *Conducting scientifically crafted child custody evaluations*. Thousand Oaks, CA: Sage.
- Gould, J., & Martindale, D. (2007). *The art and science of child custody evaluations*. New York, NY: Guilford Press.
- Gould, J., & Stahl, P. (2001). Never paint by the numbers: A response to Kelly and Lamb (2000), Solomon and Biringen (2001), and Lamb and Kelly (2001). *Family Court Review*, 39(4), 372–376.
- Greenberg, L. R., Doi Fick, L., & Schnider, R. (2012). Keeping the developmental frame: Child-centered conjoint therapy. *Journal of Child Custody*, 9(1–2), 39–68.
- Greenberg, L. R., Gould, J. W., Gould-Saltman, D. J., & Stahl, P. (2003). Is the child's therapist part of the problem? What judges, attorneys and mental health professionals need to know about court-related treatment for children. *Family Law Quarterly*, 37, 241–271.
- Hetherington, E., & Kelly, J. (2002). *For better or worse: Divorce reconsidered*. New York, NY: Norton.

- Hodges, W. F., (1991). *Interventions for children of divorce: Custody, access, and psychotherapy* (2nd ed.). Oxford, England: Wiley.
- In re Marriage of LaMusga, 88 P. 3d 81—Cal. Supreme Court 2004.
- Jaffe, P. G., Johnston, J. R., Crooks, C. V., & Bala, N. (2008). Custody disputes involving allegations of domestic violence: The need for differentiated approaches to parenting plans. *Family Court Review*, 46, 500–522.
- Kelly, J. B. (2012). Risk and protective factors associated with child and adolescent adjustment following separation and divorce: Social science applications. In K. Kuehnle & L. Drozd (Eds.), *Parenting plan evaluations: Applied research for the family court* (pp. 49–84). New York, NY: Oxford University Press.
- Kelly, J. B., & Emery, R. (2003). Children’s adjustment following divorce: Risk and resilience perspectives. *Family Relations: Interdisciplinary Journal of Applied Family Studies*, 52(4), 352–362.
- Kelly, J. B., & Johnson, M. P. (2008). Differentiation among types of intimate partner violence: Research update and implications for interventions. *Family Court Review*, 46(3), 476–499.
- Kelly, J. B., & Johnston, J. R. (2001). The alienated child: A reformulation of parental alienation syndrome. *Family Court Review*, 39(3), 249–266.
- Kelly, J. D., & Lamb, M. E. (2000). Using child development research to make appropriate custody and access decisions for young children. *Family Court Review*, 38(3), 297–311.
- Kelly, R. K., & Ramsey, S. H. (2009). Child custody evaluations: The need for systems-level outcome assessments. *Family Court Review*, 47(2), 286–303.
- Kirkland, K., and Kirkland, K. L. (2001). Frequency of child custody evaluation complaints and related disciplinary action: A survey of the Association of State and Provincial Psychology Boards. *Professional Psychology: Research and Practice*, 32(2), 171–174.
- Kuehnle, K., & Connell, M. (Eds.). (2009). *The evaluation of child sexual abuse allegations: A comprehensive guide to assessment and testimony*. Hoboken, NJ: Wiley.
- Kuehnle, K., & Drozd, L. (Eds.). (2012). *Parenting plan evaluations: Applied research for the family court*. New York, NY: Oxford University Press.
- Lamb, M. E., Sternberg, K., & Thompson, R. (1997). The effects of divorce and custody arrangements on children’s behavior, development, and adjustment. *Family & Conciliation Courts Review*, 35(4), 393–404.
- Lamb, M. E., Hershkowitz, I., Orbach, Y., and Esplin, P. (2008). *Tell me what happened: Structured investigative interviews of child victims and witnesses*. Hoboken, NJ: Wiley.
- Lewis, T. (2010). *Exploring the Best Interest of the Child standard*. Paper presented at continuing legal education program of the Rutter Group, Los Angeles, CA.
- Ludolph, P., & Dale, M., (2012). Attachment in child custody: An additive factor, not a determinative one. *Family Law Quarterly*, 46(1), 1–40.
- McCann, J., Flens, J., Campagna, V., Collman, P., Lazzaro, T., & Connor, E. (2001). The MCMI-III in child custody evaluations: A normative study. *Journal of Forensic Psychology Practice*, 1(2), 27–44.
- Maccoby, E., & Mnookin, R. (1992). *Dividing the child: Social and legal dilemmas of custody*. Cambridge, MA: Harvard University Press.
- McIntosh, J., & Smythe, B. (2012). Shared-time parenting: An evidence-based matrix for evaluating risk. In K. Kuehnle & L. Drozd (Eds.), *Parenting plan evaluations: Applied research for the family court* (pp. 155–187). New York, NY: Oxford University Press.

- Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C. (2007). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers*. New York, NY: Guilford Press.
- Millon, T., Millon, C., Davis, R., & Grossman, S. (2009). *MCMI-III manual* (4th ed.). Minneapolis, MN: Pearson Education.
- Morey, L. C. (2007). *The Personality Assessment Inventory professional manual*. Lutz, FL: Psychological Assessment Resources.
- Otto, R. K., Edens, J. F., & Barcus, E. (2000). The use of psychological testing in child custody evaluations. *Family and Conciliation Courts Review*, 38, 312–340.
- Parkinson, P., Cashmore, J., & Single J. (2010). The need for reality testing in relocation disputes. *Family Law Quarterly*, 44, 1.
- Rorschach, H. (1998). *Psychodiagnosics: A diagnostic test based on perception* (10th ed.). Cambridge, MA: Hogrefe (Original work published 1921).
- Schepard, A. (2004). *Children, courts, & custody*. Cambridge, England: Cambridge University Press.
- Solomon, J., & Biringen, Z. (2001). Another look at the developmental research: Commentary on J. B. Kelly and Lamb's "Using child development research to make appropriate custody and access decisions for young children." *Family Court Review*, 39, 355–364.
- Sparta, S. N., & Koocher, G. P. (Eds.). (2006). *Forensic mental health assessment of children and adolescents*. New York, NY: Oxford University Press.
- Stahl, P. (1994). *Conducting child custody evaluations: A comprehensive guide*. Thousand Oaks, CA: Sage.
- Stahl, P. (1999). *Complex issues in child custody evaluations*. Thousand Oaks, CA: Sage.
- Stahl, P. (2004). Understanding and evaluating alienation in high-conflict custody cases. *Wisconsin Journal of Family Law*, 24, 1.
- Stahl, P. (2005, April). The benefits and risks of child custody evaluators making recommendations to the court: A response to Tippins and Wittmann. *Family Court Review*, 43(2), 260–265.
- Stahl, P. (2006). Avoiding bias in relocation evaluations. *Journal of Child Custody*, 3(3–4), 109–124.
- Stahl, P. (2010). *Conducting child custody evaluations: From basic to complex issues*. Thousand Oaks, CA: Sage.
- Strachan, A. M., Lund, M. E., & Garcia, J. A. (2010). Assessing children's perceptions of family relationships: An interactive instrument for use in custody disputes. *Journal of Child Custody*, 7(3), 192–216.
- Sullivan, M. (2004). Ethical, legal, and professional practice issues involved in acting as a psychologist parent coordinator in child custody cases. *Family Court Review*, 42(3), 576–582.
- Taylor, N. J., Gollop, M., & Henaghan, R. M. (2010). *Relocation following parental separation: The welfare and best interests of children*. Research report for the New Zealand Law Foundation. Dunedin, NZ: University of Otago Centre for Research on Children and Families and Faculty of Law.
- Tippins, T., & Wittmann, J. (2005). Empirical and ethical problems with custody recommendations: A call for clinical humility and judicial vigilance. *Family Court Review*, 43(2), 193–222.

- Wallerstein, J. S., & Kelly, J. B. (1980). *Surviving the breakup: How children and parents cope with divorce*. New York, NY: Basic Books.
- Wallerstein, J. S., & Tanke, T. (1996). To move or not to move: Psychological and legal considerations in the relocation of children following divorce. *Family Law Quarterly*, 30, 305.
- Warshak, R. A. (2001). Current controversies regarding parental alienation syndrome. *American Journal of Forensic Psychology*, 19(3), 29–59.
- Weiner, I. B., & Hess, A. K. (Eds.). (2006). *The handbook of forensic psychology* (3rd ed.). Hoboken, NJ: Wiley.

CHAPTER 7

Conducting Personal Injury Evaluations

LISA DRAGO PIECHOWSKI

UNLIKE criminal proceedings in which the government charges an individual with having committed an unlawful act, civil proceedings involve disputes between private parties (Melton et al., 2007). Many such cases involve allegations that one party (the plaintiff) has been injured by another party (the defendant) and seeks compensation for the damage that was done. Such actions are known as tort or personal injury cases, which are the subject of this chapter.

Although many personal injury cases involve claims of physical injuries or financial losses, allegations of “psychological injury” or “emotional distress” can also be elements of a tort action. Koch, Douglas, Nicholls, and O’Neill (2006) defined psychological injuries as “stress-related emotional conditions resulting from real or imagined threats or injuries” caused by a third party (p. 3). In cases involving claims of emotional distress, psychologists may be retained as experts, most often to evaluate the plaintiff and offer an opinion on the nature and extent of the psychological damage incurred by the plaintiff and the causal relationship (if any) between these injuries and the actions of the defendant. This chapter discusses the legal, ethical, and practical aspects of performing forensic psychological evaluations in personal injury cases.

LEGAL CONTEXT OF PERSONAL INJURY EVALUATIONS

It is important that psychologists undertaking evaluations in personal injury cases have knowledge of the legal parameters of civil litigation. This includes an understanding of tort law, the civil procedure, and the rules governing expert testimony.

TORT LAW

Tort law forms the legal framework for personal injury cases (Greenberg, 2003). A “tort” is a civil wrong, but it is not necessarily an illegal act. Sexual assault, for

example, is both a crime and a possible basis for a tort action. In contrast, a person who slips on ice outside a neighbor's front door and breaks a leg may be able to sue the neighbor for the harm suffered. It is unlikely, however, that the neighbor would be charged for having ice on the doorstep. Melton et al. (2007) point out that, in addition, torts are not directly related to moral conduct insofar as an act may be immoral but not tortious, and provide the example of a bystander failing to save a drowning child. Although this may be morally reprehensible, the person could not be sued for failure to act.

In *A Concise Restatement of Torts*, the American Law Institute (ALI, 2000) defines tortious conduct as including both acts of commission and acts of omission. For an act to be tortious, it must meet four criteria: duty, dereliction, damage, and causality. *Duty* refers to an obligation by the defendant to the plaintiff. Consider a psychologist's obligation to keep confidential a patient's disclosures. If the psychologist tells someone else what the patient confided, the psychologist would not be living up to this duty. If a psychologist is having coffee with a friend, however, and the friend discloses she is having an affair, the psychologist has no legal duty to keep that information confidential.

The second required element is a *dereliction* or breach. Dereliction occurs when the tortfeasor (wrongdoer) fails to fulfill a duty. For example, the owner of a day care center fails to properly screen potential employees and hires an individual who has a history of sexually abusing children, or a doctor fails to review a patient's chart and prescribes a medication to which the patient is allergic. In both of these examples, the duty owed (to the children at the day care center, to the doctor's patient) was not fulfilled.

Damage, which can include monetary losses, physical injuries, or emotional distress, is the third element of a tort. A tort action requires that, in addition to the dereliction of duty, some damage must have occurred. Suppose, for example, that when the patient just described attempted to get the prescription filled at the pharmacy, the pharmacist recognized that the patient had an allergy to the medication and called the physician, who then prescribed an alternative medication. Despite the physician's dereliction of duty, the patient suffered no damage. The patient, therefore, would have no cause for a tort action.

The final element of a tort is *causality*, which requires that the damage incurred be directly related to the tortfeasor's dereliction of duty. This concept is also known as proximate cause. The actions (or inactions) of the tortfeasor must be a primary reason or a "substantial factor" in bringing about the harm (ALI, 2000). In the previous example, suppose that, due to the physician's error and the need to obtain an alternative medication, the patient was delayed for an extra 30 minutes at the pharmacy. Upon leaving the pharmacy, the patient was struck by a car operated by a drunk driver and suffered severe physical injuries. Despite the physician's dereliction of duty in prescribing the wrong medication, it is more likely that the actions of the drunk driver and not those of the physician would be considered the proximate cause of the patient's injuries.

A tort action seeks to make the plaintiff “whole”—that is, to return the plaintiff to the state he or she was in prior to the defendant’s wrongful actions. If the defendant is judged to be responsible for the injury suffered by the plaintiff, the plaintiff will be awarded damages in the form of monetary compensation. The amount of the award is determined by the degree of damage suffered by the plaintiff and the behavior of the defendant that resulted in the plaintiff’s injury.

Tortious acts can be intentional or negligent. *Intent* means that the actor desired to cause the consequence of the act (ALI, 2000). Sexual assault is an example of an intentional tort. A woman who sees someone lurking in her yard and points a gun at that person, in order to frighten him so that he will leave her yard, is acting with intent when the person runs away, because inducing fear was the intention of her act. If, however, the person runs into the street in front of a moving car and is killed, his death was not the intent of the actor, despite the ultimate outcome.

Negligence is defined as “a departure from a standard of conduct demanded by the community for the protection of others against unreasonable risk” (ALI, 2000, p. 33). Both acts of commission and acts of omission can be negligent. It might be argued that pointing a gun at someone would constitute an unreasonable risk. Failing to act can also be negligent. For example, a preschool teacher who failed to notice that a pupil had wandered out of the school might be considered negligent.

PROCESS OF CIVIL LITIGATION

Both federal and state courts have rules governing the process of civil litigation. In the federal court system, these rules are known as the Federal Rules of Civil Procedure (FRCP). All states and the District of Columbia have similar sets of rules governing civil procedure in their jurisdictions. (See DeMatteo, Kessler, & Strohmaier, Chapter 3 this volume, for information on how to find laws in various jurisdictions.) Greenberg (2003) noted that, despite variations, most jurisdictions follow a similar pattern, which consists of several stages, including pleadings, discovery, trial, and judgment. At any stage of this process prior to a judgment being rendered, all or part of the case may be dismissed or the parties may agree to a settlement, thus ending the proceedings.

The civil action begins when the plaintiff files a legal pleading known as the complaint, which outlines the grounds for the action and the events giving rise to the dispute. The complaint identifies the defendant and describes why the plaintiff believes he or she is entitled to damages. The defendant is then notified of the pending litigation and provides a response, or answer. In addition, the defendant may take other actions, including filing a motion to dismiss, filing a counterclaim whereby the defendant makes a claim against the plaintiff, or filing a notice of an affirmative defense that seeks to negate the plaintiff’s right to recover.

The second phase of personal injury litigation is known as discovery. This phase allows the parties to obtain evidence and can be accomplished in a number of ways, including requests for the production of documents, the use of interrogatories

(a series of written questions to which the opposing party must respond), depositions (sworn testimony of a party or witness taken before trial), and physical or mental examinations of the plaintiff. This is the point at which forensic psychologists typically enter the process. If the plaintiff alleges that he or she incurred psychological damage or suffers from emotional distress as a result of the defendant's actions, the defendant has the right to independently assess the nature and degree of psychological damage the plaintiff alleges. This is authorized by FRCP Rule 35 (or its equivalent in a state court system). The plaintiff's attorney may also choose to have the plaintiff evaluated as a way of providing further evidence supporting the plaintiff's claimed injuries or to dispute the findings of the defendant's evaluation.

Depending on the outcome of the evaluation, one party may choose to present the psychologist's report and testimony as evidence at trial. Before this can be done, the psychologist must be "disclosed," or identified as an expert witness, and the opinions the psychologist intends to testify to must be described in what is called an expert disclosure. This document identifies the following:

- The psychologist
- The psychologist's qualifications including a list of publications
- The subject matter of the psychologist's testimony
- The facts and opinions to which the psychologist is expected testify
- Grounds for the psychologist's opinions
- Materials the psychologist relied on in forming these opinions
- A list of all cases in which the psychologist has testified in the previous 4 years
- Compensation the psychologist is receiving

The psychologist also provides a written report describing the findings from the examination and the opinions he or she reached. Following this disclosure, the psychologist is likely to be "deposed" (questioned under oath) by the opposing attorney and will be required to provide the opposing attorney with documents, including test data, notes, and correspondence related to the evaluation.

The next phase of litigation is the trial. In civil litigation, the plaintiff has the burden of proof. This means the plaintiff is required to prove that it is more likely than not that the defendant's actions resulted in damage to the plaintiff. This "standard of proof" is known as preponderance of the evidence. The plaintiff's case is presented first. The attorney for the plaintiff will call and directly examine witnesses. In direct examination, the attorney asks nonleading, open-ended questions, such as "Doctor, can you describe your examination of the plaintiff?" rather than "Doctor, isn't it true that the plaintiff has posttraumatic stress disorder (PTSD)?" The purpose of direct examination is to provide the witness the opportunity to give evidence that will support the plaintiff's case. Following direct examination of each witness, the defendant's attorney will cross-examine the witness. Cross-examination is intended to raise questions about the reliability of the witness's evidence and/or to expose weaknesses or contradictions in the witness's testimony. During cross-examination, leading questions may be asked, such as "Doctor, isn't it true that no scale on the

Minnesota Multiphasic Personality Inventory–2 (MMPI-2) can prove that someone is responding truthfully on the test?” After the plaintiff’s case is concluded, the defense presents its case. Witnesses will be called and directly examined, followed by cross-examination by the plaintiff’s attorney.

After both the plaintiff’s and the defendant’s arguments have been presented, the judgment phase begins. The trier of fact (the jury or judge) will consider the credibility of the evidence and render a verdict. In civil litigation, there are three considerations: (1) Did the defendant breach a duty that was owed to the plaintiff? (2) Was the plaintiff injured by the actions of the defendant? (3) If so, what damages (financial compensation) should be awarded to the plaintiff? Either party may choose to appeal the verdict.

EXPERT WITNESS IN CIVIL LITIGATION

In the federal court system, the meaning of the term *expert witness* is defined in Rule 702 of the Federal Rules of Evidence (FRE) as a witness who is “qualified as an expert by knowledge, skill, experience, training, or education” and can assist the trier of fact to “understand the evidence or to determine a fact in issue” using “scientific, technical or other specialized knowledge.” Expert witnesses, unlike lay witnesses, may testify in the form of opinions and inferences if the testimony is (a) based on sufficient facts and data; (b) the product of reliable methods and principles; and (c) if the underlying methods and principles have been applied reliably to the facts of the case (FRE 702, 703).

In personal injury cases, a psychologist may be retained as an expert by either the plaintiff or the defendant. The plaintiff can elect to be examined by an expert in order to provide evidence of emotional damage or in an attempt to dispute the findings of the defendant’s expert. If the findings of the examination do not appear to be helpful to the plaintiff’s case, it is likely that the plaintiff will not request a written report from the expert or disclose the expert as a potential witness. If the examiner’s opinion appears to support the plaintiff’s position, however, he or she will be asked to write a report and will be disclosed as an expert.

As noted previously, Rule 35 of the FRCP provides that the defendant can force the plaintiff to undergo an examination when the plaintiff claims psychological injury as a basis for receiving damages. In such cases, the plaintiff is entitled to receive a written report describing the examiner’s findings. This report must be detailed and disclose all the procedures used, the opinions reached, and the foundations for each opinion. In return, the plaintiff is required to provide the defendant any reports under the plaintiff’s control addressing these same issues.

ETHICAL ISSUES IN PERSONAL INJURY WORK

When psychologists are engaged in forensic work, they have increased exposure to legal risks (Knap & VandeCreek, 2012). The parties involved in the legal system

(attorneys, judges, clients, etc.) have no obligation to safeguard the professional or legal interests of the experts participating in the process. Consequently, psychologists may face ethical challenges engendered by the adversarial nature and the differing professional responsibilities, agendas, and goals of those working in the legal system (Shuman & Greenberg, 1998). The Ethical Principles of Psychologists and Code of Conduct (EPPCC; American Psychological Association [APA], 2010) contains both aspirational goals (General Principles) and enforceable rules of conduct (Ethical Standards) for psychologists. Because the ethical standards are written broadly so as to be applicable to psychologists working in a range of professional roles and circumstances, psychologists doing forensic work may benefit from the additional guidance provided by the Specialty Guidelines for Forensic Psychology (SGFP; APA, 2013; Guidelines are reprinted as the appendix to this volume with permission of the APA). Unlike the Ethical Standards of the EPPCC, which mandate behavior, the SGFP are “aspirational” and were designed to “improve the quality of forensic psychological services; enhance the practice and facilitate the systematic development of forensic psychology; encourage a high level of quality in professional practice; and encourage forensic practitioners to acknowledge and respect the rights of those they serve” (p. 1). Specific issues that are likely to arise for psychologists performing personal injury evaluations are discussed next. A more detailed discussion of ethical considerations in forensic practice can be found in Weiner and Hess (Chapter 4 this volume).

OBJECTIVITY

In the midst of the adversarial legal system, ethics demand that psychologists remain objective, impartial, and fair. Psychologists have a duty to the court to offer “testimony that is reliable, helpful, honest and objective” (Kane, Nelson, Dvoskin, & Pitt, 2013, p. 151). This commitment to objectivity must be maintained regardless of which party retained the psychologist, how sympathetic or unsympathetic the plaintiff may be, or the nature of the relationship the psychologist has with the attorneys in the case. One way to facilitate objectivity is for the psychologist to establish consistent evaluation procedures that are utilized regardless of the specific circumstances or retaining party. By following a preestablished protocol, the psychologist is more likely to avoid bias in test selection, the scope or length of the interview, and decisions about including collateral sources. Objectivity is also enhanced by a methodical approach to data interpretation characterized by identifying and testing alternative hypotheses, acknowledging and considering inconsistencies in the data, and avoiding selectively attending to data that support the evaluator’s opinion while ignoring or minimizing conflicting data.

ROLES AND ROLE CONFLICT

Conducting an evaluation in a personal injury case requires that the psychologist understand his or her role in this process. Whether the psychologist is retained by

the plaintiff or the defendant, his or her role is to provide an accurate and objective assessment of the plaintiff. This role is distinctly different from the role of a treating psychologist and the role that the attorneys play.

As described by Greenberg and Shuman (1997), the role of the therapist is to demonstrate support, acceptance, and empathy toward the patient and seek to understand the patient's perceptions and feelings. The therapist's knowledge of the patient's world usually is based exclusively on the patient's self-report and is likely to be skewed by the patient's biases, assumptions, and attitudes. The therapist is less concerned about the objective truth of the patient's circumstances than the patient's subjective understanding of these circumstances. Because of these differences, psychologists should avoid acting as both treatment provider and objective evaluator for the same individual. This does not preclude a therapist from testifying as a fact witness (as opposed to an expert witness) about the patient's treatment, although consideration should be given as to how such testimony might affect the therapeutic relationship.

Attorneys are required to advocate vigorously for their clients and expect the experts they retain to do the same (Shuman & Greenberg, 1998). The attorney's job is to persuade the trier of fact to find in favor of the party the attorney represents. This creates a conflict for the forensic evaluator, who must balance the demands of the advocacy-based legal system with psychological ethics that require impartiality and not partisanship (APA, 2010, EPPCC 3.05, 3.06). The fact that the attorney is paying the expert and may hold out the promise of future work has the potential to further cloud these issues.

The roles a psychologist can play in a personal injury evaluation are not limited to fact witness and expert witness. A psychologist might be retained as a consultant by an attorney. Consultants do not evaluate the plaintiff or offer expert testimony. The role of the consultant is to provide expertise to the attorney in preparing and litigating the case. This might include reviewing and commenting on the plaintiff's treatment records, assisting with jury selection, reviewing a report summarizing the examination completed by the opposing party's expert, reviewing the depositions of the opposing party's experts, providing questions for cross-examination of opposing experts, and conducting research.

Professional ethics demand that psychologists avoid entering into relationships that could compromise their objectivity, competence, or effectiveness (APA, 2010, EPPCC 3.05, 3.06). Because of this requirement, it is important to consider the potential impact of influences such as the therapeutic bond, financial remuneration, and personal persuasion. Psychologists will find it easier to navigate these influences if they avoid multiple roles (therapist, examiner, consultant) and remain focused on reaching opinions based on an unbiased and objective view of the evidence. Once such opinions are reached, the psychologist can vigorously and persuasively advocate for his or her opinions, keeping in mind the distinction between advocating for an opinion and advocating for a party. As described by Kane et al. (2013), "The

expert must resist this pressure, remaining impartial and advocating for his or her opinion, not for his or her retaining attorney” (p. 151).

ADEQUATE BASIS FOR OPINIONS

Psychologists are required to base the opinions in reports and testimony on information and techniques sufficient to substantiate their findings (APA 2010, EPPCC 9.01). This is consistent with the requirements of the legal system that expert testimony meet certain standards, including that it be based on sufficient facts and data and reliable methods (FRE 702). This means that the evaluator must collect data that are comprehensive, relevant, and valid. In addition, data must be interpreted appropriately. The inferences made must be reasonable and based on the data that were collected. The SGFP stress the importance of utilizing multiple sources of data and attempting to corroborate data by comparing information across sources (APA, 2013). In circumstances when such cross-validation is not possible, resulting limitations should be identified. If the psychologist is not able to obtain sufficient data to reach an opinion on a given question, no opinion should be offered.

It is also important to focus on collecting data that are relevant to the legal question at hand. Relevant, in this context, is defined as “evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence” (FRE 401). Only relevant evidence can be admitted (FRE 402). This distinction requires that the psychologist clearly understand the facts that are at issue and how data can be collected to address those facts specifically. Collecting data that are not relevant wastes time and the financial resources of the retaining party. It also serves to confuse the issues and make it more difficult to reach and clearly explicate opinions.

PREVENTING MISUSE OF FINDINGS

Psychologists have an ethical obligation to ensure that their reports and testimony do not contain statements that are untrue or are likely to mislead others (APA, 2010, EPPCC 1.01, 5.01). Section 11.01 of the SGFP provides more detailed guidance: “forensic practitioners do not distort or withhold relevant evidence or opinion in reports or testimony.... Forensic practitioners do not, by either commission or omission, participate in the misrepresentation of their evidence, nor do they participate in partisan attempts to avoid, deny or subvert the presentation of evidence contrary to their own position or opinion” (APA, 2013, p. 16). As noted earlier, attorneys are obligated to vigorously advocate for the parties they represent. As such, attorneys attempt to present evidence in a light most favorable to their clients’ positions. Forensic psychologists must remain alert to attempts to present their findings in ways that may misleading the judge or jury by way of overemphasizing certain points or failing to acknowledge conflicting data. The forensic psychologist must also be sure that the retaining attorney does not misrepresent the psychologist’s

credentials, training, or experience. It is important that the psychologist carefully review the expert disclosure before it is submitted to be sure that it does not misrepresent his or her opinions or overstate the certainty with which those opinions are held.

NATURE OF PSYCHOLOGICAL DAMAGE

Courts have long allowed claims based on physical harm, and all jurisdictions now allow claims of emotional harm proximately related to physical injuries. However, courts have been less receptive to claims based solely on psychological damage absent a physical injury (Kane et al., 2013). According to Koch, O'Neill, and Douglas (2005), courts have long been skeptical of claims of psychological injury because of fears that, without objective markers for the existence of mental health conditions, such disorders would be easy to fabricate. The ALI defines *harm* in the context of tort law as "the existence of loss or detriment of any kind to a person resulting from any cause" (p. 2). By extrapolation, this definition suggests that psychological damage has two elements: (1) that there is a change in the plaintiff's emotional functioning, and (2) that this change can be attributed, at least in part, to the actions of another party.

EMOTIONAL FUNCTIONING

It is important to point out that the term *psychological damage* is not synonymous with psychiatric diagnosis. Greenberg, Shuman, and Meyer (2004) differentiated between the clinical issue of assigning a specific diagnostic label and using a diagnostic classification as the basis for a legal determinant: "[The] law does not make diagnosis an essential element of a claim or defense. Instead, legal criteria for these actions are functional and concern themselves with impairment or capacity, without regard to diagnosis" (p. 2). Furthermore, the *Diagnostic and Statistical Manual of Mental Disorders* cautions about its use in nonclinical settings. As described earlier in this chapter, the fit between clinical diagnostic data and legal questions is "imperfect" (American Psychiatric Association, 2000, p. xxxii).

In establishing the existence of psychological damage, functional capacity is more relevant than a diagnostic label. Grisso (2003) defined functional capacity as that which an individual can do or accomplish, as well as the knowledge, understanding, or beliefs that may be necessary for that accomplishment. Since functional capacity is distinct from diagnosis, it cannot be assumed that the presence of a particular condition is necessarily related to a specific level of functioning. Greenberg et al. (2004) warned that substituting diagnosis for an analysis of functioning can be misleading, does not serve the purposes of the court, and has the potential to distort an objective assessment of the plaintiff. They noted that tort cases require a functional analysis of the plaintiff in order to understand how, if at all, the defendant's actions have affected the plaintiff's life.

With the understanding that establishing a diagnosis is not required for a claim of psychological injury, there are some conditions that are commonly claimed by

plaintiffs in tort cases. Witt and Weitz (2007) reported that the most common symptoms seen following motor vehicle accidents include chronic pain, depression, and anxiety (including PTSD). Melton et al. (2007) reported that the condition variously known as traumatic neurosis/PTSD/acute stress disorder (ASD) is the most commonly observed mental injury in personal injury cases. Koch et al. (2005) listed PTSD, ASD, and major depressive episode as conditions that may be the subject of personal injury litigation. Kane et al. (2013) pointed to PTSD as the most common diagnosis in personal injury cases.

PTSD and ASD occupy a unique position in the realm of psychiatric diagnoses, in that these conditions are by definition caused by an external event or circumstance. Other conditions, such as depression and panic disorder, can be triggered by a traumatic experience, such as the death of a loved one, but they often arise in the absence of exogenous factors. In terms of personal injury evaluations, forensic experts are more likely to be presented with a diagnosis of PTSD than ASD, as the symptoms of ASD, according to the diagnostic criteria, resolve within a four-week period. If the symptoms persist beyond this period, a diagnosis of PTSD would be considered.

The type of trauma necessary for a diagnosis of PTSD is clearly outlined in Criterion A of the *DSM*. The trauma must involve experiencing or witnessing an event characterized by actual or threatened death or serious injury or a threat to the integrity of self or others leading to a response of intense fear, helplessness, or horror. Consider the following two scenarios:

Scenario 1: While receiving treatment at an inpatient psychiatric facility, a woman was awakened during the night by a male employee who restrained her in her bed while he groped her breasts and genitals. He warned her not to tell anyone, or he would “fix it” so she would have to stay longer in the hospital, and besides, “no one would take the word of a crazy person.”

Scenario 2: A man who was employed as an executive at a large corporation for a number of years was assigned to report to a new supervisor. He found himself frequently “butting heads” with his new boss, as they had very different approaches and personal styles. After 4 months, the man was called into his boss’s office without warning and fired. He was escorted to his office by security, allowed a few minutes to collect his personal belongings, and then escorted out of the building.

Both of these scenarios describe disturbing events that would be profoundly upsetting to most people. There is, however, a qualitative difference between the circumstances of Scenario 1 and those of Scenario 2. The events described in the first scenario are consistent with the type of trauma outlined in Criterion A, whereas the events described in the second scenario do not meet this criterion, as they do not involve actual or threatened death or serious injury or a threat to the integrity of self. Koch et al. (2005) recommended considering five questions about a traumatic event:

1. Was there a life-threatening injury as a result of the traumatic event?
2. Was there a physically disabling injury that could conceivably limit employment or accomplishment of important life goals?
3. Did the individual witness a death or severe injury?
4. Was there a realistic fear that the individual's life was in danger despite no obvious traumatic injury?
5. Was there some threat to their physical integrity (e.g., threatened or actual coercive sexual contact)? (p. 137)

If one or more of these questions is answered in the affirmative, it is likely that the event in question would meet the requirements of Criterion A.

It is important to remember, however, that the fact that an individual has experienced an extreme traumatic stressor does not necessarily mean that the individual will undoubtedly develop PTSD. Brunello et al. (2001) estimated that one-third of the population will be exposed to a trauma of this magnitude at some point in their lives, yet only 10% to 20% will develop PTSD. PTSD is twice as common among women as among men (Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995). The traumas most commonly associated with PTSD are combat exposure among men and rape and sexual molestation among women (Kessler et al., 1995). In the general population in the United States, the prevalence of PTSD has been estimated at 8% (American Psychiatric Association, 2000).

In addition to meeting the bar set by Criterion A, a diagnosis of PTSD requires the presence of additional symptoms described in Criteria B through F. These include persistently reexperiencing the traumatic event, avoiding stimuli associated with the trauma, and experiencing persistent symptoms of increased arousal. As noted by Brunello et al. (2001), several of the symptoms of PTSD are not specific to that disorder and overlap with other mental health conditions, including major depressive disorder, anxiety disorders, and substance abuse, which makes diagnosis more complicated. Nemeroff et al. (2006) suggested that the presence of Criterion C symptoms (i.e., avoidance and numbing) in the months following the trauma was most predictive of the development of PTSD.

It should be noted that, even if symptoms of PTSD do appear, they may not persist for months or years. Koch et al. (2005) noted that, among most representative samples of PTSD sufferers, there is as much as 50% spontaneous remission within the first year after experiencing the traumatic event. However, as many as 10% of those who do develop PTSD remain chronically distressed.

CAUSALITY

The second element in the legal definition of *harm* is causality—the attribution of changes in the plaintiff's condition to a specific event. Koch et al. (2005) noted that it is not necessary to prove that an event in question is the sole cause of the plaintiff's psychological injury, but it does have to be a contributing cause. In some

cases, existing symptoms might be worsened or exacerbated but not caused by the event. Conceptually, determining causality seems simple: Compare the plaintiff's current functioning with his or her functioning prior to the event. If there is a notable decrement in functional capacity, it must have been caused by the event in question. Unfortunately, the practical reality is far from simple. Assessing causality is clearly the most difficult task facing the forensic examiner in personal injury work. Greenberg, Otto, and Long (2003) pointed out that no psychological test can reliably assess what the plaintiff's functioning was like prior to the events in question or separate the effects of one trauma from another. Establishing the plaintiff's pre-injury functioning in the absence of contemporaneously created, extensively detailed, objectively based records is extremely challenging.

Assessing causality is made more difficult if the plaintiff manifested compromised functioning or experienced other traumatic events prior to the events in question. Consider the following example.

The plaintiff, a 28-year-old woman, was in a car accident in which two people died. She filed a lawsuit alleging, among other things, that she suffered from PTSD as a result of the accident. As a child, the plaintiff was sexually abused by her alcoholic father. By high school, she was frequently skipping classes, developed substance abuse problems, and cut her arms and legs to cope with stress. She dropped out of school and moved in with an older boyfriend who physically assaulted her. She eventually moved out and was working part time in a fast food restaurant at the time of the accident.

The woman described in this example has a history of multiple traumas predating the car accident. It is quite possible that her history of abuse led to the development of symptoms of psychological distress. Her history also shows evidence of impaired functioning prior to the accident, as evidenced by her difficulty in school (skipping classes, dropping out) and reliance on inadequate coping strategies (substance abuse, cutting). Despite this history, however, it is also possible that this woman suffered additional psychological damage as a result of the car accident. The difficulty for the examiner would be attempting to sort out which of her current symptoms (if any) are attributable to the accident and whether her functioning has been further compromised as a result.

Consider another scenario:

The plaintiff, a 36-year-old man, was in a convenience store when an armed robber entered the store and threatened to shoot the clerk unless he gave him cash. The clerk complied and the robber left. The man subsequently filed a lawsuit against the owner of the convenience store, alleging that, as a result of this experience, he had developed panic disorder and depression so severe he was unable to leave his house. As a result, he lost his job as an accountant, a position that he had held for 8 years. As a child, the man had been physically abused by his stepfather over a period of 10 years. Despite this, he had been

successful in school, had finished college, and had earned a graduate degree. At the time of the robbery, he owned a home and was engaged.

This case is an example of an eggshell plaintiff. Despite his history of unimpaired functioning, the man's experience of violent abuse at the hands of his stepfather made him especially vulnerable (like an eggshell is to cracking) to the stress of the armed robbery. Although his reaction was extreme, he may be entitled to compensation from the defendant for all the damage he suffered, even though it is in excess of what the "average" person would have experienced.

PRACTICAL ISSUES

Personal injury evaluations involve managing relationships with different parties including the plaintiff, the defendant, and their legal representatives. The nature of the evaluator's relationship with each party, including the degree of contact and the flow of communication, varies depending on which party has retained his or her services. It is important for the evaluator to understand how to effectively manage these relationships.

DEALING WITH ATTORNEYS

In most cases, the forensic examiner is retained by the attorney who represents one of the parties in the litigation. Contact usually is initiated by telephone. The attorney typically identifies the party he or she represents, describes the case in general terms, and inquires about the expert's experience and knowledge as it relates to the issues involved. The attorney also identifies the other parties and attorney(s) involved in the litigation to determine if the expert has a conflict of interest. The attorney then asks about the expert's interest in accepting the referral. The expert should clarify the role he or she is expected to play—that is, testifying expert or consultant—to ensure that there are no misunderstandings. Fees should be discussed in this initial conversation. The expert should request a financial retainer from the attorney along with a letter of agreement outlining the scope of the expert's projected involvement and payment arrangements. The expert should request that the retaining attorney provide all available records relating to the case, but the expert should not begin working on the case until the retainer and letter of agreement have been received. It is important to understand that the attorney, not the examinee, is the expert's client. As mentioned earlier, if the expert is to examine the plaintiff and offer testimony, the expert's commitment is to objectivity and impartiality, not to advocating for the retaining attorney's client.

DEALING WITH EXAMINEES

Examinees may present differently depending on whether the examiner was retained by their own attorney or by the opposing attorney. When the examiner is

retained by the plaintiff's attorney, the plaintiff may approach the examination with the notion that the expert has been hired to "help him out" or to "prove her case." Sometimes plaintiffs assume the examiner will continue to treat them after the initial evaluation. By contrast, when the examiner has been retained by the defense, the plaintiff may be guarded, defensive, or even hostile as a result of assuming the examiner has been hired to attack his or her case. Because of this, regardless of by whom the examiner has been hired, it is important for the examiner to help the plaintiff understand that he or she is objective and unbiased.

Before beginning the examination of the plaintiff, unless the examination has been court-ordered, the expert must obtain the plaintiff's informed consent. This is true even if the expert has been retained by the plaintiff's own attorney. The plaintiff should be informed about the following:

- The context of the evaluation (i.e., litigation)
- Who retained the examiner
- That the examiner will be assessing and not treating the plaintiff
- What the examination will consist of
- That the examiner will provide his or her opinion to the retaining attorney
- That the examiner might be called on to write a report and testify at deposition or trial

If the plaintiff refuses to participate, the examiner should contact the retaining attorney to discuss how to proceed.

Throughout the examination, the expert needs to walk a fine line in terms of building sufficient rapport to conduct a productive evaluation while remaining clear that the purpose of the meeting is not therapeutic. The examiner can facilitate this process by expressing a desire to understand the plaintiff's point of view, by emphasizing the expert's role as an independent fact finder, and by treating the plaintiff in a professional and respectful manner.

CONDUCTING THE EVALUATION

The evaluation of the plaintiff should address the specific questions presented by the retaining attorney. These questions usually include one or more of these:

- Does the plaintiff show evidence of a psychological injury? Sometimes this question will include asking the expert to comment on a specific diagnosis or to formulate a diagnostic impression.
- If the plaintiff shows symptoms or injuries, are they causally related to the actions of the defendant? The expert may be asked to specify if the defendant's actions were the sole cause, the primary cause, or a contributing cause of the plaintiff's injuries. The expert may be asked to identify other events or circumstances that may have contributed to the plaintiff's injuries. The expert

may also be asked if the plaintiff's own actions or inactions contributed to the damage he or she has suffered.

- How has the plaintiff's functioning changed as a result of the claimed injury? This question usually refers to changes in areas such as work capacity, educational performance, social relationships, family responsibilities, and self-care.
- Will the plaintiff need treatment? The expert may be asked to identify the type of treatment and the frequency and duration of the treatment that is being recommended.

As discussed earlier in this chapter, accurate opinions can be formulated only when the expert has collected data that are both relevant and sufficient. Data collection is especially challenging in a personal injury evaluation, as the expert is expected to form an understanding of the plaintiff's past as well as present condition. The use of multiple sources of data is essential in this process (Heilbrun, 2001). By utilizing a variety of sources, the expert will be able to compare information across sources so as to obtain a more complete and objective understanding of the plaintiff's condition and functional capacity.

DOCUMENTS AND BACKGROUND INFORMATION

It is recommended that the expert review records prior to meeting with the plaintiff (Witt & Weitz, 2007). Doing this allows the expert to gain an understanding of the plaintiff and the circumstances leading to the lawsuit. It also allows the expert to identify areas to explore in greater detail when the plaintiff is interviewed. The records involved in any case will vary considerably in volume and nature, depending on the circumstances of the plaintiff, the details of the litigation, and the plaintiff's history. Typically, available records will fall into one of three categories: medical and mental health treatment records, educational and work records, and legal documents.

Especially when there is a large volume of records, it is often helpful to begin by organizing the records chronologically and creating a timeline of events, both before and after the claimed injury. Doing this provides an overview of the plaintiff's situation and will call attention to any gaps in the information that has been provided. Written records can be especially useful in trying to understand the plaintiff's baseline functioning, particularly when the examiner has access to school transcripts, standardized test scores, or psychological or psychoeducational evaluations that were performed prior to the events in question. Treatment records, particularly if they are available both prior and subsequent to the claimed injury, can be valuable in understanding changes in the plaintiff's emotional functioning.

In reviewing legal documents associated with the case, it is often best to start with the formal complaint. This will lay out, from the plaintiff's perspective, the events leading up to the lawsuit, including the nature of the injuries for which the plaintiff

is seeking compensation. The plaintiff's deposition can be another source of useful information. Depositions tend to be quite detailed and cover topics including the plaintiff's functioning prior to the alleged injury, the events leading to the alleged injury, and the plaintiff's condition subsequent to the alleged injury. Depositions of the plaintiff's mental health treatment providers can help in understanding the plaintiff's condition and progress in treatment. This is especially true if the actual treatment notes are handwritten and difficult to decipher and/or lacking in detail.

CLINICAL INTERVIEW

The purpose of the clinical interview is twofold: to gather information from the plaintiff's perspective about his or her history, events surrounding the injury, and events subsequent to the injury; and to give the expert an opportunity to directly observe the plaintiff's appearance, speech, affect, and behavior. Sufficient time should be allowed for these purposes. A thorough interview usually takes a minimum of 3 hours due to the breadth of information that should be covered.

Given the amount of information to be covered, it is important for the expert to maintain control over the structure of the interview. The examiner should focus on obtaining detailed descriptive information from the plaintiff about his or her functioning before and after the injury rather than asking the plaintiff to provide conclusive statements about his or her condition. For example, rather than asking the plaintiff "Do you think you have PTSD?" the examiner might ask the plaintiff to describe his or her daily life both before and after the events in question. It is useful to start with open-ended questions and then probe for specific details or examples as necessary.

Self-reported data can vary in accuracy for a variety of reasons, including the unreliability of memory and limitations in insight. Several authors have suggested that the accuracy of self-reported data appears to be influenced by situational factors such as involvement in litigation or compensation seeking. Lees-Haley et al. (1997) and Williams, Lees-Haley, and Djanogly (1999) reported that, in compensation situations, examinees were more likely to report superior premorbid functioning and poorer current functioning, and they were more likely to exaggerate the number and severity of their symptoms than examinees who were not seeking compensation.

As an adjunct to the face-to-face interview and direct observation of the plaintiff, self-report questionnaires and surveys can be used to record and document the plaintiff's reported symptoms, complaints, and functional limitations. Self-report questionnaires (e.g., symptom checklists) should not be confused with psychological tests, in that the former were designed for treatment planning, have high face validity, and lack the means for assessing the validity of the examinee's responses. Because of these limitations, self-report questionnaires are useful for collecting information, but they should not be considered objective evidence of valid symptoms.

PSYCHOLOGICAL TESTING

When psychologists are asked to examine a personal injury plaintiff, it is usually assumed that the examination will involve psychological testing. Psychological testing can be a valuable source of information that can be used to both formulate and confirm hypotheses as well as to disconfirm hypotheses about psychological constructs relevant to the legal issue (Heilbrun, 1992). Melton et al. (2007) described psychological testing as more relevant in personal injury evaluations than in any other forensic context, as many of the constructs measured by psychological tests are directly related to the assessment of distress and impairments in functional abilities. Witt and Weitz (2007) suggested that psychological testing can be useful in three ways:

1. Testing may be helpful in identifying the presence of personality disorders. Such disorders are, by definition, present from early adulthood. Therefore, unless the plaintiff is a child or an adolescent, symptoms of a personality disorder could be assumed to have been present and may have had an influence on the plaintiff's functioning and adjustment, prior to the injury that is the subject of the litigation.
2. Psychological testing provides a standardized method for assessing the plaintiff's current symptoms, by comparing the plaintiff's symptom report to the reports of individuals in normal or clinical samples.
3. Psychological testing can help in determining if the plaintiff is exaggerating symptoms or otherwise dissimulating, through the use of stand-alone measures of response style as well as validity scales imbedded in multiscale inventories, such as those found in the MMPI-2 (Butcher et al., 2001).

Despite the potential usefulness of psychological testing, no standard battery has been consistently identified in the literature for use in personal injury evaluations, and Greenberg et al. (2003) pointed out that not all personal injury examinations require the same assessment instruments. In making decisions about which, if any, psychological tests to use, these authors stressed the importance of using measures that validly assess the constructs they are designed to assess and noted that psychological tests are a method for generating hypotheses rather than for directly answering psycholegal questions. In addition, Koch, Douglas, Nicholls, and O'Neill (2006) cautioned that many of the tests commonly employed in clinical and forensic assessment may not have been evaluated scientifically in an ecologically valid manner for the population being assessed in a personal injury evaluation. Several researchers have studied the frequency in which specific psychological tests are used in various forensic contexts (Archer, Buffington-Vollum, Stredny, & Handel, 2006; Boccaccini & Brodsky, 1999; Lally, 2003). The MMPI was the most frequently named instrument in these surveys. Boccaccini and Brodsky, for example, found the MMPI was used in 89% of emotional injury cases by the psychologists who responded to their survey.

Ultimately, the decision regarding which tests to select for a given evaluation is left to the examiner. A good starting point is to consider the psycholegal questions the evaluation is to address. Generally, these questions involve: (a) understanding the plaintiff's pre-event functioning, (b) understanding the plaintiff's current functioning, and (c) determining the credibility of the plaintiff's presentation during the evaluation. As pointed out by Koch, Nader, and Haring (2009), although psychological testing can be very useful in evaluating the plaintiff's current mental health functioning, there is no scientific support for the use of psychological testing to make definitive statements about an individual's functioning at some time in the past. Psychological testing may, however, be helpful in gathering information about sustained patterns of functioning that, when used in conjunction with data obtained from collateral sources, can provide some insight into the plaintiff's pre-injury capacities. In assessing the plaintiff's current emotional functioning, multiscale inventories, such as the MMPI-2 (Butcher et al., 2001) and the Personality Assessment Inventory (PAI; Morey, 1991), are particularly useful as these instruments facilitate an assessment of a broad range of psychopathology. In addition, these tests include sophisticated scales for the detection of exaggerated or other distorted response styles.

When selecting instruments to address the plaintiff's response style, the nature of the plaintiff's reported symptoms should be considered. In addition to the embedded validity scales in instruments like the MMPI and the PAI, stand-alone measures such as the Structured Interview of Reported Symptoms (SIRS; Rogers, Sewell, & Gillard, 2010), the Miller Forensic Assessment of Symptoms Test (M-FAST; Miller, 2001), and the Structured Inventory of Malingered Symptomatology (SIMS; Widows & Smith, 2005) can be used. Rosen and Powel (2003) explored the use of a neuropsychological forced-choice symptom validity test (Portland Digit Recognition Test) to detect symptom exaggeration in PTSD. They recommended including one of these instruments in forensic evaluations of PTSD if the plaintiff reports problems with memory or concentration. Regardless of the methods employed, it is important to remember that no test exists that can "prove" the plaintiff is malingering. Well-validated measures of symptom overendorsement (e.g., MMPI validity scales) can provide useful information about the plaintiff's approach to the assessment, but they are not sufficient grounds for inferring malingering (Koch et al., 2009).

COLLATERAL DATA

The use of collateral or third-party data is an established principle of forensic mental health assessment (Heilbrun, 2001). Collateral data can help to fill in information that is absent in the plaintiff's self-report, help the expert to form a clearer picture of the plaintiff's functioning prior to the injury, and corroborate information obtained from the plaintiff during the interview and from psychological testing.

There are many potential sources of collateral data. The specific sources selected will vary depending on the type of information the expert is seeking, the plaintiff's

situation, and the availability and willingness of sources to provide data. In general, the plaintiff's authorization should be sought before contacting or interviewing collateral sources; however, the plaintiff's consent may not be necessary when the examination has been court-ordered.

Collateral data, like other forms of data, are vulnerable to distortion and misinterpretation. Heilbrun, Warren, and Picarello (2003) described issues that can limit the accuracy of information obtained from collateral sources. These include reluctance to participate in the evaluation, bias, lack of specific expertise, suggestibility, and memory loss. Similar to the approach used when interviewing the plaintiff, interviews with collateral sources should focus on obtaining descriptions of the plaintiff's behavior as observed by the collateral source rather than conclusions about the plaintiff's condition. This information is best solicited with open-ended questions about the source's observations of the plaintiff followed by probing for specific examples to support the general statements.

INTERPRETATION OF FINDINGS

As noted by Witt and Weitz (2007), personal injury evaluations are challenging because "the evaluator is attempting to retrospectively reconstruct the plaintiff's mental condition before, during, and after an event or series of events." (p. 6). Although an assessment of current psychological functioning is important, the main thrust of the evaluation is on discerning whether a change in adjustment occurred at some time in the past and whether this was caused by some prior event. Koch et al. (2009) pointed out that, although it is possible to evaluate the plaintiff's current emotional functioning and provide descriptive information about his or her past functioning based on information provided by the plaintiff, written records, and collateral sources, there is no scientifically sound method for determining that the plaintiff's current condition was definitively caused by a specific event in the past. That said, the expert is often expected to offer some opinions as to the plaintiff's current condition (damage) and the relationship of this condition to the actions of the defendant (causality). Because causality is an "ultimate issue" question, not all courts permit experts to testify about causality.

ESTABLISHING DAMAGE

As discussed earlier in this chapter, establishing damage is not synonymous with assigning a diagnostic label. In order to establish damage, three factors must be considered:

1. Does the plaintiff manifest symptoms of psychological distress that affect his or her ability to function?
2. Are these symptoms and functional impairments valid?
3. Are these symptoms and functional impairments different from what the plaintiff was experiencing prior to the events leading to the lawsuit?

In addressing these questions, it is helpful to aggregate information from interviews, records, and testing regarding the plaintiff's past and present functioning and to consider each reported symptom and functional impairment separately. For example, if the plaintiff claims that since his accident, he has been too depressed to leave the house most days, the expert can review the plaintiff's treatment records, employment records, and information obtained from collateral sources to establish if the impairments the plaintiff is claiming were present prior to the accident. Next, the expert can examine test data, collateral information, and current treatment records to determine if these impairments are currently present. The expert should note specific examples and/or test findings that address each symptom and impairment in functioning that the plaintiff reports. Doing this allows the expert to corroborate specific pieces of information rather than attempting to confirm or disconfirm a particular diagnosis.

If requested by the retaining attorney, once the expert has reviewed these data and has reached some opinions about symptoms and impairments that are currently present, the expert can consider whether the plaintiff's symptoms relate to a specific diagnosis. This can be accomplished by reviewing each diagnostic criterion to determine if evidence exists that the plaintiff meets that criterion.

As noted earlier in this chapter, damage must be related to a change in the plaintiff's functioning, not just the existence of a valid diagnosis. The expert must consider how the plaintiff's ability to carry out important functions in his or her daily life has changed as a result of the emotional injury. Consideration of functional impairments should always be based on the capacity of the plaintiff prior to the injury.

ADDRESSING CAUSALITY

Psychological problems are almost always multidetermined. Genetic predisposition, early experiences, physiology, social support, substance abuse/exposure, culture, stress, injury, and disease may all play a role to one degree or another. Therefore, it is impossible to say with certainty that a particular event was the sole cause of an individual's symptoms. In attempting to address issues of causality, four factors can be considered:

1. Did the plaintiff's symptoms arise subsequent to the events in question?
2. Did the plaintiff experience other traumas that might have contributed to the development of the symptoms?
3. Are other conditions present (e.g., substance abuse, medical problems) that could explain, at least in part, the symptoms the plaintiff is experiencing?
4. Has there been a notable change in the plaintiff's functional capacity subsequent to the events in question?

By considering this information, the evaluator can begin to formulate opinions about the likelihood that the observed changes in the plaintiff are proximately related to the actions of the defendant.

COMMUNICATING FINDINGS

Communication of the expert's opinion may be accomplished in several different ways. These include the expert disclosure, a written report, a deposition conducted by the opposing party, and testimony at trial.

EXPERT DISCLOSURE

Once the examination is complete and the expert has formed opinions about the plaintiff, the retaining attorney should be contacted by telephone. The examiner should not write a report until this conversation takes place. After speaking to the expert, the attorney may decide that the examiner's opinion will not be helpful to his or her case and will not want a report. If the attorney determines that the expert's testimony would be useful, the attorney will request a written report. The attorney may also discuss an expert disclosure, which may be submitted prior to the report being completed. Although this document will most often be prepared by the attorney, the expert should carefully review the disclosure before it is submitted to ensure that his or her opinion has not been misrepresented or misconstrued.

WRITTEN REPORT

According to FRCP Rule 26, the expert report must contain six things:

1. A complete statement of all opinions the witness will express and the basis and reason for them
2. The facts or data considered by the witness in forming them
3. Any exhibits that will be used
4. The witness's qualifications, including a list of all publications authored in the previous 10 years
5. A list of all the cases in the previous 4 years in which the witness testified as an expert (both at trial and depositions)
6. A statement of the compensation to be paid to the witness

As long as this required information is included, the actual format and organization of the report is left to the expert. A suggested format that includes all the required elements follows. The reader may also wish to consult Weiner (Chapter 21 this volume) for a more in-depth discussion of forensic report writing.

Subject of Expert Opinion. The report should begin with a paragraph identifying the party who retained the expert and a brief statement describing the purpose of the evaluation.

Summary of Expert Opinion. A list of the expert's opinions that will be presented in the report should follow.

Qualifications. Then the expert should provide one or two paragraphs describing his or her academic and experiential credentials as related to the case. The final sentence should state that a copy of the expert's curriculum vitae (CV) is attached to the report as an addendum. The CV must contain a listing of the expert's publications.

Data Sources. This section should list all the records the expert reviewed, the examination procedures, and any collateral interviews.

Summary of Information Obtained From Records. This section summarizes the content of the records reviewed by the expert.

Examination of the Plaintiff. This section contains behavioral observations of the plaintiff, summary of the history of the events as reported by the plaintiff, the plaintiff's description of his or her injuries and functional impairments, and psychometric test results.

Expert Opinion. This section is a full description and rationale for each opinion listed under the "Summary of Expert Opinion." Each opinion should be supported by examples from the records, interviews, and/or test finding.

Signature. The report must be signed by the expert.

Addenda. Addenda should include the expert's CV, a list of cases in which sworn testimony has been provided in the previous 4 years, and an invoice or other documentation of the compensation the expert has received (or expects to receive) from the retaining attorney.

Sometimes the retaining attorney will ask for a draft report before the final version is sent. It is acceptable for an attorney to point out grammatical or factual errors in a draft report or to suggest that something be explained more fully or clearly. It is not acceptable, however, for the attorney to suggest changes in the examiner's opinions.

The examiner has a duty to carefully maintain all documentation, notes, and raw data from the evaluation, including information in electronic format, as most of this material is discoverable. Some communications between the retaining attorney and the expert may fall under the work-product exemption from discovery. In the federal court system, discovery is governed by FRCP 26. The work-product exemption allows an attorney to protect materials prepared in anticipation of litigation from discovery by the opposing attorney, including material prepared by expert witnesses. It does not protect the examiner's opinions, qualifications, prior testimony, compensation received, or data provided to the expert by the attorney—all of which are subject to discovery. Rule 26 was amended in 2010 to provide greater protection from discovery. Under this revision, draft reports and communications between the attorney and the expert may be exempt from discovery. It is important to remember that each jurisdiction has different rules regarding discovery. Ultimately it is up to the judge to determine what is and is not exempt from discovery. Because of this fact, the best practice is to assume that everything is discoverable unless told otherwise by the judge.

DEPOSITIONS

It is very likely that the opposing attorney will want to depose the expert in order to discover additional information about the examiner's opinions and the foundation for these opinions. This is also an opportunity for the opposing attorney to size up the expert in order to get a sense of how effectively the expert would testify at a trial. Typically the examiner will receive a subpoena commanding appearance for a deposition and to bring everything in the examiner's possession relating to the case. A deposition, in which the expert will provide sworn testimony, might last for as little as an hour or as long as several days. The opposing attorney (who will ask the questions), the retaining attorney, and a court reporter (who will record the proceedings) will be present. Sometimes the plaintiff attends the deposition as well. The rules of the deposition are explained at the outset. Briefly, all answers must be verbal. If a question is not understood, the witness should ask for clarification before answering. The retaining attorney may enter an objection for the record, but the witness is expected to answer the question regardless. It is important for the witness to listen carefully to each question and to answer truthfully and succinctly.

Occasionally, the opposing attorney may attempt to provoke the witness to see how the witness might react on the stand. If this happens, the best course of action is to remain calm and avoid reacting emotionally. The use of anger or humor is not recommended.

At the end of the deposition, the witness is asked if he or she wants to waive signature. The best answer to this question is no, in response to which the deposition transcript will be sent to the witness, who will have an opportunity to correct any errors that were made in recording his or her testimony or errors made by the witness. Typically, the expert's time spent at the deposition is paid by the attorney conducting the deposition, whereas preparation time usually is paid by the attorney who retained the expert.

COURTROOM TESTIMONY

Many personal injury cases do not end up going to trial. The parties can agree to a settlement at any time, including during the trial. (The reader may also wish to consult Otto, Kay, and Hess [Chapter 22 this volume], which provides more detailed information about expert testimony.) If the expert is required to testify in court, it is helpful to review the deposition transcript beforehand. If there are any discrepancies between the testimony given at deposition and the expert's testimony at trial, the opposing attorney will almost certainly point this out.

The examiner's report should form the basis for his or her testimony. The retaining attorney will directly examine the examiner using open-ended questions. In general, it is usually a good idea to review the report with the attorney ahead of time to ensure that the attorney understands the opinions and will ask questions that will facilitate the expert explaining his or her opinion to the judge and jury.

During testimony, the expert should speak clearly and avoid using jargon. Ideally, the expert should exude a competent, respectful, professional, yet approachable demeanor. After the direct examination is completed, the opposing attorney will cross-examine the expert. The cross-examination is intended to expose flaws in the expert's opinion or to raise questions about the expert's credibility. It is important for the expert to listen to all questions carefully before answering. If either attorney objects during direct or cross-examination, the expert should say nothing until the objection has been resolved by the judge. If the objection is sustained, the question must be answered. Whether responding to questions under direct or cross-examination, the expert should always keep in mind that his or her primary obligation is to serve the court by giving truthful, unbiased testimony.

CONCLUSIONS

Evaluating claims of emotional distress in personal injury cases creates unique challenges for forensic psychologists. Many of these challenges stem from the conflicting demands, duties, and expectations between the professions of law and psychology. Lawyers are obligated to vigorously advocate for their clients, while psychologists must remain objective, impartial, and fair. For this reason, it is imperative that psychologists performing these evaluations fully understand the legal context in which the evaluations take place. Lawyers and judges are not responsible for ensuring that psychologists remain faithful to the ethical demands of their profession. Psychologists must be vigilant in avoiding multiple roles and conflicts of interest that could undermine their ability to maintain objectivity and must defend against the partisan misuse or mischaracterization of their work.

When examining plaintiffs in personal injury cases, the focus must be on functional capacity rather than diagnosis. In order to prevail, the plaintiff must prove that he or she has been damaged, which is manifested by changes in the plaintiff's functioning and not by the assignment of a diagnostic label. A multimethod approach to evaluation is recommended, aggregating data from diverse sources in order to obtain a more accurate picture of the plaintiff's condition and functioning. Given the high stakes involved in tort litigation, consideration must be given to the possibility of exaggeration or feigning of symptoms. This is best explored through using a combination of embedded validity scales, stand-alone symptom validity measures, and the cross-validation of data from different sources. No one test finding should be used in isolation to reach the conclusion that the plaintiff is malingering.

Psychologists must be aware of the limits of their science, especially in terms of extrapolating from the present examination of the plaintiff to inferences about the plaintiff's functioning at some time in the past or predictions about his or her future condition. In order to be admissible, the evidence given by the expert witness must be relevant to the legal questions at issue and based on sound science. It is important to not overstate the certainty of one's opinions, to acknowledge conflicting data, and to give consideration to alternative hypotheses.

REFERENCES

- American Law Institute. (2000). *A concise restatement of torts*. St. Paul, MN: Author.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- American Psychological Association. (2010). *Ethical principles of psychologists and code of conduct* (2002, amended June 1, 2010). Retrieved from <http://www.apa.org/ethics/index.aspx>
- American Psychological Association (2013). Specialty guidelines for forensic psychology. *American Psychologist*, 68(1), 7–19. doi:<http://dx.doi.org/10.1037/a0029889>
- Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment*, 87, 84–94.
- Boccaccini, M. T., & Brodsky, S. L. (1999). Diagnostic test usage by forensic psychologists in emotional injury cases. *Professional Psychology: Research & Practice*, 30, 253–259.
- Brunello, N., . . . Zohar J. (2001). Posttraumatic stress disorder: Diagnosis and epidemiology, comorbidity and social consequences, biology and treatment. *Neuropsychobiology*, 43, 150–162.
- Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W. G., & Kaemmer, B. (2001). *MMPI-2: Manual for administration and scoring* (rev. ed.). Minneapolis: University of Minnesota Press.
- Federal Rules of Evidence Handbook, 2012–2013 Edition*. (2012). Newark, NJ: Matthew Bender.
- Greenberg, S. A. (2003). Personal injury examinations in torts for emotional distress. In A. M. Goldstein (Ed.), *Forensic psychology* (pp. 233–257). Vol. 11 in I. B. Weiner (Ed.-in-Chief), *Handbook of psychology*. Hoboken, NJ: Wiley.
- Greenberg, S. A., & Shuman, D. W. (1997). Irreconcilable conflict between therapeutic and forensic roles. *Professional Psychology: Research and Practice*, 28(1), 50–57. doi:<http://dx.doi.org/10.1037/0735-7028.28.1.50>
- Greenberg, S. A., Otto, R., & Long, A. (2003). The utility of psychological testing in assessing emotional damages in personal injury litigation. *Assessment*, 10(4), 411–419.
- Greenberg, S. A., Shuman, D. W., & Meyer, R. G. (2004). Unmasking forensic diagnosis. *International Journal of Law and Psychiatry*, 27(1), 1–15.
- Grisso, T. (2003). *Evaluating competencies: Forensic assessments and instruments* (2nd ed.). New York, NY: Kluwer Academic/Plenum Press.
- Heilbrun, K. (1992). The role of psychological testing in forensic assessment. *Law and Human Behavior*, 16(3), 257–272. doi:<http://dx.doi.org/10.1007/BF01044769>
- Heilbrun, K. (2001). *Principles of forensic mental health assessment*. New York, NY: Kluwer.
- Heilbrun, K., Warren, J., & Picarello, K. (2003). Third-party information in forensic assessment. In A. M. Goldstein (Ed.), *Forensic psychology* (pp. 69–86). Vol. 11 in I. B. Weiner (Ed.-in-Chief), *Handbook of psychology*. Hoboken, NJ: Wiley.
- Kane, A., Nelson, W., Dvoskin, J. A., & Pitt, S. E. (2013). Evaluation for personal injury claims. In R. Roesch & P. A. Zapf (Eds.), *Forensic assessments in criminal and civil law* (pp. 148–160). New York, NY: Oxford University Press.
- Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. B. (1995). Posttraumatic stress disorder in the national comorbidity survey. *Archives of General Psychiatry*, 52(12), 1048–1060.
- Knap, S. J., & VandeCreek, L. D. (2012). *Practical ethics for psychologists: A positive approach* (2nd ed.). Washington, DC: American Psychological Association.

- Koch, W. J., Douglas, K. S., Nicholls, T. L., & O'Neill, M. L. (2006). *Psychological injuries: Forensic assessment, treatment, and law*. New York, NY: Oxford University Press.
- Koch, W. J., Nader, R. & Haring, M. (2009). The science and pseudoscience of assessing psychological injuries. In J. Skeem, K. Douglas, & S. Lilienfeld (Eds.), *Psychological science in the courtroom: Controversies and consensus* (pp. 263–283). New York, NY: Guilford Press.
- Koch, W. J., O'Neill, M., & Douglas, K. S. (2005). Empirical limits for the forensic assessment of PTSD litigants. *Law & Human Behavior, 29*, 121–149.
- Lally, S. (2003). What tests are acceptable for use in forensic evaluation? A survey of experts. *Professional Psychology: Research & Practice, 34*(5), 491–498.
- Lees-Haley, P., Williams, C., Zasler, N., Margulies, S., English, L., & Stevens, K. (1997). Response bias in plaintiffs' histories. *Brain Injury, 11*, 791–799.
- Melton, G., Petrila, J., Poythress, N., & Slobogin, C., Lyons, P., & Otto, R. K. (2007). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (3rd ed.). New York, NY: Guilford Press.
- Miller, H. A. (2001). *M-FAST: Miller Forensic Assessment of Symptoms Test and professional manual*. Odessa, FL: Psychological Assessment Resources.
- Morey, L. C. (1991). *The Personality Assessment Inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Nemeroff, C. B., Bremner, J. D., Foa, E. B., Mayberg, H. S., North, C. S., & Stein, M. B. (2006). Posttraumatic stress disorder: A state-of-the-science review. *Journal of Psychiatric Research, 40*(1), 1–21.
- Rogers, R., Sewell, K. W., and Gillard, N. D. 2010. *Structured Interview of Reported Symptoms: Professional manual* (2nd ed.). Odessa, FL: Psychological Assessment Resources.
- Rosen, G. M., & Powel, J. E. (2003). Use of a symptom validity test in the forensic assessment of posttraumatic stress disorder. *Journal of Anxiety Disorders, 17*(3), 361–367. doi:[http://dx.doi.org/10.1016/S0887-6185\(02\)00200-1](http://dx.doi.org/10.1016/S0887-6185(02)00200-1)
- Shuman, D. W., & Greenberg, S. A. (1998). The role of ethical norms in the admissibility of expert testimony. *The Judges' Journal: A Quarterly of the Judicial Division, 37*, 4–43.
- Widows, M. R., and Smith, G. P. 2005. *Structured Inventory of Malingered Symptomatology professional manual*. Odessa, FL: Psychological Assessment Resources.
- Williams, C. W., Lees-Haley, P. R., & Djanogly, S. E. (1999). Clinical scrutiny of litigants' self-reports. *Professional psychology: Research and practice, 30*(4), 361–367.
- Witt, P. H., & Weitz, S. E. (2007). Personal injury evaluations in motor vehicle accident cases. *Journal of Psychiatry and Law, 35*, 3–24.

CHAPTER 8

Identifying and Treating Educational Disabilities

DANIEL J. RESCHLY

ASSESSMENT of educational disabilities is a major and sometimes controversial role of psychologists in the United States. Although many types of psychologists may be involved in the identification and treatment of children and youth with disabilities, this is a major role for the approximately 33,000 school psychologists employed by public educational agencies in the United States (www.ideadata.org). This chapter discusses the legal influences on the identification and treatment of disabilities in educational settings, including the nature of these disabilities, classification systems, and public policy that has shaped current practices.

GROWTH OF PSYCHOLOGISTS IN ASSESSMENT OF EDUCATIONAL DISABILITIES

Assessment of educational disabilities has been prominent in school psychology since its earliest days. Arnold Gesell, most likely the first person to use the title of school psychologist, was hired by the Connecticut State Department of Education in 1913 to examine school-age children suspected of being mentally deficient¹ (now intellectually disabled [ID]) (Fagan, 1987a, 1987b). Throughout the 20th century,

1. The American Association on Mental Retardation changed to the American Association on Intellectual and Developmental Disabilities in 2008. “Rosa’s Law” was signed by President Barack Obama in 2010, changing terminology in all federal documents from *mental retardation* to *intellectual disability* (ID). In all subsequent sections, the term *ID* is used rather than *mental retardation*.

The Diagnostic and Statistical Manual of the Mental Disorders-5 (DSM-5; American Psychiatric Association, 2013) also uses the term *intellectual disability*. The *DSM-5* terminology is consistent with recent proposed changes in the World Health Organization *International Classification of Diseases and Related Health Problems, 11th Revision* (Carulla et al., 2011).

increases in the number of school psychologists closely followed the expansion of special education services for students with disabilities (SWD). By 1975, there were at least a few school psychologists in every state, and although the national ratio was in excess of 4,500 students per school psychologist, some states had achieved ratios in the range of 2,000 to 3,000 students per psychologist (Kicklighter, 1976). School psychology employment expanded rapidly as a result of the changes discussed in the next section.

In 1975, a federal law was established requiring states to provide special education services in order to qualify for federal funding (Education of the Handicapped Act [EHA], 1975). This law has been reauthorized periodically over the past 35 years and is now titled the Individuals with Disabilities Education Act (IDEA; 1990, 1997, 2004). Although there is no federal special education legal mandate per se, since a state can decline the federal funding, all states have chosen to implement the law in order to receive the IDEA funding. The EHA and state legislative mandates regarding special education were enacted in the mid-1970s as a response to litigation in federal courts that established the right of students with disabilities to an appropriate education at public expense and other protections (see later section on evolution of legal influences) (Reschly & Bersoff, 1999; Yell, 2012).

Together the federal and state legislation led to an increase in identified SWD from about 3.3 million in 1976–1977 to approximately to 6.5 million SWD between the ages of 3 and 21 in the 2011–2012 school year (U.S. Department of Education [USDE], www.ideadata.org). As special education enrollments grew over the past 25 years, so too did the number of psychologists employed in the schools (see Figure 8.1). The USDE has collected data since 1977 on the employment of personnel providing services to SWD (see www.ideadata.org). From the first year these data were collected until the most recent year for which results are reported (2010), the number of psychologists employed by public educational agencies and serving

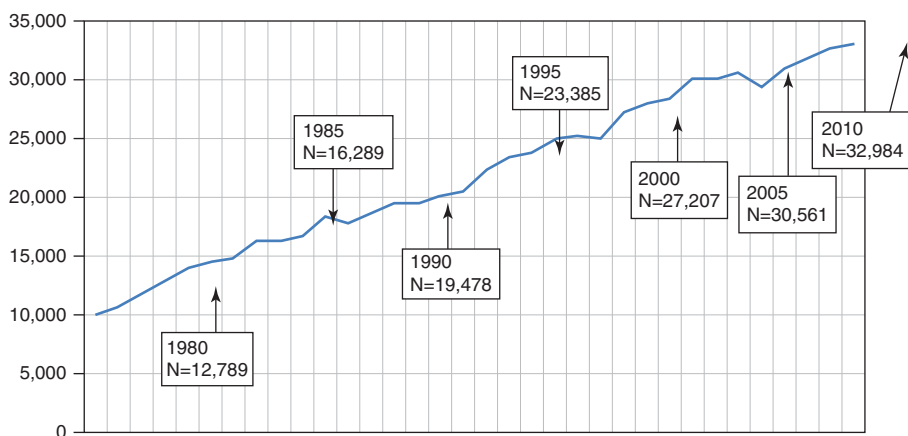


Figure 8.1 Growth of School Psychology 1977 ($N = 9,950$) to 2010 ($N = 32,984$)
 Source: Office of Special Education Programs, Annual Reports, www.ideadata.org

SWD grew from approximately 9,500 to approximately 33,000—over a threefold increase. Clearly, the growth in the number of psychologists serving SWD is closely related to the expansion of the number of identified SWD and special education programs.

EVOLUTION OF LEGAL INFLUENCES

Federal and state litigation and legislation establish the rights of SWD and their parents to individually designed educational programs based on a comprehensive evaluation. Although most areas of psychological practice are influenced by legal requirements, the assessment of educational disabilities as a prerequisite to the provision of mandated accommodations and special education and related services clearly is one of the most heavily regulated. These legal requirements have developed through dynamic and continuing cycles of litigation and legislation, each exerting reciprocal influences on the other (Reschly & Bersoff, 1999; Yell, 2012). Among those rights was a “full and individual” assessment that met explicit requirements and that, in most states, was conducted primarily by school psychologists.

LITIGATION: RIGHT TO EDUCATION AND PLACEMENT BIAS PROTECTIONS

Litigation was the first step in establishing the right of children ages 3 to 21 to appropriate educational services that met complex requirements. Detailed discussion of the early litigation has appeared in multiple sources (Reschly & Bersoff, 1999; Yell, 2012) and therefore will be discussed only briefly in this chapter. Two kinds of cases in the late 1960s and early 1970s profoundly influenced education services for SWD and the psychological services provided to them.

Right to Education. Two cases in the federal district courts—*Mills v. Board of Education* (1972) and *Pennsylvania Association for Retarded Children (PARC) v. Commonwealth of Pennsylvania* (1972)—established the right of SWD to educational services. Prior to these landmark cases, many SWD were either excluded from the public schools by local, district, and state policies or were provided educational services without extensive assessment to ensure the programs were matched to individual needs. In the cited cases, the plaintiffs asked federal district courts to apply the 14th Amendment concepts of equal protection and due process and force the states to provide appropriate educational services to *all* students. The courts agreed with the plaintiffs’ arguments, deciding that exclusion of SWD students from public schools constituted a violation of equal protection of the laws and due process. These early decisions led to the EHA and to state special education mandates. Basic principles that were established regarding these services are discussed later in the chapter in the review of federal and state legislation.

Placement Bias Litigation. Placement bias litigation, also applying the concepts of equal protection and due process, but advocating *fewer* rather than more special education services appeared in the federal district courts in the late 1960s and early 1970s, focusing on the overrepresentation of racial minority students in special education programs. Central to these cases were allegations of inappropriate assessments, use of biased instruments, and excessive reliance on intelligence testing. Three cases—*Diana v. State Board of Education* (1970), *Guadalupe Organization v. Tempe Elementary School District No. 3* (1972), and *Larry P. v. Riles* (1972, 1974, 1979, 1984, 1986, 1992) were filed on behalf of minority students placed in special programs, noting that minority students were placed in programs for students with mild ID at rates from 1.5 to 3 times the rate of placement for nonminority students.

The *Diana* and *Guadalupe* cases were settled by consent decrees that resulted in implementation of a number of safeguards in the identification of SWD, including determination of primary language, assessment instruments administered using procedures consistent with the student's primary language, reliance on nonverbal rather than verbal ability measures with English-language learners, greater reliance on adaptive behavior measures in diagnosis of mild ID, and due process protections. Some of the phrases that appeared in the *Diana* or *Guadalupe* consent decrees appear verbatim today in federal regulations and state education agencies (SEAs) special education rules (see later section on state and federal legislation).

The *Larry P.* case was somewhat different, in that matching the psychological assessment to the primary language of the home was not a central issue. The plaintiffs were African American students, and intelligence tests became the primary focus in the case. *Larry P.* resulted in injunctions in 1972 and 1974, restraining first the San Francisco Unified School District and, later, all California school districts, from using IQ tests with African American students that were ruled as having a discriminatory effect with African American students.

There is a clear irony in the early litigation. In one type of case, school districts were cited by the federal courts for violating children's equal protection and due process rights because special education services were *not* provided (*Mills*, 1972; *PARC*, 1972); in another type of case, the same constitutional principles were the basis for rulings that assessments by psychologists were inappropriate, leading to excessive and discriminatory placement of minority students in special education (*Diana*, 1970; *Guadalupe*, 1972; *Larry P.*, 1972, 1974, 1979, 1984, 1986, 1992). The same issues exist today regarding psychological services and the overrepresentation of minority students in special education (see placement bias and overrepresentation section).

Much more could be discussed about the *Larry P.* litigation, a case that appears to have fewer direct implications for psychological services today. *Larry P.* (1979) banned the use of intelligence tests in California if the result of testing was diagnosis of African American students with mild ID and placement in special education programs that were inferior to instruction in general education. The ban was expanded in 1986 to include all uses of intelligence tests with African American

students in California public schools. The ban was implicitly modified in *Crawford et al. v. Honig* (1992), and further court proceedings were anticipated in response to Judge Peckham's 1992 order to the California Department of Education to identify which of their current special education programs were "dead end and inferior" (p. 15). No response was made to this order, Judge Peckham died, and neither the original plaintiffs nor the defendants have initiated further proceedings in the case.

Larry P. is significant today because it illustrates the potential impact of the courts on psychological assessment and decision making. Other federal courts reached different decisions regarding the issues of intelligence test biases and overrepresentation of minority students in special education (*Marshall et al. v. Georgia*, 1984, 1985; *Parents in Action on Special Education (PASE) v. Hannon*, 1980; *S-1 v. Turlington*, 1971, 1981, 1986), blunting the impact of *Larry P.*

STATE AND FEDERAL LEGISLATION

In response to right to education consent decrees in *Parc* and *Mills* and the possibility of the federal courts mandating special education services, legislatures passed mandatory special education bills in every state by the mid-1970s. SWD had to be served by the public schools, and the educational programs had to meet certain standards. Generally the legislation provided state monies for these programs, thus alleviating part of the burden of providing expensive special education and related services incurred by local educational agencies (LEAs). Rather than attempting further discussion of state legislative mandates, the key principles in a subsequent federal law with which all states now comply is discussed. This federal law also incorporated the major principles from the pre-1975 litigation.

EHA/IDEA established federal funding mechanisms to support states' provision of services to SWD pending compliance with key principles that largely originated in the right to education and placement bias court cases. SEAs are required to monitor implementation of these principles in LEAs in order to receive federal funding. The key principles had significant influences on the development and implementation of psychological services:

- Free appropriate education
- Least restrictive environment
- Individualized education program
- Due process procedural safeguards
- Evaluations and reevaluations (formerly protection in evaluation procedures)

Free Appropriate Education. The most important of the general principles is the right of all students with disabilities to free, appropriate, publicly supported education programs (FAPE). Subsequent interpretation of EHA-IDEA by the courts has established that this right applies to all students, even those who have no apparent learning disability (*Timothy W. v. Rochester*, 1988, 1989). The effect of this

principle is that many more students are now diagnosed as disabled and provided special education services, and students with more severe disabilities—previously excluded from public schools—are present in school settings where they were to receive a full range of assessment and programming services. Psychological services were markedly expanded as more students were referred for eligibility determination needing full and individual evaluations. Moreover, students with more severe disabilities that required specialized services entered the public schools, leading to demands for more sophisticated psychological services.

For example, in the last decade, increasing numbers of school psychologists have augmented their basic graduate preparation by completing the Board Certified Behavior Analyst requirements (<http://www.bacb.com/index.php?page=1>) as a means to intervene effectively with more complex, challenging behavior.

Least Restrictive Environment. A second important EHA-IDEA principle is least restrictive environment (LRE), that is:

That to the maximum extent appropriate, children with disabilities . . . are educated with children who are non-disabled. That special classes, separate schooling or other removal of children with disabilities from the regular educational environment occurs only when the nature or severity of the disability is such that education in regular classes with the use of supplementary aids and services cannot be achieved satisfactorily. (Code of Federal Regulations [CFR] 34.300.114)

LRE had the effect of greater integration of SWD with general education students and greatly increasing the complexity of psychological assessments. It is important to note that neither the law nor any court mandates a full inclusion requirement that all students with disabilities be educated in general education environments all of the time. IDEA conceptualizes educational environments on a continuum from participation in the general education classroom for 80% or more of the school day to the most restrictive placements of self-contained classes with other SWD to placement in institutional settings. The trend over the past 30 years has been toward less restrictive placements and greater participation in general education settings. In 1980, approximately 35% of SWD spent 80% or more of the school day in general education settings. By 2011, that proportion changed from 35% to 60% (www.ideadata.org, December 2, 2012). The trend toward greater inclusion and participation in the general education curriculum produced greater demands for behavioral interventions for SWD to facilitate their functioning in general education settings.

Individualized Education Program. SWD are also guaranteed an individualized educational program (IEP) that is to be reviewed annually and modified as appropriate. The emphasis on greater individualization increased demands for direct assessment of academic skills and behaviors in natural classroom settings to establish

baseline levels of performance from which IEP goals can be developed and progress monitored.

Additional IEP requirements (IDEA, 1997, 2004) were established requiring assessment of progress in the general educational curriculum and participation of SWD in state and federally mandated achievement testing (34 CFR 300.320–324). A small proportion, approximately 2%, of the general student population can be excused from the mandated achievement testing due to the severity of their disabilities. Excused students must be provided alternative assessments that are aligned with the general education curriculum in the areas of communication, mathematics, social studies, science, and life skills.

Development and alignment of the alternative assessments with the general education curriculum are extraordinarily difficult. For example, consider the scaling and construct validity problems with representing achievement in communication skills from extremely low (e.g., eye movements in response to social stimuli) to extremely high levels (e.g., advanced written composition). Scale integration algorithms can be applied to this problem, but the critical issue is whether the same construct is being measured from the very low to very high skill levels in the example. Much work remains to be done in order to develop the underlying conceptual and technological basis for the alternative assessments (Bolt & Ysseldyke, 2008; Elliott, McKeivitt, & Kettler, 2002; Kettler et al., 2011; McDonnell, McLaughlin, & Morison, 1997).

Due Process Procedural Safeguards. The due process protections established in the courts, then mandated in state and federal legislation, guarantee the rights of children with disabilities—in the case of minor children, their parents or guardians—to five things:

1. Advance notice of decisions a school is contemplating
2. Information in an understandable form
3. Reasons the decision is contemplated
4. Participation in decision making
5. Discretion to approve or reject decisions

Complex hearing procedures are available to schools and parents to resolve disputes.

Children with disabilities or, depending on age, their parents have access to nearly everything a psychologist might do as part of the assessment of educational disabilities, including the right to review, but not necessarily copy, copyrighted test protocols. This allows parents and their legal advocates to scrutinize and challenge the work of psychologists in legal proceedings. IDEA and Family Education Rights and Privacy Act (34 CFR 99) mandates in regard to sharing assessment findings, including the documents and procedures used to produce the findings, and legal requirements that psychologists ensure test security, place psychologists in a difficult position, and easy resolutions to these dilemmas do not exist. Psychologists

must study the available ethical codes as well as materials from professional associations on the resolution of such dilemmas (American Psychological Association, 2010; National Association of School Psychologists, 2010).

One of these procedural safeguards is the right of parents to obtain an individual educational evaluation (IEE) conducted by qualified personnel outside of the educational setting if they disagree with the public school evaluation. Under certain circumstances, the costs of the IEE must be paid by the educational agency. The federal courts recently affirmed this right in an Alabama case challenging whether parents could obtain the IEE (*Phillip and Angie C. v. Jefferson Board of Education*, 2011). Regardless of who pays for the IEE, the school must consider its results and recommendations in formulating the IEP. IEEs are sometimes provided by psychologists educated in other specialties (clinical or counseling) or in interdisciplinary clinics of various kinds (e. g., local mental health, medical institutions). In either case, there are significant challenges in producing *educationally relevant* results and recommendations in evaluations conducted outside of educational settings.

Evaluations and Reevaluations (formerly Protection in Evaluation Procedures). The legal requirements that have the greatest influence on assessment appear in the IDEA (2004) Evaluations and Reevaluations regulations at 34 CFR 300.301–311. The key provisions of these regulations that are applicable to all evaluations of students referred for consideration of disability status are listed next.

- Full and individual evaluation before special education or related services are provided.
- “Tests and other evaluation materials used to assess a child are selected and administered so as not to be discriminatory on a racial or cultural basis; and are provided and administered in the child’s native language or other mode of communication, unless it is clearly not feasible to do so” (34 CFR 300.304 (c)). The law does not provide a definition of discrimination, nor are guidelines supplied to guide decisions about feasibility of matching native language. Both are considerable challenges. An implicit definition of discrimination and overrepresentation was included in IDEA (2004) and is discussed in the changing overrepresentation criteria section.
- A variety of assessment tools and strategies are used to gather relevant functional and developmental information about the child from multiple sources that are relevant to determination of disability status and formulation of the IEP.
- Tests and other assessment procedures are valid for the purpose for which they are used and administered by appropriately trained personnel.
- Tests and evaluation materials are appropriate to determine specific educational needs, not just “a single intelligence quotient” (34 CFR 300.304 (b)), and are selected and administered to assess the domain of behavior intended, not the effects of the disability on the domain. For example, this provision

would protect visually impaired and blind children from being administered an individual IQ test that requires visual spatial reasoning using visual stimuli.

- “The child is assessed in all areas related to the suspected disability, including, *if appropriate*, health, vision, hearing, social and emotional status, general intelligence, academic performance, communicative status, and motor abilities” (emphasis added) (34 CFR 300.304).
- The evaluation must be sufficiently comprehensive to identify all the child’s special education and related services needs (34 CFR 300.304 (c)).
- Technically sound instruments and strategies are used that may assess the relative contribution of cognitive and behavioral factors, in addition to physical or developmental factors, and provide relevant information that directly assists persons in determining the educational needs of the child.
- In IDEA (2004), the mandatory triennial reevaluation of disability status was made optional, depending on a determination of the need for additional data through an extensive review involving parents and professionals.
- Initial evaluations and, if needed, reevaluations of disability status must be based on consideration of classroom-based assessments and observations, eligibility for disability status, present levels of performance and educational needs, documented need for special education, and, in a reevaluation, whether additions or changes in special education and related services are needed to participate in the general education curriculum.
- Eligibility decisions must be based on the individual evaluation, made by a group of persons including parents, using the categories of disability defined in the law at 34 CFR 300.8 (see later discussion on the legal basis for classification of students with disabilities). The disability cannot be due to lack of instruction in reading or mathematics or limited English proficiency.
- The educational agency must document careful consideration of information from a variety of sources including “aptitude and achievement tests, parent input, teacher recommendations, physical condition, social or cultural background, and adaptive behavior” (34 CFR 300.304 (c)). If the child is determined to be eligible for disability classification and needs special education, an IEP must be developed that meets extensive requirements (34 CFR 300.320–328).

A virtual revolution in rights of SWD has occurred over the past 25 years, and conceptions of best psychological practices continue to evolve (Thomas & Grimes, 2008). Numerous legal requirements have been established where relatively little legal requirements existed previously, influencing directly the psychological services in schools. Many of these requirements, as noted before, are ambiguous and subject to different interpretations. The general effect, however, is clear. The work of psychologists in the schools regarding assessment of educational disabilities is governed by extensive legal requirements, especially in the diagnosis of disability status and in the determination of special education need.

COMPLIANCE MONITORING

Implementation of the EHA/IDEA statutes is monitored by federal and state authorities with the ultimate sanction of denying funding to SEAs and LEAs that fail to implement the law adequately. In fact, withholding or withdrawing funding rarely occurs, and the amounts involved typically are a small proportion of overall federal or state support. Moreover, federal monitoring of SEAs and LEAs since 1975 has resulted in findings of numerous flaws in the implementation of the EHA/IDEA, a pattern that continues today; however, no SEA has ever been found to fail completely or to defy the IDEA principles. The main influence of compliance monitoring comes not from withholding funds but through public embarrassment for SEA and LEA authorities through public disclosure of violations. This mechanism is highly motivating to education officials. The cycles vary, but monitoring usually occurs in intervals of 3 to 5 years and is carried out by the USDE Office of Special Education Programs (OSEP).

In 2004, OSEP notified states that specific numerical outcome indicators were established and would be used in the future to monitor compliance with IDEA. The outcome indicators monitor states' SWD progress in improving inclusion in general education settings, achievement on state high-stakes tests, timeliness of evaluations and IEP development, and reduction/elimination of overrepresentation of minority students. Recently published reports include results of compliance monitoring focusing on outcomes using four categories of implementation: Meets Requirements, Needs Assistance, Needs Intervention, Needs Substantial Intervention (USDE, 2012).

SECOND-GENERATION LITIGATION AND LEGISLATION

Over the past 35 years, the rights to education and placement bias litigation and legislation have continued to evolve. Subsequent legislation and litigation have not established new legal principles; rather, the legal proceedings have refined the meaning of the basic principles described previously. For example, the right of SWD to an appropriate education was particularly ambiguous. What is appropriate: the best possible education? education equal to the educational opportunities of students without disabilities? or education consistent with application of best educational practices?

Right to Education and LRE. In the *Board of Education v. Rowley* (1982), "appropriateness" was defined as "reasonably calculated to produce educational benefit" (p. 3049) rather than ideal, most effective, or best educational program. The *Rowley* court further ruled that the law required access rather than guaranteeing any degree or level of benefit. The right to education principle continues to be litigated frequently in due process hearings and in state and federal courts (e.g., *Cedar Rapids Community School Dist. v. Garret F.*, 1999; *Forest Grove School District v. T.A.*, 2009).

Similar to the right to education, the LRE principle is ambiguous, with differing interpretations of what the following phrase means in educational and psychological practice: “to the maximum extent appropriate, children with disabilities . . . are educated with children who are non-disabled.” Four criteria were established (*Board of Education v. Holland*, 1992, 1994) to determine whether a student with a disability can be placed by the school outside of the general education classroom in a more restrictive setting, such as a special class:

1. The educational benefits available to the child in a regular classroom supplemented with appropriate aids and services, as compared to the educational benefits of a more restrictive special education placement
2. The nonacademic benefits to the handicapped child of interaction with non-handicapped children
3. The effect of the presence of the handicapped child on the teacher and other children in the regular classroom
4. The costs of supplementary aids and services necessary to mainstream the handicapped child in the regular classroom setting

The LRE decisions must be data based, establishing the need for more behavioral assessment through observation of natural environments to determine the effects of a student with a disability on other students and the degree to which the SWD benefits from being in the general education environment. Further legal refinement of the LRE inclusion requirements continues to the present (Yell, 2012).

Placement Bias and Overrepresentation. The EHA-IDEA principle of nondiscrimination continues to be refined in the federal courts and in other adjudicative forums. The placement bias litigation that resulted in either court-approved consent degrees or judicial opinions from 1975 to 2010 is discussed briefly here because relatively little activity has occurred over the past 10 years. Thorough discussions appear elsewhere (Reschly & Bersoff, 1999; Yell, 2012). The few placement bias cases established in the last 15 years generally resulted in decisions contrary to plaintiffs’ contentions of intelligence test bias and discrimination due to overrepresentation of minority students in special education (*Amber Blunt et al. v. Lower Merion School District et al.*, 2011; *Coalition to Save Our Children v. State Board of Education*, 1995).

Concepts of Fairness. The conception of fairness adopted by a court has a profound impact on decisions dealing with disproportionate representation. Two competing conceptions of fairness create enormous tensions that are apparent in the placement bias litigation and in other judicial analyses of disproportionate representation (e.g., *Gratz et al. v. Bollinger et al.*, 2003; *Grutter v. Bollinger et al.*, 2003). *Equal results* means that the same outcomes have to be achieved for all sociocultural groups. In order to achieve equal results, differential treatment of individuals often is necessary (e.g., adding points for diversity, consideration of a broad set of criteria

that values minority status as part of a diversity goal, changing requirements for disability identification with minority students). The main alternative is an *equal treatment* conception of fairness, which requires that all individuals with similar characteristics be treated in the same ways regardless of race, ethnicity, social class, or gender. Equal treatment allows disproportionate outcomes by race, ethnicity, social class, or gender as long as individuals are treated comparably.

The inherent weakness of equal treatment is that long-standing patterns of disproportionate outcomes continue to exist, including, for example, differential placement of students by race in special education or admission to professional schools. There is no easy solution to the problems of disproportionate representation in several domains of public and private life in the United States, but recognizing the tension between competing notions of fairness and the inherent importance of both conceptions is important to understanding legal analyses and public policy. These tensions significantly affect legal interventions regarding special education disability diagnosis.

Research on equal treatment in the assessment of disabilities is rarely conducted, due to the reluctance of SEAs and LEAs to allow research on individual decision making related to identification of minority students as SWD. The few rigorous studies that exist indicate that equal treatment criteria are met in current assessment services with African American and White students suspected of being disabled or who are disabled (*Coalition to Save Our Children v. State Board of Education*, 1995; Hosp & Reschly, 2004; MacMillan, Gresham, Lopez, & Bocian, 1996). Of course, the null hypothesis of equal treatment cannot be proven because of the impossibility of conducting research in all possible situations.

Plaintiffs' efforts to litigate issues related to assessment and identification of minority students with disabilities are markedly undermined by changes in criteria for discrimination in the federal courts. Skiba, Eckes, and Brown (2009–10) described the increasingly narrow criteria used in federal courts to adjudicate claims of discrimination, including the particularly difficult requirement applied to many cases that *intent* to discriminate must be proved by plaintiffs. Clearly, overrepresentation or disparate impact is now insufficient in the federal courts, creating significant headwind against claims of discrimination in psychological assessment and decision making.

Changing Overrepresentation Criteria. The IDEA (1997, 2004) supplemented the nondiscrimination requirement in the Evaluations and Reevaluations of the regulations with the addition of requirements to monitor and respond to significant disproportionality. State and local educational agencies were required to collect and examine data to determine if significant disproportionality existed in the categorical designation of children or in the LRE placement options. If significant overrepresentation existed, the SEA or LEA was required to revise policies and procedures (34 CFR 300.646).

These regulations are the closest provision in the law to date for defining a conception of discrimination and criteria determining whether nondiscriminatory assessment and placement exist. Guidance on what accounts for “significant disproportionality” has been discussed extensively, but OSEP officials have not announced specific criteria due to their interpretation of the *Gratz* and *Gruter* decisions that appear to prohibit the use of specific numerical procedures to produce more minority participation in higher education. In prior litigation, court-approved consent decrees applying similar criteria have established narrow ranges for differences between disability identification rates between minority and nonminority groups (e.g., *Mattie T. v. Johnson*, 2003). However, these court cases are not directly relevant, leaving to the states discretion in defining what the term *significant* means regarding disproportionality. Most states to date have adopted excessively lenient criteria (Albrecht, Skiba, Losen, & Middleberg, 2011).

LEGAL BASIS FOR CLASSIFICATION OF STUDENTS WITH DISABILITIES

Legal requirements determine the criteria used to assess disability eligibility and special education needs. *Conceptual definitions* specifying the key domains of behavior for 13 disabilities are provided in IDEA at 34 CFR 300.8. States have discretion in whether to use the exact nomenclature in IDEA, but all SEAs and LEAs must ensure that students with these kinds of disabilities are provided appropriate educations that meet IDEA requirements. For example, intellectual disability is defined in IDEA, but SEAs have discretion to use different terminology, such as mental disability and cognitive impairment. The general trend is toward more consistent adoption of federal conceptual definitions in state special education rules (Reschly & Hosp, 2004).

Classification criteria specifying the pattern and levels of performance for disability diagnosis are provided for only one of the IDEA disabilities in federal statute or regulations: specific learning disabilities (SLDs) (34 CFR 300.307–311). Classification criteria for all other disabilities are left to the states, and large classification criteria differences between the states have existed throughout the history of EHA/IDEA (Patrick & Reschly, 1982; Reschly & Hosp, 2004). Psychologists diagnosing educational disabilities in the context of special education must use the criteria usually specified in state rules for special education rather than other authoritative sources, such as the text revision of the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR)*; American Psychiatric Association, 2000).

The SLD regulations that existed from 1977 to 2004 were changed in IDEA (2004), effective in October 2006. The SLD identification requirements prior to 2006 required the establishment of a “severe discrepancy between intelligence and academic achievement” (IDEA, 1997, 34 CFR 300.541 (a)), necessitating in 48 of 50 states prior to 2006 the administration of an individually administered intelligence test and multiple achievement tests (Reschly & Hosp, 2004). The states employ widely discrepant criteria regarding the size of the severe discrepancy and the statistical

method for its determination. Criteria for SLDs are particularly important, because it is the disability category with the highest identification.

The IDEA regulations published in 2006 permit two primary methods to determine eligibility in SLDs: determination of a pattern of strengths and weaknesses that a team determines to be related to SLDs, or determination of the student's response to intervention using scientifically based interventions and progress monitoring. Each of the methods involves considerable judgment on the part of professionals as well as the collection of a wider variety of data by school psychologists in eligibility determination and reevaluations. Although there appears to be greater ambiguity in and greater degrees of professional judgment in making SLD eligibility decisions, the identification of SLDs has declined over the past 8 years (Snyder & Dillow, 2011).

The IDEA disability conceptual definitions reflect a mixture of medical and social system models (Reschly, 1996) as well as disability conceptions from other sources, such as the American Association on Intellectual and Developmental Disabilities (Schalock et al., 2010) and *DSM-IV-TR*. The latter is particularly influential on the identification of attention-deficit hyperactivity disorder (ADHD; often served in IDEA under the category of "other health impaired") and autism spectrum disorder.

DISTRIBUTION OF DISABILITIES ACROSS CATEGORIES AND PROBABLE ETIOLOGIES

Identification data can be interpreted more readily if a distinction is made between low- and high-incidence disabilities, a distinction that largely parallels the different etiologies of educational related disabilities, biologically based disabilities, and functional/behavioral deficits that do not have an identifiable biological cause. The distinctions to be described and in Table 8.1 are approximate but yield further understanding of the variations among SWD and the differential costs and psychological services associated with different kinds of SWD. Five of the 13 disabilities defined in IDEA (2004; 34 CFR 300.8) clearly are low-incidence disabilities (3 per 1,000 or less) with biological bases (see Table 8.1). Disabilities with clear biological foundations are deaf/blind, hearing impaired and deaf, multiple disabilities, orthopedic impairment, and visual impairment/blindness. The overall identification for the five low-incidence disabilities is less than 1% of the public school enrollment.

Four disabilities have a moderate identification varying from 0.7% to 0.9%. Three of these disabilities do not fit easily into either the biological or functional behavior etiologies. Autism with severe symptoms appearing before age 3 is best conceptualized as a biologically based disability. Milder forms of autism (Asperger's syndrome) generally are diagnosed after school entry and likely fit better into disabilities that do not have a clear biological basis and are best understood as functional/behavioral. Intellectual disability likewise is a mixture of disability models. The mild level of intellectual disability defined by intellectual functioning in the IQ = 60 to 75 range typically has no underlying identifiable biological basis, whereas more severe levels (e.g., Down syndrome) nearly always have a biological

Table 8.1
SWD Identification by Disability Category/Etiology and Typical Psychological Services

Disability Category	Low, Moderate, High Prevalence	Identifiable Biological Basis	U.S. 2009–10 Identification	Psychological Services Typically Provided
Deaf/Blind	Low	Yes	0.0%	Initial eligibility by medical specialists.
Hearing Impaired and Deaf	Low	Yes	0.2%	Psychologists involved with determination of
Multiple Disabilities	Low	Yes	0.3%	levels of performance across multiple areas of
Orthopedic Impairment	Low	Yes	0.1%	functioning and academic and behavioral needs
Traumatic Brain Injury	Low	Yes	0.1%	and interventions. Some psychologists
Visual Impairment and Blindness	Low	Yes	0.1%	specialize in low-incidence disabilities; however, many psychologists spend little time with low-incidence disabilities.
Autism	Moderate	Mix of causes	0.8%	Psychologists have multiple responsibilities, including participation in eligibility
Developmental Delay	Moderate	Mix of causes	0.7%	determinations and determination of academic
Intellectual Disability	Moderate	Mix of causes	0.9%	and behavior needs and interventions. Psychologists often develop behavioral interventions to deal with challenging behaviors.
Other Health Impaired	High	Mix of causes	1.4%	Psychologists extensively involved, leading the eligibility determination process and in determining current academic and behavior needs and interventions. Psychologists spend about 65% of their time with high-incidence disabilities. Psychologists typically lead the response to intervention process focused on prevention, early identification/treatment, and disability diagnosis if necessary.
Emotional Disturbance	Moderate	Increasingly Functional/Behavioral		
Specific Learning Disability	High	Functional/Behavioral	0.8%	
		Functional/Behavioral	4.9%	
Speech Language Impairment		Functional/Behavioral	2.9%	Little involvement except for students with speech/language impairments with comorbid disabilities.

basis. Developmental delay is used with children ages 3 to 9 whose disability status is ambiguous. Some children with developmental delay have biologically based disabilities, whereas other disabilities are more functionally/behaviorally based. Emotional disturbance nearly always is explained best as a functional/behavioral disability. The combined identification of the disabilities in the mixed etiology categories is 3.2% of the preschool to 12th grade (P–12) public school enrollment.

Three disabilities have identification rates over 1% of the P–12 enrollment: other health impaired, specific learning disability, and speech or language impairment. Prevalence in other health impaired has changed significantly over the last decade from about 0.4% to 1.4%, as it has become a mixture of children with significant biological disorders, such as epilepsy, diabetes, and asthma, who require specially designed instruction and related services, and children with ADHD, a condition without an agreed-on etiology. It is highly likely that ADHD now accounts for 75% or more of the children in the other health impaired category.

High-incidence disabilities typically are not identified prior to school entrance. Chronic achievement problems, particularly in reading, sometimes accompanied by disruptive behavior, lead to referral to special education by teachers and psychological services related to eligibility determination and programming academic and behavioral interventions.

The identification of low-incidence disabilities likely is relatively constant across states. All states have close to 1% of the overall enrollment needing high-cost special education services. In addition, some students in the mixed etiology disability categories require very costly special education and related services while the costs for other students in the moderate-prevalence categories are much lower. Moreover, disability identification of some students in the moderate- and high-identification categories can be prevented through strong general education programming, thus preventing referral and special education placement. The large variations among states to be discussed arise far more from high- than from low-identification disabilities.

SEA and LEA Variations in SWD Identification. SEAs and LEAs vary dramatically in SWD identification (Singer, Palfrey, Butler, & Walker, 1989). Large and unexplained LEA variations occur even among LEAs within the same state with demographically similar student populations and equivalent resources, presumably applying the same disability identification criteria. In some commentaries, poverty is implicated as a contributor to higher SWD identification. Research indicates that LEA poverty rates do not explain variations from district to district. In fact, urban districts with high poverty rates and high racial/ethnic diversity appear to have lower SWD identification. The highest SWD identification actually occurs in affluent suburban schools (USDE, 2001).

Large state-to-state variations in identification of SWD exist. Figure 8.2 was constructed using the identification data from Snyder and Dillow (2011, Table 48). The highest SWD identification was in Rhode Island (18.1%), the lowest in Texas

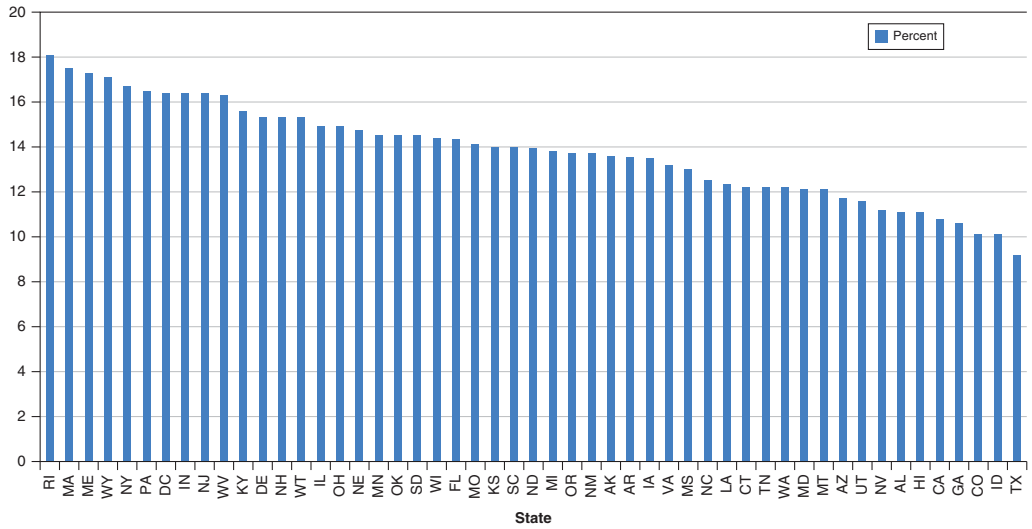


Figure 8.2 Identification of Students With Disabilities by State
 Source: www.ideadata.org

(9.2%). The median was 13.9% with half the states above and below that level. The national weighted mean in 2009–2010 was 13.1% (weighted by state SWD numbers and state P–12 public school enrollment).

Few efforts have been made to explain the large state variations in identification, and none has been successful. Obviously, the SWD identification is related to the number of full and individual evaluations and reevaluations conducted by psychologists. A particularly interesting element is the relationship between the numbers of school psychologists and SWD identification rates. No relationship exists between the ratio of psychologists and the number of SWD. For example, Connecticut has the most favorable ratio of about 500 students to each psychologist. However, that state’s SWD identification level is significantly below the national median SWD state identification rate. Psychologists’ roles, especially the degree to which broader services are provided beyond assessments for eligibility, are very much influenced by ratios of students to psychologists (Reschly, 2000).

SYSTEM REFORM TRENDS AND IMPLICATIONS FOR PSYCHOLOGICAL SERVICES

The system reform trends discussed in this chapter and in prior editions of the *Handbook of Forensic Psychology* (Reschly, 2006) have been substantially implemented in many settings and are in various stages of adoption in nearly all SEAs and LEAs. The terminology varies by place, either response to intervention or, more recently, multi-tiered systems of support (see Figure 8.3). The key principles of response to intervention (RTI) are early identification and intervention with academic and behavioral problems and application of scientifically based interventions that have large effect sizes (Kavale, 2005; Mayer, Sulzer-Azaroff, & Wallace, 2011; Shinn &

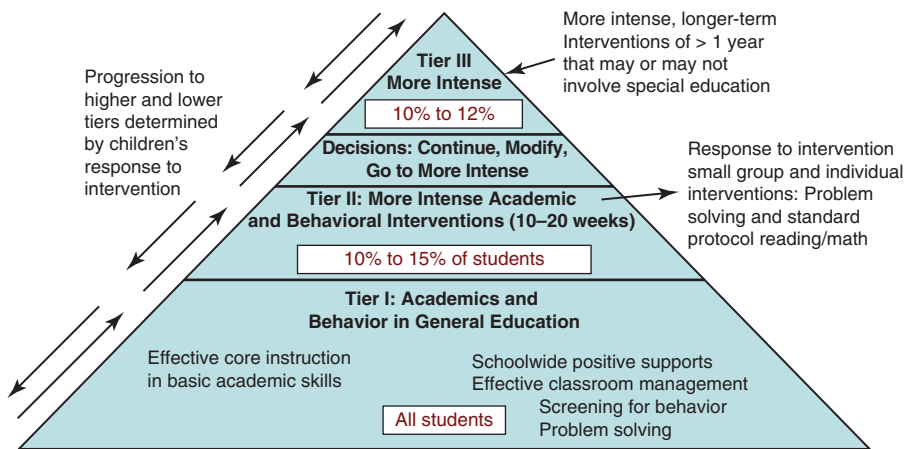


Figure 8.3 Multi-Tiered Systems With Tiers Varying in Intervention Intensity and Measurement Precision

Walker, 2010), are monitored closely, and, if progress is insufficient, strengthened through intervention modifications (Reschly & Bergstrom, 2009). The goal is to bring the best of psychological knowledge to the practical solution of school based academic and behavioral problems.

Psychologists' roles have changed significantly in RTI services-based systems, focusing less on use of standardized tests of ability and achievement and far more on direct measures of skills in natural settings that are useful for determining educational need, designing interventions, monitoring progress, evaluating outcomes, and determining eligibility if the interventions are unsuccessful. The criteria for psychological services increasingly emphasize change in child/youth skills and competencies rather than largely descriptive reports of deficits and needs. This approach to psychological services is more consistent with the experimental than with the correlational tradition in psychology as well as with the short-run empiricism described by Cronbach (1975). The short-run empiricism, or what now is called problem solving, is a promising replacement for interventions guided by aptitude by treatment interactions that have failed and continue to fail in the literature (Pashler, McDaniel, Rohrer, & Bjork, 2009).

CONCLUSIONS

Legal influences on school psychology expanded enormously over the past four decades and will continue to evolve. Further influence through the gradual evolution of case law and the enactment of legislation should be anticipated. The uneasy relationship between the courts and psychologists and the occasional misuse and distortion of psychological evidence by the courts require substantial additional efforts toward mutual understanding. Greater appreciation on the part of psychologists for the essential role of the courts in determining the educational rights of students with disabilities is needed, as is greater understanding by the courts of the strengths and limitations of psychoeducational assessment. Better assessment can produce better evidence that will in turn improve legal decisions that affect the lives of children and youth.

REFERENCES

- Albrecht, S. F., Skiba, R. J., Losen, D. J., & Middleberg, L. (2011). Federal policy on disproportionality in special education: Is it moving us forward? *Journal of Disability Studies, 20*, 1–12.
- Amber Blunt, et al. v. Lower Merion School District, et al. U.S. District Court for the Eastern District of Pennsylvania, Civil Action No. 07–3100 (2011).
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

- American Psychological Association. (2010). Ethical principles of psychologists and code of conduct (including 2012 Amendments). Retrieved from <http://www.apa.org/ethics/code/index.aspx?item=14>
- Board of Education v. Holland, 786 F. Supp. 874 (E.D. CA 1992) (9th Cir. 1994).
- Board of Education v. Rowley, 102 S. Ct. 3034 (1982).
- Bolt, S. E., & Ysseldyke, J. (2008). Accommodating students with disabilities in large-scale testing: A comparison of differential item functioning (DIF) identified across disability type. *Journal of Psychoeducational Assessment, 26*, 121–138.
- Carulla, L. S., Reed, G. M., Vaez-Azizi, L. M., Cooper, S. A., Leal, R. M., Bertelli, M., & Saxena, S. (2011). Intellectual developmental disorders: Towards a new name, definition, and framework for “mental retardation/intellectual disability” in ICD 11. *World Psychiatry, 10*(3), 175–180. Retrieved from www.ncbi.nlm.nih.gov/pmc/articles/PMC3188762
- Cedar Rapids Community School Dist. v. Garrett F., 526 U.S. 66 (1999).
- Coalition to Save Our Children v. State Board of Education (1995), Civil Action Nos. 56-1816-1822-SLR, U.S. District Court for the District of Delaware.
- Crawford, et al. v. Honig, No. C-89-0014 RFP U. S. District Court, Northern District of California, Complaint for Declaratory Judgment, May 1988; Order, September 29, 1989; Memorandum and Order, August 31, 1992.
- Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist, 30*, 116–127.
- Diana v. State Board of Education, No. C-70-37 RFP U.S. District Court, Northern District of California, Consent Decree, February 3, 1970.
- Education of the Handicapped Act (1975). 20 U.S.C. 1400-1485, 34 CFR-300.
- Elliott, S. N., McKevitt, B. C., & Kettler, R. (2002). Testing accommodations research and decision-making: The case of “good” scores being highly valued but difficult to achieve for all students. *Measurement and Evaluation in Counseling and Development, 35*, 153–166.
- Fagan, T. K. (1987a). Gesell: The first school psychologist, Part I. The road to Connecticut. *School Psychology Review, 16*, 103–107.
- Fagan, T. K. (1987b). Gesell: The first school psychologist. Part II: Practice and significance. *School Psychology Review, 16*, 399–409.
- Family Educational Rights and Privacy Act, 20 U.S.C. § 1232g; 34 CFR Part 99 (1974).
- Forest Grove School District v. T.A., 557 U.S. No. 08–305 (2009).
- Gratz et al. v. Bollinger et al., No. 02-516. U.S. Supreme Court, Opinion, June 23, 2003.
- Gutter v. Bollinger et al., No. 02-241. U.S. Supreme Court, Opinion June 23, 2003.
- Guadalupe Organization v. Tempe Elementary School District No. 3, No. 71–435 (D. Ariz., January 24, 1972) (consent decree).
- Hosp, J. L., & Reschly, D. J. (2004). Disproportionate representation of minority students in special education: Academic, demographic, and economic predictors. *Exceptional Children, 70*, 185–199.
- Individuals with Disabilities Education Act (1990, 1997, 2004). 20 U.S.C. Chapter 33, Sections 1400–1485. (Statute)
- Kavale, K. A. (2005). Effective intervention for students with specific learning disability: The nature of special education. *Learning Disabilities, 13*, 127–138.
- Kettler, R. J., Rodriguez, M. C., Bolt, D. M., Elliott, S. N., Beddow, P. A., & Kurz, A. (2011). Modified multiple-choice items for alternate assessments: Reliability, difficulty, and differential boost. *Applied Measurement in Education, 24*, 210–234.

- Kicklighter, R. H. (1976). School psychology in the U.S.: A quantitative survey. *Journal of School Psychology, 14*, 151–156.
- Larry P. v. Riles (1972, 1974, 1979, 1984, 1986, 1992). 343 F. Supp. 1306 (N. D. Cal. 1972) (preliminary injunction) *aff'd* 502 F. 2d 963 (9th Cir. 1974); 495 F. Supp. 926 (N. D. Cal. 1979) (decision on merits) *aff'd* (9th Cir. no. 80–427 Jan. 23, 1984). Order modifying judgment, C-71–2270 RFP, September 25, 1986. Memorandum and Order, August 31, 1992. *aff'd* F. 3d#485 (9th Cir. 1994)
- MacMillan, D. L., Gresham, F. M., Lopez, M. F., & Bocian, K. M. (1996). Comparison of students nominated for prereferral interventions by ethnicity and gender. *Journal of Special Education, 30*, 133–151.
- Marshall et al. v. Georgia. (1984, 1985). U.S. District Court for the Southern District of Georgia, CV482–233, June 28, 1984; *aff'd* (11th Cir. No. 84–8771, October 29, 1985). (Appealed as NAACP v. Georgia.) (Note: The court of appeals decision was published as Georgia State Conference of Branches of NAACP v. State of Georgia.)
- Mattie T. et al. v. Johnson et al. Consent decree. U.S. District Court, Northern District of Mississippi, No. DC-75–31-S, December 23, 2003.
- Mayer, G. R., Sulzer-Azaroff, B., & Wallace, M. (2011). *Behavior analysis for lasting change* (2nd ed.). Cornwall-on-Hudson, NY: Sloan.
- McDonnell, L. M., McLaughlin, M. J., & Morison, P. (1997). *Educating one & all: Students with disabilities and standards-based reform*. Washington, DC: National Academy Press.
- Mills v. Board of Education, 348 F. Supp. 866 (D. D. C. 1972).
- National Association of School Psychologists. (2010). *Principles for professional ethics*. Bethesda, MD: Author.
- Parents in Action on Special Education v. Joseph P. Hannon. U.S. District Court, Northern District of Illinois, Eastern Division, No. 74 (3586), July 1980.
- Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2009). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest, 9*, 105–117.
- Patrick, J., & Reschly, D. (1982). Relationship of state educational criteria and demographic variables to school-system prevalence of mental retardation. *American Journal of Mental Deficiency, 86*, 351–360.
- Pennsylvania Association for Retarded Children v. Commonwealth of Pennsylvania, 343 F. Supp. 279 (E. D. Pa. 1972).
- Phillip and Angie C. v. Jefferson County Board of Education, U.S. District Court, Northern District of Alabama, Case 2:07-CV-00756-LSC, August 2011.
- Reschly, D. J. (1996). Identification and assessment of children with disabilities. *Future of Children: Special Education for Children with Disabilities, 6*(1), 40–53.
- Reschly, D. J. (2000). The present and future status of school psychology in the United States. *School Psychology Review, 29*, 507–522.
- Reschly, D. J. (2006). Legal influences on the identification and treatment of educational disabilities. In I. B. Weiner & A. K. Hess (Eds.), *Handbook of forensic psychology* (3rd ed., pp. 167–189.). Hoboken, NJ: Wiley.
- Reschly, D. J., & Bergstrom, M. K. (2009). Response to intervention. In T. B. Gutkin & C. R. Reynolds (Eds.), *The handbook of school psychology* (4th ed., pp. 434–460). Hoboken, NJ: Wiley.
- Reschly, D. J., & Bersoff, D. N. (1999). Law and school psychology. In C. R. Reynolds & T. B. Gutkin (Eds.), *The handbook of school psychology* (3rd ed., pp. 1077–1112). New York, NY: Wiley.

- Reschly, D. J., & Hosp, J. L. (2004). State SLD policies and practices. *Learning Disability Quarterly*, 27, 197–213.
- S-1 v. Turlington, Preliminary Injunction, U.S. District Court, Southern District of Florida, Case No. 79–8020-Civ-CA WPB, June 15, 1979. *aff'd* U.S. Court of Appeals, 5th Circuit, January 26, 1981, 635 F 2d 342 (1981). Trial on Merits, May 19–June 4, 1986. Order on Motion to Dismiss, No. 79–8020-Civ-Atkins, U.S. District Court, Southern District of Florida, October 9, 1986.
- Schalock, R. L., Borthwick-Duffy, S. L., Bradley, V. J., Buntinx, W. H. E., Coulter, D. L., Craig, E. M., . . . Yeager, M. H. (2010). *Intellectual disability: Definition, classification, and system of supports* (11th ed.). Washington, DC: American Association on Intellectual and Developmental Disability.
- Shinn, M. R., & Walker, H. M. (Eds.) (2010). *Interventions for achievement and behavior problems in a three-tier model including RTI* (2nd ed.) Bethesda, MD: National Association of School Psychologists.
- Singer, J. D., Palfrey, J. S., Butler, J. A., & Walker, D. K. (1989). Variation in special education classification across school districts: How does where you live affect what you are labeled? *American Educational Research Journal*, 26, 261–281.
- Skiba, R. J., Eckes, S. E., & Brown, K. (2009/10). African-American disproportionality in school discipline: The divide between best evidence and legal remedy. *New York Law School Review*, 54, 1071–1112.
- Snyder, T. D., and Dillow, S. A. (2011). *Digest of Education Statistics 2010* (NCES 2011–015). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- Thomas, A., & Grimes, J. (Eds.). (2008). *Best practices in school psychology IV* (4th ed.). Bethesda, MD: National Association of School Psychologists.
- Timothy W. v. Rochester, Sch. Dist., IDELR, 559:480 (D.N.H., 1988, Case No. C-84–733-L); IDELR 441:393 (1st Cir., 1989, Case No. 88–1847).
- United States Department of Education. (2001). *To Assure the Free Appropriate Public Education of All Children with Disabilities: Twenty-Second Annual Report to Congress on the Implementation of the Education of the Individuals with Disabilities Education Act*. Washington, DC: Office of Special Education Programs, Author.
- United States Department of Education. (2012). *Determination letters on state implementation of IDEA*. Washington, DC: Author. Retrieved from <http://www2.ed.gov/fund/data/report/idea/partbspap/prevideaetermfs/2012ideafactsheet-determinations.pdf>
- Yell, M. L. (2012). *The law and special education* (3rd ed.). Boston, MA: Pearson.

CHAPTER 9

Assessing Civil Capacities

MICHELE GALIETTA, ALEXANDRA GARCIA-MANSILLA, AND BARBARA STANLEY

THE right to self-determination is a bedrock principle of the American legal system. This right, however, is not unbounded. A state may restrict an individual's right to self-determination in cases where significant harm is likely to come to the individual or others (e.g., civil commitment) or in cases when an individual is determined to lack the capacity to make a particular decision. Whereas children are not assumed to be competent under the law, adults are presumed competent to manage their affairs and make important life decisions, unless some evidence exists to suggest a lack of capacity to make a particular decision. In the past, the mere presence of mental disability or cognitive impairment might have been deemed sufficient evidence that an individual lacked capacity to make certain decisions. Early conceptualizations of competence in the civil realm, particularly in the area of guardianship, were quite global (e.g., decisions about competence applied broadly to any type of decision making). The resulting infringement on personal liberties was far reaching (American Bar Association Commission on Law and Aging/American Psychological Association, 2008; hereinafter ABA/APA, 2008). However, legal, medical, and psychological thinking has evolved considerably. Assessments of competence are now understood to be specific to the requirements of a particular decision or task. Competence is understood to be context-specific, and it is in fact possible to be competent in one area and not another. In areas where the question of competency arises, the law specifies particular elements to be evaluated in making determinations of competence (Perlin, Champine, Dlugacz, & Connell, 2008).

Mental health practitioners are frequently asked to offer opinions to assist in legal decision making regarding the capacities of individuals across a wide range of contexts. The impact of such opinions can be far-reaching and may have serious implications for the civil liberties and rights of the individuals in question. The past 25 years have seen vast changes in the area of assessment of legal competence. The development of clearer theoretical conceptualizations of the construct

of competency provided a firm foundation for research on behavioral components of the legal term *competence* across multiple domains. Mental health professionals working to assist the courts around these issues can rely on this broad body of research to direct their work.

This chapter briefly reviews criticisms of early psychological contributions to the courts in the area of competence. The chapter then reviews current research and practices regarding the evaluation of a variety of competencies in the civil domain, specifically decision making in treatment and research, guardianship, financial capacity, and testamentary capacity, as well as other capacities (e.g., to drive a vehicle). The need for professionals skilled in competency assessments in civil areas is increasing, in large part due to the greater number of older individuals in society.

HISTORY

In 1986, Thomas Grisso wrote a groundbreaking volume on the evaluation of competencies wherein he reviewed the significant criticisms of forensic mental health professionals and their contributions to legal questions of competence (Grisso, 1986). Grisso's summation of the criticisms of forensic practitioners has been colloquially referred to as the "five I's of discontent with forensic practice": ignorance, irrelevance, intrusion, insufficiency, and incredibility (Grisso, 2003, p. 19).

The first criticism stemmed from the observation that practitioners frequently provided testimony that was not directly relevant to legal questions. In fact, Grisso noted that many mental health professionals who conducted forensic evaluations were unaware of the legal questions and issues integral to such proceedings. One explanation was that there is not a simple correspondence between psychological or medical information and legal capacities. In other words, knowing that an individual has early Alzheimer's dementia may indicate that there is reason to question capacity, but it does not directly answer the question of whether the individual possesses the functional abilities to be able to perform a particular task.

Grisso also criticized mental health professionals for often testifying beyond the scope of their expertise and ultimately intruding on the domain of the legal decision maker. The rationale is that legal decisions require a balancing of individual and societal rights. Since such determinations often involve ethical, moral, or societal choices, they fall beyond the purview of clinical science, the domain of the mental health practitioner. In other words, the psychologist in the hypothetical Alzheimer's case just described can, with the proper focus, assess the individual's functional abilities and the likely effect of any deficits on decision-making abilities. However, the determination of if and how the individual's rights should be curtailed is for the court to determine. Such a determination entails a balancing of the individual's personal interests with the possible detrimental effects to the individual if he or she makes a poor decision (e.g., risks associated with a particular treatment) and possible negative effects on others in society (e.g., risks associated with operating a motor

vehicle). Although this balancing falls within the domain of law, determinations of incapacity in medical settings are quite frequently *de facto* determinations made by the treating medical professional or a consulting medical professional (Kolva & Rosenfeld, 2012, p. 22).

Although competence in the legal sense may be a dichotomous determination (i.e., competent or not), from the perspective of mental health practitioners, individuals are not “competent” or “incompetent.” Rather, they possess degrees of various functional abilities relevant to the capacity to make particular decisions or perform certain tasks. It has been noted that these abilities are related in complex ways to decision making (Grisso, 2003). Some have advocated that the degree of functional ability necessary to make decisions should bear some relation to the importance of the decision or to the risk of harm involved in the decision or task, but others disagree (Melton, Petrila, Poythress, & Slobogin, 2007).

Grisso also expressed concern about the lack of scientific basis or data that informed the opinions of mental health professionals in such contexts. Mental health professionals have been criticized for making determinations based on personal experience (e.g., anecdotes), single samples of information, or instruments with unknown reliability or validity.

So, where is the field today with respect to the aforementioned critiques? Practices in the area are dramatically improved. The Ethical Principles of Psychologists (APA, 2002) as well as the Specialty Guidelines for Forensic Psychology (APA, 2013) emphasize that professionals must understand the legal question at issue in each case as well as any relevant legal and professional standards (e.g., guidelines established by their profession). There has been a growing recognition in the field of psychology in general about the importance of the quality of data that form assessment opinions. Additionally, changes in the legal system have increased the demand for valid and reliable data (Groscup, Penrod, Studebaker, Huss, & O’Neil, 2002). There has been an explosion of research in the area, such that the construct of competence has been clarified, recommended procedures for evaluations have been described (e.g., ABA/APA, 2008), and, for many types of competencies, specific valid instruments have been developed. However, it is important to note that, despite these advances, mental health professionals without specific expertise in psycholegal issues continue to engage in some of the unhelpful practices referenced by Grisso, and it is still commonplace to see opinions based on limited information or the use of unreliable or invalid assessment techniques.

CONSTRUCT OF COMPETENCE

The most dramatic contribution to our current understanding of evaluating competencies can be traced back to Grisso’s formulation (2003) of a standardized model for mental health practitioners to assess legal competencies. This model served as the basis for the formulation of a host of empirical questions related to the assessment of capacity. Grisso’s model started from the legal perspective, demanding an

understanding of the law's view of relevant competencies. The model synthesized the legal perspective with the "scientific, empirical and ethical standards of mental health professionals' disciplines" (p. 20). This represented a significant departure from the early work of mental health professionals who focused on assessing symptoms and diagnoses and neglected inquiry into the specific abilities necessary to perform the task at hand (e.g., execute a will, consent to or refuse treatment, enter into a contract) and how such might be limited by underlying impairments.

In brief, Grisso's (2003) model (which has been refined in comparison to what he proposed in 1986) posits that all legal competencies have five components: functional, causal, interactive, judgmental, and dispositional. For any legal capacity, there are related functional abilities necessary to make a decision or to successfully perform the tasks in question. If there are deficits in those functional abilities, the law seeks to understand their causes. The interactive component refers to the fact that competencies are contextual. In other words, particular situations create the necessity for specific functional abilities. For instance, the salience of significant memory impairment may differ for an individual whose health requires consistent compliance with a complicated medication regimen than for a similarly impaired individual without these demands. Finally, judgmental and dispositional aspects of competency proceedings require making direct links (ideally based on observation of functional capacity for relevant tasks) between the individual's capacities and the demands of the situation. For example, in the case just mentioned, the mental health practitioner could describe the impact of memory impairment on the individual's ability to manage his or her finances. Grisso clearly recommends that mental health practitioners refrain from offering ultimate opinions concerning legal questions. More recently, a collaborative working group of the ABA and APA (2008) expanded on this model to create a conceptual framework for assessing the civil capacities in older adults. They identified nine elements inherent in assessing any particular capacity:

1. Legal standard
2. Functional elements
3. Diagnosis
4. Cognitive underpinnings
5. Psychiatric or emotional factors
6. Values
7. Risk considerations
8. Steps to enhance capacity
9. Clinical judgment of capacity

CAPACITY EVALUATIONS

In order to conduct a competency evaluation, an examiner must be knowledgeable about the legal definition and standard (Grisso, 2003; Perlin et al., 2008).

Legal standards are explicated in state statutes and case law precedent in the particular jurisdiction where the issue arises. Evaluators must be familiar with the legal question and associated standards so that they can identify and assess the associated functional abilities, using whatever valid assessment techniques are available (Perlin et al., 2008). Competency assessments must include all of the considerations that apply to other psychological assessments. In other words, testing conditions, the reliability and validity of any measures utilized, cultural considerations, and limitations need to be considered and communicated clearly. Identification of all symptoms/impairments and the effects they have on relevant functional abilities, if any, is an essential element of capacity evaluations. Examiners should describe the causal relationship between specific features of the illness or diagnosis and particular functional impairments relevant to the legal standard.

BARRIERS TO CONDUCTING EVALUATIONS

Significant barriers exist to conducting helpful capacity assessments. Perhaps the most significant barriers stem from temporal fluctuations in ability (e.g., due to fatigue) or other factors (e.g., due to medication). If this is suspected, evaluating the examinee over multiple days and at different times may be important. Testing conditions such as noise, lighting, and privacy also impact evaluations. Particularly with older adults or those with cognitive impairments, fatigue and impression management may be factors (ABA/APA, 2008).

Also having the potential to impede accurate assessment are the examiners' judgments about the quality of an individual's expressed choice (e.g., whether the choice an individual is making is consistent with or valued by others, including the examiner). In fact, the decision-making abilities of individuals may not be challenged if the decisions they are making are consistent with what those around them believe to be in their best interests. In reality, questions of competency and final judgments of incompetence are more likely to emerge for those who make unpopular or unusual choices or choices that are not consistent with those of people around them (e.g., family members, treating health-care professionals, financial advisors) or the examiner (Perlin et al., 2008).

ASSESSMENT TECHNIQUES AND TOOLS

Whenever available, examiners should use assessment techniques with demonstrated validity. Choosing an assessment tool involves consideration of the ability of the instrument to provide construct relevant information. Clinical diagnosis may or may not be associated with impaired capacities, but diagnosis is relevant to most evaluations, because it may provide causal attributions about particular deficits. For example, it would be important to note that, although a person is diagnosed with schizophrenia, his or her decision to refuse a particular medication appears sound and is not related to paranoia or an inability to appreciate implications of the

decision. Similarly, global assessments of cognitive functioning may be helpful but not sufficient to answer the psycholegal question(s) at hand. When available, specific well-validated forensic assessment instruments (i.e., those designed to address specific functional abilities relevant to legal questions or constructs) should be utilized. These are presented in the “Specific Civil Capacities” section.

SPECIFIC CIVIL CAPACITIES

Perlin et al. (2008) describes civil law as pertaining to all aspects of an individual’s private life. Thus, questions of competence beyond those described in this chapter are likely to arise, particularly as technological advances breed new areas of civil law. Consider a physician asked to implant eight embryos into a woman with some mental health problems. Bioethical issues aside, a decision about the appropriateness of this intervention would certainly require consideration and perhaps formal assessment of the woman’s functioning and decision-making abilities. Similarly, clinicians may be asked to assess capacities in the civil realm that are highly specific and relatively new. As a result, individual mental health practitioners may have the expertise to make some competence assessments and not others.

Discussion of the assessment of specific civil capacities included in this section follows a general framework employed in discussing the specific capacities presented: identification of the legal standard, assessment of functional capacities, consideration of most relevant diagnoses and associated symptoms, review of important contextual factors, and reference to valid assessment tools. In cases where other capacities arise, this framework can be readily applied.

Guardianship. Guardianship (referred to as conservatorship in some jurisdictions) allows the state to appoint someone (i.e., a guardian) to exercise the rights of a person who has been adjudicated as lacking such ability. The doctrine of *parens patriae* justifies state intervention, if the basis of the intrusion is to protect a vulnerable individual (Melton et al., 2007). With the exception of children, who are presumed to lack the capacity to make many important decisions on the basis of their age (Perlin et al., 2008), capacity is presumed, and the individual initiating the petition for guardianship must prove that the person is incapacitated (ABA/APA, 2008). In most jurisdictions, any interested party can initiate guardianship proceedings (Perlin et al., 2008), and the standard of proof for such findings is most often clear and convincing in consideration of the potential loss of liberty that may ensue (Kolva & Rosenfeld, 2012).

Many jurisdictions provide for plenary guardianships that allow the guardian to take responsibility for all aspects of an individual’s decision making (i.e., making decisions about finances, health care, residence, and legal matters) as well as limited guardianships, which are more respectful of autonomy and provide a more narrow scope for the surrogate decision maker (Perlin et al., 2008). It should be

noted that the rights at risk in guardianship proceedings vary widely. Some states are quite specific detailing numerous rights at stake in such proceedings (Florida Statutes Chapter 744). Some statutes provide qualifications that the guardianship be necessary or use less restrictive alternative language (ABA/APA, 2008). However, some have noted that plenary guardianship arrangements remain more common and have described barriers to adoption of limited guardianship in practice (Melton et al., 2007). However, guardianship proceedings have been criticized for their lack of procedural protections. For instance, although notification to individuals is required, they may not have the right to counsel, to jury trial, or in some cases to be present at the hearing (Melton et al., 2007). Drogin and Barrett (2010) have noted additional problems with the oversight and monitoring, including a need for guardian training, procedures to assess the veracity of guardian reports, and lack of adequate funding for oversight functions.

The legal standards for incapacity vary by jurisdiction. Historically, states typically required the presence of a disabling condition and evidence that the condition caused inability to adequately manage one's affairs (ABA/APA, 2008). Although some jurisdictions continue to maintain these very gross standards (Melton et al., 2007), in recent years, the standards have generally moved toward more specific functional definitions that include combinations of these factors: disabling condition, functional behavior, and cognitive functioning (ABA/APA, 2008, p. 21).

The functional abilities required to manage all aspects of one's affairs comprise the broadest range of skills; thus, professionals may need to utilize both broad and specific measures of functioning. Some types of capacities to be assessed include activities of daily living (i.e., eating, dressing) and more specific complex activities (i.e., managing household responsibilities).

Financial Capacity. In cases in which a person lacks capacity to manage his or her financial affairs, a guardian, guardian ad litem, or conservator may be appointed. Given the intrusion into liberty that is involved, the standard of proof is quite high, namely, clear and convincing evidence. As in other areas of civil law, jurisdictions vary considerably in terms of legal standards for financial incapacity. The *Assessment of Older Adults with Diminished Capacity: A Handbook for Psychologists* (ABA/APA, 2008) states that, in contrast to vague standards,

A better and far more specific criterion is set forth in Section 410(2) of the Uniform Guardianship and Protective proceedings Act (UGPPA), which states that a court may appoint a conservator if the court determines that "the individual is unable to manage property and business affairs because of an inability to receive and evaluate information or make decisions, even with use of appropriate technical assistance *and* the individual has property that will be wasted or dissipated unless management is provided or money is needed for the support, care, education, health and welfare of the individual or of individuals who are entitled to the individual's support." (p. 72)

Specific diagnoses that may be associated with impairment in essential functioning in this area include various dementias, Parkinson's disease, cerebrovascular accidents, traumatic brain injury, mental retardation, and autism spectrum disorders, as well as mental disorders such as schizophrenia, bipolar disorder, and substance abuse disorders and cognitive losses associated with normal aging.

Functional abilities associated with financial capacity include performance-type skills, such as the ability to count money, read and interpret financial statements, write checks, and execute other financial transactions, as well as more abstract abilities requiring more complicated judgment (e.g., weighing risks in the case of business endeavors and behaving consistently with personal values about financial decisions; Griffith et al., 2003). The majority of research in this area has been undertaken by Daniel Marson and his working group, who have defined financial capacity as the ability to manage money and financial assets in ways that meet a person's needs and is consistent with his or her values and self-interest (Marson, Tribel, & Knight, 2012).

The ABA/APA working group (2008) specified structured questions that can be added to a clinical interview; these include "What is your financial history? Are you in any debt? . . . In spending money, what are your highest priorities? . . . Is there anyone you specifically would *not* want to be involved in helping to make financial decisions on your behalf?" (p. 77). Additionally, there are several global performance-based measures of activities of daily living that include specific items that tap financial capacity (e.g., the Adult Functional Adaptive Behavior Scale [AFABS], Pierce, 1989; Direct Assessment of Functional Status [DAFS], Lowenstein et al., 1989; Independent Living Scales [ILS], Loeb, 1996). However, the best-developed assessment instrument is the Financial Capacity Instrument (FCI; Marson et al., 2000). The FCI has a standard protocol for administration and scoring and assesses functional abilities in six domains:

1. Basic monetary skills (e.g., counting money and making change)
2. Financial conceptual knowledge (e.g., defining financial concepts)
3. Cash transactions (e.g., tipping)
4. Checkbook management
5. Bank statement management
6. Financial judgment (e.g., identifying risk for mail or telephone fraud)

The FCI differentiates financial management abilities in individuals with varying degrees of cognitive impairment (Marson et al., 2009).

Forensic evaluators need to consider important contextual factors when conducting financial capacity assessments. For example, it is important for the evaluator to gather information about an individual's lifelong financial habits and values in order to determine if current financial behaviors are a departure from the individual's baseline (e.g., some individuals have never kept good financial records, while others had meticulous records and an indication of a change may indicate

deterioration of functioning). In addition, information about the financial demands on an individual (e.g., the complexity of his or her daily financial activities and asset portfolio) is important in terms of knowing the level of ability individuals must possess in order to manage their affairs. It may be possible to enhance capacity with aids or supports. It is important to describe not only the examinee's deficiencies but the abilities that remain intact. Such descriptions of capacity can assist in legal judgments and may influence the scope of legal intervention (e.g., plenary versus limited guardianship in states where discretion is possible).

Testamentary Capacity. The capacity of an individual to create or change a will is known as testamentary capacity. In general, in order for a will to be considered valid, one must at the time of the will's creation understand the nature and purpose of a will, who the natural heirs are, and the nature and extent of one's assets. It should be noted, however, that the relative weight given to the elements just described may vary in different jurisdictions (Marson & Hebert, 2008; Perlin et al., 2008).

Disorders that are likely to influence cognitive capacities in this area include various dementias and related disorders with memory and executive functioning deficits, Parkinson's disease, traumatic brain injuries, and developmental disorders. Major mental disorders, such as schizophrenia, delusional disorder, and bipolar disorder, can interfere with capacity if they impair rational decision making (e.g., an individual harboring a delusion that a family member is trying to poison him or her).

Although clinicians may be asked to assess testamentary capacity contemporaneously, it is more common for questions about testamentary capacity to arise after the individual is deceased. The role for mental health professionals may be more challenging when the individual is deceased, but often a great deal may be gleaned from gathering available information about the person at the time the will was executed. These "psychological autopsies" typically involve review of medical records, financial records, and multiple sources of collateral information (e.g., depositions of persons familiar with the testator's functioning at and around the time the will was authored; other writings completed by the testator at and around the time the will was executed). Helpful guidelines for retrospective evaluations of this type can be found in *Assessment of Older Adults with Diminished Capacity: A Handbook for Psychologists* (ABA/APA, 2008, pp. 87–88) and in practice-oriented texts written by others (Drogin & Barrett, 2010; Mart & Alban, 2010). One challenge is that, if a person has fluctuating periods of lucidity, a will is considered valid if it was executed while the person was rational (Perlin et al., 2008).

Another issue that commonly arises after the testator is deceased is the question of undue influence, which can be understood as the use of power to deceptively gain control over the decision making of the individual (Marson, Hithwaite, & Hebert, 2004; Marson, Triebel, & Knight, 2012). Various models have been presented to

describe the process by which strangers, caregivers, or family members exert undue influence over an individual. Most are similar and include creating or maintaining isolation of the individual, creating a dependence on the coercive individual, and using various means to manipulate the individual and the loss of assets by the individual. When many of these elements appear to have been present, the will may be deemed invalid even if the person possesses the requisite cognitive abilities. Although there are no formal assessment instruments published for testamentary capacity, authors have created structured worksheets for assessments in this area (see ABA/APA, 2008). It should be noted that the threshold for competence is generally held to be fairly low, so long as no duress or undue influence was present. Drogin and Barrett (2010) advise careful examination of multiple sources of information in order to assess for the presence of undue influence.

Capacity to Accept or Refuse Treatment or Participate in Research Studies. Another important area of civil law has to do with the right to accept or decline medical treatments or the opportunity to participate in medical, psychosocial, or other types of research. Both the capacity to make treatment decisions and the capacity to make decisions about participation in research studies are rooted in the process of informed consent. Understanding the requirements of informed consent is essential in order to understand the assessment of capacity in this area.

Doctrine of Informed Consent. The doctrine of informed consent is derived from two sources. Informed consent to treatment has primarily been influenced by litigation and subsequent case law. The foundation of this body of law lies in common law, which subjects physicians to tort liability in cases where unauthorized medical procedures have occurred (Ludlum, 1972; Rovosky, 1990).

In contrast, the doctrine of informed consent to research evolved through a series of policies, regulations, and professional codes. The first guidelines and standards for human experimentation were developed in response to the atrocities perpetrated by Nazi experimenters during World War II (National Research Council [NRC] Committee on National Statistics, 2003, p. 60). They established the requirement that research participants must provide voluntary, informed consent. In 1964, the Declaration of Helsinki (World Medical Assembly, 1964, 1975) expanded on the guidelines, emphasizing the need for fully informing research participants of the “aims, methods, anticipated benefits and potential hazards of the study and the discomfort it may entail.” However, unethical and ethically troubling practices were apparently fairly commonplace at the time, even in prestigious research institutions (Beecher, 1966).

Concern over such practices prompted the U.S. Public Health Service to create a policy on human research participant protection in 1966 (NRC, 2003). Problems with the implementation of the 1966 policy and media coverage of several instances of abuses of participants in biomedical research led to the

development of additional federal guidelines and regulations protecting human research participants (Department of Health, Education and Welfare [DHEW], 1973; Department of Health and Human Services [DHHS], 1981). Some of the abuses stemmed from the fact that individuals were either unaware that they were research subjects or were only partially informed about the experiment in which they were participating (e.g., experiments at the Jewish Chronic Disease Hospital; the Tuskegee Syphilis experiments, Levine, 1981; and the Tearoom Trade Studies, Humphreys, 1970).

The 1970s saw the creation of additional protections (e.g., Institutional Guide to DHEW Policy on Protection of Human Subjects [DHEW, 1973]) and the advent of ethical guidelines for research by professional organizations (e.g., APA, 1977). Another important development at this time was the publication of the Belmont Report (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979), which established the philosophical basis for the ethical treatment of individuals and outlined three essential principles to be considered in conducting research: respect for persons, beneficence, and justice. This report has become the foundation for contemporary institutional review boards (IRBs) and human subject protections.

Government regulation of research has continued to grow and become increasingly more expansive and complex. In 2002, the Office for Human Research Protections (OHRP) was established and charged with oversight of IRBs as well as providing education and guidance on human research participant protection to both federal and nonfederal bodies and institutions (NRC, 2003). At the time of this writing, the OHRP is currently accepting public comment on revisions to the Common Code in order to further enhance the protections of research participants and to improve policies and procedures for investigators.

Informed Consent Defined. Although the informed consent doctrine has evolved over several decades and from a variety of sources, there is consensus on the three major elements of the doctrine: adequate disclosure of information, voluntariness, and competency (Meisel, Roth, & Lidz, 1977; Melton et al., 2007; Roberts et al., 2002). Both case law developments and expanded federal regulations have resulted in increased scrutiny of treatment providers and researchers with regard to the presence and adequacy of informed consent procedures. Although empirical studies have indicated that informed consent in the real world is often far from optimal, it is generally agreed that informed consent must be solicited for all treatments and for most research studies. Consent is not required in some studies (e.g., studies using archival data or ethically permissible studies requiring the use of deception). Written documentation that informed consent has occurred is standard.

Adequate Disclosure. Patients and research participants require sufficient information in order to make a choice about course of treatment or whether to participate in research (Culver & Gert, 1982; Stanley & Stanley, 1981). Federal regulations

indicate that, at minimum, disclosure to research participants should include the following:

1. A statement that the study involves research and a description of the purposes, procedures, and duration of the research
2. A description of the foreseeable risks and discomforts
3. Explanation of potential benefits to the participant
4. Available alternatives
5. Extent of confidentiality
6. Extent of compensation and treatment for injuries
7. Name of an appropriate contact person in case questions or problems arise
8. A statement that participation is voluntary and that participants may withdraw at any time without penalty or loss of benefits to which they are otherwise entitled (NRC, 2003, p. 82)

In the treatment setting, usually the first four items (i.e., purpose of treatment and procedures, risks, benefits, and alternatives) are most salient.

Descriptions and documents must explain the study or treatment in simple, nontechnical language. Additionally, participants must have individuals available to answer questions as well as to assess whether the individual comprehends the information presented adequately enough to make a reasoned decision.

Voluntariness. Following adequate disclosure, the second element of informed consent is voluntariness. It requires that the individual be free from coercive influences and undue pressure in deciding whether to participate in treatment or research (Culver & Gert, 1982; Meisel et al., 1977; Roberts et al., 2002). Potential for coercion may be minimized by following procedures such as allowing a certain amount of time to elapse between the doctor–patient discussion and the patient’s actual decision; giving patients the opportunity to temporarily remove themselves from the environment that may be exerting influence; and, for research, making available a professional not directly involved in the research project for participant consultation (President’s Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, 1982).

Competency. As in all civil competencies discussed in this chapter, competency to consent to treatment or research is specific to a particular decision. Although there is no universal legal standard for competence to consent to treatment or research, most case law includes these standards:

- The ability to communicate a choice
- The ability to understand relevant information
- The ability to provide rational reasons for one’s decision

- The ability to appreciate the situation and probable consequences associated with various options
- The ability to manipulate relevant information in a rational manner (Appelbaum & Grisso, 1988)

Each capacity decision is embedded in a unique context and involves transaction between situational demands and the individual abilities of the decision maker. It is important to note that, in some cases, the functional abilities associated with standards of competence fluctuate over time. Thus, assessments are necessarily temporal in nature.

In addition to conditions with potential to impede capacity, such as dementias, developmental disorders, and major mental disorders such as schizophrenia or bipolar disorder, certain classes of individuals may be at greater risk for deficits in functional abilities—related capacity (e.g., elderly adults, children and adolescents).

Health-Care Proxies. In cases where capacity is impaired, individuals may appoint a health-care proxy. A health-care proxy (also known as a power of attorney) allows a competent individual to appoint a person to make health-care decisions on his or her behalf, should the person become incapacitated at any point in the future. As noted, the individual must be competent at the time the proxy is established. However, the capacity to appoint someone to make treatment decisions on one's behalf generally involves knowing the person who will be the agent and expressing the desire to have that person represent one's health-care interests. The capacity to make health-care decisions for oneself requires a host of other abilities, including the ability to appreciate information and weigh the potential risks and benefits of various treatment options. Thus, even individuals without capacity to make treatment decisions may be considered competent to execute a health-care proxy.

In addition, prospective procedures, such as living wills (i.e., personal directives, advance directives), allow individuals to dictate specific instructions regarding their health-care wishes, should they become incapacitated (i.e., to forgo life-sustaining treatments under specific circumstances; Emmanuel & Emmanuel, 1989). Some have noted that it is often problematic to use abstract scenarios to make decisions. Individuals who attempt to make decisions based on their family members' responses to hypothetical scenarios are frequently confused about what they would wish under different (real-life) circumstances (Annas, 1991; Brett, 1991). In response to these criticisms, Levi and Green (2010) developed a computer-assisted decision aid that educates individuals about advance directives, assists them in prioritizing decision-making factors, and incorporates an individual's values and goals into meaningful advance directive instructions. Although such measures optimize the likelihood that the patient's wishes will be followed, one criticism is that frequently there

is no clear assessment, notification, or documentation for when such procedures commence, despite the fact that they should become active only in the event a person is determined to be incapacitated.

Instruments to Assess Capacity to Consent to Treatment. For many years, capacity was determined by unstructured clinical judgment. Such judgments suffered from poor reliability (Marson, Ingram, Cody, & Harrell, 1995a). The sole instrument used in the assessment of early decision-making capacity was the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975), a brief screening measure for cognitive impairment. However, studies have found very weak relationships between MMSE scores and measures of decision-making ability (Barton, Mallik, & Orr, 1996; Etchells et al., 1997; Janofsky, McCarthy, & Folstein, 1992).

As understanding of the construct of competence evolved, formal instruments were developed to assess some or all of the standards of competency. Early instruments suffered from limitations such as assessing only a limited number of the legal standards or assessing general information rather than information specific to the actual decision the person may be facing (i.e., neglecting context).

Although numerous instruments have been developed for use in this context, normative data are limited or absent for many of them. The three instruments described next were selected because they evaluate each of the four standards: the ability to communicate a choice, understanding, the ability to appreciate information in relation to their case, and the ability to reason (i.e., rationally compare alternatives and their consequences). Additionally, each instrument has favorable psychometric properties and has considerable normative data.

The MacArthur Competence Assessment Tool, Treatment (MacCAT-T; Grisso & Appelbaum, 1998) assesses decision making for a particular decision rather than general informed consent issues. The MacCAT-T is administered in a semistructured interview that is tailored to the specific decision the patient faces, which allows the instrument to be used for an unlimited range of decisions. Because the MacCAT-T is rooted in Grisso and Appelbaum's theoretical conceptualization of competence, the authors allowed for the possibility that deficits in one domain of decision-making capacity may lead one to consider a patient incompetent even if other domains of cognitive functioning are intact, and vice versa. Hence, rather than generating a total score, the MacCAT-T yields separate ratings for each of four aspects of the patient's decision making: ability to express a choice, ability to understand treatment-related information, appreciation of the significance of the information for the particular situation, and ability to rationally compare alternatives in light of their relative consequences.

Marson and his colleagues designed the Capacity to Consent to Treatment Instrument (CCTI) to assess decision-making capacity in older adults (Marson et al., 1995a, 1995b). The CCTI utilizes two hypothetical vignettes, one related to neoplasm (i.e., cancer) and the second describing cardiac problems. It uses patient responses to questions about these vignettes to rate patients on five legal standards considered

to reflect an ascending level of difficulty. The five standards assessed by the CCTI are:

1. Ability to make a treatment choice
2. Ability to make a reasonable choice
3. Ability to appreciate the consequences involved with various choices
4. Ability to provide rational reasons for choosing a particular decision
5. Ability to understand the treatment situation and choices

This last standard requires synthesis of treatment information as well as a fairly sophisticated, contextualized understanding of each treatment choice.

Moye and colleagues (2008) developed the Assessment of Capacity to Consent to Treatment (ACCT) for use with individuals with neurocognitive or neuropsychiatric deficits. The measure was designed to be relatively brief. The instrument begins with structured interview questions to elicit patient values and preferences relevant to treatment decisions. Three hypothetical vignettes are utilized to assess capacities in understanding appreciation, reasoning, and communicating a choice. The instrument has a subscale for each of the aforementioned abilities. Subscales range from 2 to 13 items. The scale has good internal consistency and interrater reliability. Since development, the ACCT interview has been normed on a number of samples.

Instruments to Assess Capacity to Consent to Participation in Research. A version of the MacCAT-T exists to assess the capacity to consent to clinical research (the MacCAT-CR; Applebaum & Grisso, 2001). This measure utilizes the same model as the MacCAT-T, enabling the investigator to customize the information disclosed to fit the study characteristics. Other research groups have offered their own such instruments, but these measures appear to be used less frequently (Joffe, Cook, Cleary, Clark, & Weeks, 2001; Miller, O'Donnell, Searight, & Barbarash, 1996).

CONCLUSIONS

Capacity evaluations in the civil realm are increasing, in large part due to the aging of the American population. This chapter has laid out a basic conceptual framework with which to approach civil capacity evaluations. As noted, the focus of evaluations varies considerably, depending on the legal question and its domain within civil law. Further, jurisdiction necessarily influences civil capacity evaluations substantially. Forensic mental health professionals working in the area are encouraged to attend carefully to such variations. The field continues to change, and there is much need for continued empirical work in this area. In particular, developing empirically based evaluation tools in areas where none exist is important. Additionally, studying procedures that might aid individuals who exhibit marginal competence or prolong capacities in older adults should be an area of increased attention.

REFERENCES

- American Bar Association Commission on Law and Aging/American Psychological Association. (2008). *Assessment of older adults with diminished capacity: A handbook for psychologists*. Washington, DC: Author.
- American Psychological Association. (1977, March). Ethical standards of psychologists. *APA Monitor*, 22–23.
- American Psychological Association. (2002). *Ethical principles of psychologists and code of conduct*. Washington, DC: Author.
- American Psychological Association. (2013). Specialty guidelines for forensic psychology. *American Psychologist*, 68, 7–19. doi:10.1037/a0029889
- Annas, G. J. (1991). The health care proxy and the living will. *New England Journal of Medicine*, 324, 1210–1213.
- Applebaum, P. S., & Grisso, T. (1988). Assessing patients' capacities to consent to treatment. *New England Journal of Medicine*, 319, 315–336.
- Appelbaum, P. S., & Grisso, T. (2001). *MacCAT-CR: MacArthur Competence Assessment Tool for Clinical Research*. Sarasota, FL: Professional Resource Press.
- Barton, C. D., Mallik, H. S., & Orr, W. B. (1996). Clinicians' judgment of capacity of nursing home patients to give informed consent. *Psychiatric Services*, 47(9), 956–960.
- Beecher, H. K. (1966). Ethics and clinical research. *New England Journal of Medicine*, 274, 1354–1360.
- Brett, A. S. (1991). Limitations of listing specific medical interventions in advance directives. *Journal of the American Medical Association*, 266, 825–828.
- Culver, C. M., & Gert, B. (1982). *Philosophy in medicine: Conceptual and ethical issues in medicine and psychiatry*. New York, NY: Oxford University Press.
- Declaration of Helsinki: Recommendations guiding medical doctors in biomedical research involving human subjects. Adopted by the 18th World Medical Assembly, Helsinki, Finland, 1964, and as revised by the 29th World Medical Assembly, Tokyo, 1975.
- Department of Health, Education, and Welfare, & the National Institutes of Health. (1973). Protection of human subjects: Policies and procedures. *Federal Register* 38(221), 31, 738–749.
- Department of Health and Human Services. (1981). Final regulations amending basic HHS policy for the protection of human research subjects. *Federal Register*, 46(16), 8366–8391.
- Drogin, E. Y., & Barrett, C. L. (2010). *Evaluation for guardianship*. New York, NY: Oxford University Press.
- Emmanuel, L. L., & Emmanuel, E. (1989). The medical directive: A new comprehensive advance care document. *Journal of the American Medical Association*, 261(22), 3288–3293.
- Etchells, E., Katz, M. R., Shuchman, M., Wong, G., Workman, S., Choudhry, N. K., . . . Singer, P. A. (1997). Accuracy of clinical impressions and Mini-Mental-State Exam scores for assessing capacity to consent to major medical treatment: Comparison with criterion-standard psychiatric assessments. *Psychosomatics*, 38(3), 239–245.
- Florida Statutes Chapter 744.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198.

- Griffith, H. R., Belue, K., Sicola, A., Krzywanski, S., Zamrini, E., Harrell, L., & Marson, D. C. (2003). Impaired financial abilities in mild cognitive impairment: A direct assessment approach. *Neurology*, *60*, 449–457.
- Grisso, T. (1986). *Evaluating competencies: Forensic assessments and instruments*. New York, NY: Plenum Press.
- Grisso, T. (2003). *Evaluating competencies: Forensic assessments and instruments* (2nd ed.). New York, NY: Kluwer Academic/Plenum Press.
- Grisso, T., & Appelbaum, P. S. (1998). *MacArthur Competence Assessment Tool for Treatment (MacCAT-T)*. Sarasota, FL: Professional Resource Press.
- Groscup, J. L., Penrod, S. D., Studebaker, C. A., Huss, M. T., & O'Neil, K. M. (2002). The effects of *Daubert* on the admissibility of expert testimony in state and federal criminal cases. *Psychology, Public Policy, and Law*, *8*(4), 339–372.
- Humphreys, L. (1970). *Tearoom trade: Impersonal sex in public places*. Chicago, IL: Aldine.
- Janofsky, J. S., McCarthy, R. T., & Folstein, M. F. (1992). The Hopkins Competency Assessment Test: A brief method for evaluating patients' capacity to give informed consent. *Hospital and Community Psychiatry*, *43*(2), 132–136.
- Joffe, S., Cook, E. F., Cleary, P. D., Clark, J. W., & Weeks, J. C. (2001). Quality of informed consent: A new measure of understanding among research subjects. *Journal of the National Cancer Institute*, *93*(2), 139–147.
- Kolva, E. A., & Rosenfeld, B. (2012). Legal perspectives on civil capacity and competence. In G. J. Demakis (Ed.), *Civil capacities in clinical neuropsychology* (pp. 17–36). New York, NY: Oxford University Press.
- Levi, B. H., & Green, M. J. (2010). Too soon to give up: Re-examining the value of advance directives. *American Journal of Bioethics*, *10*, 3–22.
- Levine, R. J. (1981). *Ethics and regulation of clinical research*. Baltimore, MD: Urban & Schwarzenberg.
- Loeb, P. A. (1996). *Independent Living Scales*. San Antonio, TX: Psychological Corporation.
- Lowenstein, D. A., Amigo, E., Duara, R., Guterman, A., Hurwitz, D., Berkowitz, N., . . . Gittelman, B. (1989). A new scale for the assessment of functional status in Alzheimer's disease and related disorders. *Journal of Gerontology, Psychological Science*, *44*, 114–121.
- Ludlum, J. (1972). *Informed consent*. Chicago, IL: American Hospital Association.
- Mahew, M. (2005). Survey of state guardianship laws: Statutory provisions for clinical evaluations. *Bifocal, Journal of the ABA Commission on Law and Aging*, *26*, 1–19.
- Marson, D. C., & Hebert, T. (2008). Testamentary capacity. In B. L. Cutler (Ed.), *Encyclopedia of psychology and the law* (Vol. 2, pp. 798–801). Thousand Oaks, CA: Sage.
- Marson, D. C., Hithwaite, J., & Hebert, K. (2004). Testamentary capacity and undue influence in the elderly: A jurisprudential therapy perspective. *Law and Psychology Review*, *28*, 71–96.
- Marson, D. C., Ingram, K. K., Cody, H. A., & Harrell, L. E. (1995a). Assessing the competency of patients with Alzheimer's disease under different legal standards. A prototype instrument. *Archives of Neurology*, *52*(10), 949–954.
- Marson, D. C., Ingram, K. K., Cody, H. A., & Harrell, L. E. (1995b). Neuropsychologic predictors of competency in Alzheimer's disease using a rational reasons legal standard. *Archives of Neurology*, *52*(10), 955–959.
- Marson, D. C., Sawrie, S., McInturff, B., Snyder, S., Chatterjee, A., Stalvey, T., . . . Harrell, L. (2000). Assessment of financial capacity in patients with Alzheimer's disease: A prototype instrument. *Archives of Neurology*, *57*, 877–884.

- Marson, D. C., Triebel, K., & Knight, A. (2012). Financial capacity. In G. J. Demakis (Ed.), *Civil capacities in clinical neuropsychology* (pp. 39–68). New York, NY: Oxford University Press.
- Marson, D. C., Martin, R. C., Wadley, V., Griffith, H. R., Snyder, S., Goode, P. S., . . . Harrell, L. E. (2009). Clinical interview assessment of financial capacity in older adults with mild cognitive impairment and Alzheimer's disease. *Journal of the American Geriatrics Society*, 57(5), 806–814.
- Mart, E. G., & Alban, A. D. (2010). *The practical assessment of testamentary capacity and undue influence in the elderly*. Sarasota, FL: Professional Resource Exchange.
- Meisel, A., Roth, L., & Lidz, C. (1977). Toward a model of the legal doctrine of informed consent. *American Journal of Psychiatry*, 134(3), 285–289.
- Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C. (2007). *Psychological evaluations for the courts* (3rd ed.). New York, NY: Guilford Press.
- Miller, C. K., O'Donnell, D. C., Searight, H. R., & Barbarash, R. A. (1996). The Deaconess Informed Consent Comprehension Test: An assessment tool for clinical research subjects. *Pharmacotherapy*, 16(5), 872–878.
- Moye, J., Karel, M. J., Edelstein, B., Hicken, B., Armesto, J. C., & Gurrera, R. J. (2008). Assessment of the capacity to consent to treatment: Current research, the "ACCT" approach, future directions. *Clinical Gerontologist*, 31(3), 37–66.
- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research*. Retrieved from <http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html>.
- National Research Council Committee on National Statistics. (2003). *Protecting participants and facilitating social and behavioral science research*. Washington, DC: National Academies Press.
- Pierce, P. S. (1989). *Adult Functional Adaptive Behavior Scale: Manual of directions* (rev. ed.). Togus, ME: Author.
- Perlin, M. L., Champine, P., Dlugacz, H. A., & Connell, M. (2008). *Competence in the law: From legal theory to clinical application*. Hoboken, NJ: Wiley.
- President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. (1982). *Making health care decisions*. Washington, DC: U.S. Government Printing Office.
- Roberts, L. W., Warner, T. D., Brody, J. L., Roberts, B., Laureillo, J., & Lyketsos, C. (2002). Patient and psychiatrist ratings of hypothetical schizophrenia research protocols: Assessment of harm potential and factors influencing participation decisions. *American Journal of Psychiatry*, 159(4), 573–584.
- Rovosky, F. A. (1990). *Consent to treatment: A practical guide* (2nd ed.). Boston, MA: Little, Brown.
- Stanley, B., & Stanley, M. (1981). Psychiatric patients and research: Protecting their autonomy. *Comprehensive Psychiatry*, 22(4), 420–427.

Conducting Child Abuse and Neglect Evaluations

LOIS O. CONDIE

IN 2010, U.S. state and local child protective service workers received an estimated 3.3 million reports of child maltreatment (abuse and neglect), at a rate of 43.8 per 1,000. Child Protective Service officials estimated that 695,000 children (9.2 per 1,000) were victims of maltreatment. The total lifetime economic burden resulting from new cases of fatal and nonfatal child maltreatment in the United States is about \$124 billion (U.S. Department of Health and Human Services, Administration for Children and Families, 2011). In 2008 dollars, the average lifetime cost for each surviving victim of child maltreatment was \$210,012, compared to a lifetime cost for stroke of \$159,846 and between \$181,000 and \$253,000 for type 2 diabetes. Child maltreatment is a serious and prevalent public health problem (Russo, 2008; Wang & Holton, 2007).

In order to conduct an evaluation of children and parents in cases of alleged child maltreatment, one needs an in-depth understanding of the etiology and impact of child maltreatment. Early research highlighted linkages between hypothesized predictive factors and maltreatment outcomes. Main effect linear models of predictive risk factors quickly fell away in the research community in favor of complex, multilevel (culture, community, family), multidirectional, multifactorial, and interactive etiological models (Alink, Cicchetti, Kim, & Rogosch, 2012; Cicchetti, 2004). Static and dynamic causal agents are hypothesized to explain the occurrence of child maltreatment. Some causes are relatively ingrained over the course of history (e.g., societal attitudes toward family privacy and autonomy, individual and collective notions of responsibility for the well-being of children, prevailing societal notions of what constitutes child maltreatment and when it is appropriate to intervene); some are situational (impoverishment, geographic access to services, neighborhood quality); some are sociocultural (family structure, acceptance and expectations of youthful parenting, definitions and tolerance of violence); and

some are individual attributes of parents and children (parental personality style, expressions of social dominance, physiological consequences of repeated exposure to stress). Increasingly, researchers are examining biological and genetic aspects of child maltreatment perpetration and victim traumatic responses (Alia-Klein et al., 2009; Alink et al., 2012; Simon et al., 2012). Definitional issues remain a challenge, particularly because maltreatment types often are defined by legal codes or social service systems and because researchers have not reached consensus on some of the nuances of definitions (see Condie, 2003; Marshall, 2012, for examples). There is, however, growing consensus over research definitions and subtypes (Cicchetti, 2004; Runyan et al., 2005).

Researchers studying risk factors, protective factors, and intervention methods have concluded that it is difficult to isolate one form of maltreatment from another in order to adequately classify or study factors that might be specific to one form of child maltreatment (i.e., physical abuse, sexual abuse, emotional/psychological abuse, neglect). Because different forms of child maltreatment tend to co-occur, pure scientific analysis of contributing factors is difficult (Cicchetti, 2004; Marshall, 2012). As a result, researchers have instead focused on maltreatment typology overlap and comorbidity, the degree and nature of maltreatment, and child protection policy analysis (Brandon, 2001; Marshall, 2012). A second problem affecting the integrity of scientific research is the degree to which social or legal definitions of child abuse meaningfully correspond to real behavior. Classification entries in state records of child maltreatment typically are recorded after negotiation and consultation with families, representatives of the justice system, and representatives of child protective systems (Bae, Solomon, Gelles, & White, 2010; Putnam-Hornstein, Webster, Needed, & Magruder, 2011). With the exception of large-scale funded projects, research samples typically are drawn from small convenience samples. Thus, there are a variety of challenges to research on the etiology and impact of child maltreatment. They are briefly mentioned here to alert evaluators to the inherent limitations in the state of the science. Researchers studying child maltreatment acknowledge the methodological difficulties; unfortunately, the difficulties are not easily overcome (MacMillan, 2005).

Theories of child maltreatment (see Belsky, 1993; Condie, 2003) include these models:

- Psychological (e.g., personality variables, emotional variables, characteristics of perpetrators)
- Sociological (societal and contextual conditions giving rise to child maltreatment)
- Criminological (social class variables, rational choice theory, self-interest motives, communal relationships, strain theory)
- Interactional (dyadic parent-child goodness of fit, communal relations)
- Genetic (epigenetics, gene/environment interactions and correlations)

No model has emerged that fully explains child maltreatment or less severe forms of problematic parenting (Belsky, 1993; Runyan et al., 2005; Simon et al., 2012). Child maltreatment, in any of its forms, is multiply influenced by a variety of determinants that coalesce through transactional processes at various levels of analysis (life course, immediate-situational, stressors-support, potentiating-protective, historical-evolutionary) in the broad context of parent-child or other caregiver-child relationships (MacMillan, 2005). No single pathognomonic factor or unique set of conditions is reliably predictive of child maltreatment. There do seem to be commonly recurring protective factors, most notably in the social domain of functioning, but those factors do not reliably protect children from maltreatment in all cases (Afifi & MacMillan, 2011; Belsky, 1993; Corse, Schmid, & Trickett, 1990; Li, Godinet, & Arnsberger, 2011). When potentiating microsocial and macrosocial factors outweigh supports and compensatory strategies, the probability of child maltreatment increases (Belsky, 1993; Cicchetti, 2004; Howes, Cicchetti, Toth, & Rogosch, 2000; Runyan et al., 2005). There are many pathways to child maltreatment.

Similarly, there is no single or uniform solution to the problem of child maltreatment. Interventions range from preventive to clinical, self-help to formal intervention, individual to macrosocial, and psychological to legal. Policies within child protective service systems range from emphasis on termination of parental rights to emphasis on family preservation strategies, and sometimes those goals take place concurrently. The targets of intervention might include a specific parent, a set of parents with common struggles, a specific child or set of children from the same family, children from similar maltreatment environments, or the neighborhood and social conditions contributing to child maltreatment risk.

RISK OF CHILD MALTREATMENT

Risk factors are factors that increase the odds that child maltreatment will occur. Because child maltreatment has a major economic and social impact, early detection is of great importance. Researchers have developed taxometric structures for predicting child maltreatment. Risk variables and mediators of risk tend to fall under the categories of developmental and psychological factors, social and community variables, and contextual variables.

DEVELOPMENTAL AND PSYCHOLOGICAL FACTORS

Evaluators benefit from a comprehensive understanding of factors contributing to and mediating risk of child maltreatment. Although developmental and psychological risk factors contribute to and mediate risk of child maltreatment, determining the potency of risk factors or combinations of risk factors is by no means straightforward. Child maltreatment has been associated with a host of developmental

and psychological variables that include examples such as parents' own histories of maltreatment, mental health and personality variables, and substance use. Developmental and psychological protective factors associated with diminished maltreatment risk include access to parental partners (whether formal or informal), social support, neighborhood cohesiveness, and access to child care (Belsky & Jaffee, 2006).

Parental Factors. Researchers studying parental factors that raise the risk of child maltreatment have focused on childhood histories of abusive and neglectful parents, personality variables, social-emotional variables, and psychological resources. For example, intergenerational researchers have focused on harsh parenting and warm-supportive parenting (Belsky, Conger, & Capaldi, 2009). Traditionally, researchers focused on maltreatment by mothers, but there is a trend toward including fathers in research on child maltreatment (Dixon, Hamilton-Giachritsis, Browne, & Ostapuk, 2007; Herring, 2009). Introducing fathers into research endeavors presents a different set of challenges compared to the study of mothers who maltreat their children. Research questions include whether there are gender differences in patterns of child maltreatment, patterns of unreliability in the presence and involvement of fathers and stepfathers, transient stepparenting, child maltreatment in the context of cohabiting/dating situations, and different patterns in the forms and severity of child maltreatment that potentially are based on gender. The salience, type, and patterns within and across the foregoing factors may be different for male and female caregivers. The relative lack of information about gender similarities or differences in child maltreatment risk is in need of remedy.

In the early decades of research examining parental risk factors, researchers hypothesized that parents who maltreated their children had their own histories of child maltreatment (Dubowitz, Hampton, Bithoney, & Newberger, 1987; E. Herrenkohl, Herrenkohl, & Toedtler, 1983). Many investigators failed to take the developmental status of the children into account, and thus an infant child might have been classified along with older adolescents as nonmaltreated even though the parents of that child had many more years of parenting ahead of them (Belsky, 1993). Few well-designed prospective studies have provided findings indicating a robust linkage between reported histories of child maltreatment and perpetration of child maltreatment (Egeland, Jacobvitz, & Papatola, 1987; Ertem, Leventhal, & Dobbs, 2000). However, the few prospective studies that have been conducted provide modest support for the hypothesis (Capaldi, Pears, Patterson, & Owen, 2003; Conger, Neppl, Kim, & Scaramella, 2003). Based on combined retrospective and prospective studies, overall support for the intergenerational transmission hypothesis has ranged from weak to modest (Belsky, 1993; Dixon & Hamilton-Giachritsis, 2009; Ertem et al., 2000; Hardt & Rutter, 2004). Most members of the scientific community abandoned the intergenerational transmission hypothesis as a primary causal factor, viewing it as too unidimensional and recognizing that it held up in only a minority of retrospective convenience sample studies

(Babiker & Herbert, 1998; Widom, 1989). Researchers consider it one of many facets of enduring child maltreatment, and the link is not viewed as inevitable (Capaldi et al., 2003; Conger et al., 2003; Dixon & Hamilton-Giachritsis, 2009; Ertem et al., 2000; Hardt & Rutter, 2004).

In prospective studies, researchers estimated the rate of intergenerational transmission to be 30% plus or minus 5% (range 7%–70%; Belsky & Jaffee, 2006; Capaldi et al., 2003; Conger et al., 2003). High-quality work is lacking that highlights conditions under which maltreatment is not transmitted across generations (Belsky et al., 2009). A history of child maltreatment is insufficient, in and of itself, to explain intergenerational transmission (Capaldi et al., 2003; Chen, Liu, & Kaplan, 2008; De Bellis, 2012). Thus, although intergenerational transmission of child maltreatment is a risk factor of interest, it has no greater salience or significance than other risk factors, and it is best understood in the context of a broad range of risk and protective factors.

Researchers instead began to ask what determines whether a victim of maltreatment grows up to become a perpetrator of child maltreatment. Learning theory mediators include modeling, direct reinforcement, coercion training (learning and incorporating methods abusers use to control victims and gain victim compliance), and inconsistency training (randomness in dispensing rewards and punishment that ultimately leads to escalation of negative behaviors not because they are desired but because the response to them is predictable and provides a false sense of control over the situation; Wahler & Dumas, 1986). Other theorists added elements of parental philosophies of discipline and harshness of personality. Thus, social learning theorists would hypothesize that maltreating parents believe in the legitimacy of strict and physical discipline (Simon, Whitbeck, Conger, & Wu, 1991), and childhood victimization leads to aggressive behavior, emotional dysregulation, and limited empathy that promotes predictable personalities and parenting styles in adulthood (Brent & Corwyn, 2002; Feshbach, 1989; Howes et al., 2000; Trickett & Kuczynski, 1986).

Attachment theorists hypothesized that perpetrators of child maltreatment developed an internal working model in childhood that incorporated maltreatment into parenting (Bowlby, 1988; Main, Kaplan, & Cassidy, 1985). Parents who did not experience parental responsiveness in childhood, but who instead experienced rejection or ambivalence, are less open to the signals and needs of their children. They have difficulty taking a child's perspective, and they feel threatened by their children's anxiety (Baer & Martinez, 2006; Cyr, Bakermans-Kranenburg, & van Ijzendoorn, 2010; Main et al., 1985). They are likely to take an adversarial perspective when responding to their children or to believe they must establish control and impose their wishes or needs. Neglecting parents are less likely to believe their relationships with their children can meet their needs, are less likely to believe they can elicit satisfactory responses from their children, and are more likely to report feelings of emptiness and depression (van Ijzendoorn, 1995).

In studies of parents victimized in childhood by maltreatment who successfully refrained from child maltreatment, attachment theorists found that parents who

recalled early maltreatment experiences and integrated them into a revised working model of parent–child relationships more successfully broke the cycle of neglect, whereas parents who ignored their negative experiences were more likely to engage in neglect (Dixon & Hamilton-Giachritsis, 2009; Main & Goldwyn, 1984). Attachment theory provides models of lawful discontinuity of child maltreatment (Belsky & Pensky, 1988). Parents maltreated in childhood who did not maltreat their own children had more extensive social support, had developed a nonabusive close relationship with one parent while growing up, were involved with an emotionally supportive spouse or dating partner, and/or were more openly angry and thus more willing to provide detailed accounts of their negative childhood experiences compared to those who engaged in maltreatment (Caliso & Milner, 1992; Egeland et al., 1987). Thus, the prediction that parents can modify their internal representation of interpersonal parent–child relationship expectations holds up in studies of perpetration and desistance of maltreatment. Researchers continue to examine whether the prediction holds up for men as well as women and attempt to better understand the processes or influences that result in desistance of maltreatment. Personal resolve not to repeat the maltreatment contributes to desistance, but it is less clear under what circumstances such resolve is sufficient (Egeland et al., 1987). There is initial support for the idea that it helps to seek out emotionally corrective close relationships (with nonabusive partners, close friends, therapists) and both formal and informal educational experiences. The question remains as to how individuals built sufficient trust in others to seek out those relationships and educational experiences (Collishaw et al., 2007; Quinton, Rutter, & Liddle, 1984).

Researchers have studied the impact of mental illness on parenting. Personality theorists initially focused on serious mental illness or distinct psychological patterns to explain persistence of child maltreatment. For example, early researchers focused on specific forms of mental illness, specifically parents with psychosis and specific personality disorders, but this approach was abandoned due to lack of support (Belsky, 1993). Research later emerged that focused on specific personality variables, psychological attributes, and symptoms correlated with maltreatment (Kim, Cicchetti, Rogosch, & Manly, 2009). There remains general agreement that few maltreating parents are overtly psychotic or otherwise mentally disturbed to a clinically significant degree. Epidemiological studies show only a modest relationship between mental health problems and parenting problems in large-scale population studies. Studies specific to care and protection samples of parents have yielded more robust data relevant to actual or probable mental health problems in parents who severely maltreat their children (De Bellis, 2012; Kohl, Jonson-Reid, & Drake, 2011). Although it is useful to examine diagnostic status as a possible variable of relevance, no characteristic patterns of symptoms are associated with child maltreatment (De Bellis, 2012; Kohl, Jonson-Reid, & Drake, 2011; Wolfe, 1985). The literature gives only nominal support to any pattern of symptoms of mental illness or personality attributes associated with child maltreatment. There is a

hypothesized modest link between lower intelligence and child maltreatment, but more research is needed to understand, in the context of prospective studies, the level and nature of risk (Valentine, 1990). Researchers have noted trends associating maltreatment with lowered impulse control, lowered self-esteem, impaired capacity for empathy, and impaired ego resiliency (Alink et al., 2012; Cicchetti & Rogosch, 2012; Feshbach, 1989; Kim et al., 2009). Studies have shown inconsistent results in associations between these factors and child maltreatment (Howes et al., 2000). Similarly, studies of associations between maltreatment and depression and/or anxiety have yielded inconsistent results (Banyard, Williams, & Siegel, 2003; Gelfand & Teti, 1990; Goodman & Tully, 2008). Although studies have found interesting associations between variants in genes and resilience in maltreated children (Cicchetti & Rogosch, 2012), behavioral genetics thus far does not hold promise in identifying heritable dimensions of personality related to child maltreatment or the obverse condition of friendly, considerate, responsive parenting (Plomin, Nitz, & Rowe, 1990). Regardless of whether the focus is psychoticism, social-emotional variables, or personality variables, differences in samples, measures, and definitions of child maltreatment systematically affect the results obtained. Research is similarly limited by the use of models measuring only a few of these variables or factors at a time. Researchers have only recently begun to turn to more sophisticated prediction models, using structural and hierarchical equation modeling. Variables that are nonsignificant in isolation can be better understood in the context of multivariate research designs. If child maltreatment is indeed multiply determined and embedded in transactional processes involving parents, children, family systems, and communities, it is best to turn to methodology that seeks contributing rather than determinative agents (T. Herrenkohl & Herrenkohl, 2007). Linkages are more easily identified in designs that allow analysis of predictor variables, interactions, and time-dependent linear equations.

One construct that does seem to hold promise is that of negative reactivity, but its robustness is dependent on attributional variables with potential to exacerbate the reactivity. Researchers focusing on internal psychological resources have examined the influences of negative reactivity and attributional style in parents (Belsky, 1993). Because most episodes of maltreatment take place during routine parent-child interactions, it has been hypothesized that parents who engage in maltreatment may be particularly reactive to aversive events or interactions (Bauer & Twentyman, 1985). Negative reactivity has been observed in studies of parental reactions to tape recordings of crying children and to both stressful and nonstressful parent-child interactions (Watson & Clark, 1992; Youngblade & Belsky, 1989). It is hypothesized that personality traits and emotional processes such as negative reactivity reflect different levels (trait and state) of influence, with neither factor sufficient by itself to turn an interaction into one involving maltreatment (Belsky, 1993). Similarly, attributions of diminished personal control, externality, and instability add another level of analysis to the influences (Bugental, Blue, & Lewis, 1990; Pidgeon & Sanders, 2009). Difficult child behavior is regarded as threatening and unmanageable,

and the parent experiences negative emotional arousal. There appears to be an interactional influence between attributional style and affective orientation. Mothers who attribute negative intent to individuals in ambiguous circumstances are likely to initiate coercive/aggressive interactions with and react negatively toward their children, compared to mothers who do not attribute negative intent (MacKinnon-Lewis, Lamb, Arbuckle, Baradoran, & Veiling, 1992). Researchers have begun to examine the combined effects of affective tendencies, cognitive biases, attributional style, negative reactivity, and specific personality traits in the context of child maltreatment research. Attributional style has a small effect in some studies, but it does not hold up for all forms of maltreatment (Gibbs, 2002). Large-scale studies are needed to better understand the combined effects of multiple levels of influences.

Child Factors. Epidemiological researchers have highlighted age and developmental level as important variables in understanding child maltreatment. Child maltreatment that comes to the attention of authorities is not evenly distributed across the developmental spectrum. Reported cases of child maltreatment decline with age in the United States (Centers for Disease Control and Prevention [CDC], 2012). Younger children are more likely to be the recipients of physical force because they spend more time with and are more dependent on their parents. Child maltreatment peaks for children between ages 3 and 8. Within this age range, children remain vulnerable, and they become publicly visible in child care agencies and schools (Belsky, 1993). Rates begin to drop after age 8, but preadolescents and adolescents remain vulnerable to child maltreatment (CDC, 2012). Results on the link between physical illness/physical disability and child maltreatment are mixed. There is an inconsistent association between physical health status or disability and child maltreatment (Jaudes & Mackey-Bilaver, 2008). Similarly, the problematic behaviors of children are inconsistently related to risk of child maltreatment, and the cause-and-effect relationship between maltreatment and child behaviors is difficult to isolate (Schulz-Heik et al., 2010). Prospective longitudinal designs that enroll children in infancy support the hypothesis that child behavior variables are a product rather than a determinant of child maltreatment (Schulz-Heik et al., 2010).

Contextual Factors. Development occurs in the context of both macrosocial influences (social systems, culture, mass media, political systems, and other broad social systems and structures) and microsocial influences (individual social agency, neighborhood and community variables, face-to-face social interaction variables, and other variables relevant to smaller social influences). These variables influence development. Examples of the larger context include societal trends that contribute to social isolation, historical trends in conceptions of parenting and childhood, and caregiving resources. Researchers examining macrosocial variables have questioned whether there are enduring societal attitudes within which maltreatment

is tolerated and may even flourish. Variables of interest include societal attitudes toward physical/domestic violence and sexual abuse, corporal punishment, and the status and rights of children (Belsky, 1993). Macrosocial analyses of child maltreatment center on societal willingness to tolerate high levels of violence. Examples include tolerance of physical punishment and sometimes infamous efforts by authorities to ignore and even cover up victim complaints of child maltreatment (Brackenridge, Bishopp, Moussalli, & Tapp, 2008; Terry, Schuth, & Smith, 2011). Historical views of children as property and tolerance of school-based corporal punishment are examples of macrosocial influences that may contribute to rates of maltreatment of children (Condie, 2003). Misconceptions concerning the appropriate care of children indirectly influence maltreatment. Evidence that the care of children is not prioritized is reflected in relatively low pay for child care providers, teachers, and other staff members responsible for educating and caring for children in American society. Thus, although the link between macrosocial variables and child maltreatment is not immediate or proximate, it is likely that sociocultural attitudes and practices, conceptions of the rights of children, and socioeconomic variables contribute to rates of child maltreatment.

Another macrosocial phenomenon of interest is competition for resources. The interests of parents and children are not always shared, harmonious, or noncompetitive (Herring, 2009). Children can pose a conflict of interest, parents and children may not share goals and interests related to food and reproductive interests, and those conflicts can lead to child maltreatment (Belsky, 1993). Nurturing responses compete with parental investments and continued reproductive interests, particularly when conditions accentuate the biological conflict of interest between parent and child (Burgess & Draper, 1989). Conditions likely to exacerbate the conflict of interest would include economic or societal instability, unpredictability, and limited resources. Consistent with this point of view is the replicated finding that impoverishment is related to child maltreatment (Brooks-Gunn & Duncan, 1997; Coulton, Korbin, & Su, 1999; Korbin, Coulton, Chard, Platt-Houston, & Su, 1998). The influence of impoverishment on child maltreatment provides support for the contextual features hypothesized to exacerbate child maltreatment. Similarly, limited parental education, unplanned pregnancy, larger family size, job loss, underemployment, and unemployment contribute to resource instability and inadequacy that, from an evolutionary perspective, might be tied to child maltreatment (Zuravin & Greif, 1989).

Social factors influence the age at which individuals become parents. Researchers have examined whether the age at which parents first give birth has relevance to risk of child maltreatment. Risk of child maltreatment is elevated in younger parents, particularly young teenagers. Sociobiological theory predicts that older parents are likely to refrain from child maltreatment due to advancing age and lessened childbearing prospects. Researchers have illustrated an inverse relationship between maternal age and child maltreatment (Whipple & Webster-Stratton, 1991), but the

effect is due primarily to elevated risk among young teenage parents and lessened risk among women reaching the termination of childbearing years (Malkin & Lamb, 1994). Evolutionary theory dovetails with developmental theory in emphasizing egocentricity in young adulthood and generativity in older adulthood. Similarly, the personality and emotional attributes contributing to maltreatment (low empathy) may contribute to reproductive capacity under adverse environmental conditions (Belsky, 1993). Whether reproductive fitness is a primary variable contributing to maltreatment in settings of limited economic resources remains an empirical question in evolutionary biological theory.

Researchers interested in microsocial variables have made progress in distinguishing physically abusive and neglectful parenting features. In particular, early studies of social interaction variables illustrated that neglectful mothers of infants were relatively unresponsive to their infants, and they had lower rates of social interaction and prosocial behavior directed toward older children (Crittenden, 1985). Physically abusive parents engaged in fewer forms of positive behaviors and engaged in controlling, interfering behavior toward their children (Howes et al., 2000; Whipple & Webster-Stratton, 1991). More recent studies have supported those findings (Appleyard, Berlin, Rosanbalm, & Dodge, 2011). Physically abusive parents are more likely than comparison parents to use physical punishment, physical control strategies, and punitive methods (threats and disapproval), and they are less likely to engage in reasoning or to show flexibility in their approaches to discipline (Appleyard et al., 2011; Trickett & Kuczynski, 1986; Whipple & Webster-Stratton, 1991). Thus, maltreatment arises in the immediate context of parent-child interaction or discipline when a parent predisposed to negative reactivity, irritability, or hostility becomes increasingly emotionally aroused, attempts to exert physical influence and control over the child, and loses control. Further learning and attachment influences, shaped by childhood victimization by maltreatment, shapes and exacerbates the behaviors (Belsky, 1993).

Societal trends toward increased isolation is a phenomenon of interest. Social support has positive influences on psychological functioning, and its absence has been linked to child maltreatment (Appleyard, Yang, & Runyan, 2010). Maltreating parents tend to have fewer friendships, to have less contact with immediate and extended family members, and to feel lonely and a sense of social isolation (Corse et al., 1990; Coulton et al., 1999). They are less likely to take advantage of available resources (Appleyard et al., 2011). Researchers have also investigated social and family conflict and competition for resources. Neighborhood quality variables also contribute to maltreatment. In a study of Chicago neighborhood quality in socioeconomically similar neighborhoods that varied in rates of maltreatment, parents coming from high-maltreatment neighborhoods had difficulty thinking of anything good to say about their neighborhoods, were unsatisfied with the quality of physical spaces housing community programs, and viewed the neighborhoods as socially disorganized and lacking in social coherence. By contrast, people from low-maltreatment neighborhoods described their neighborhoods as poor but decent,

had access to more services, and described neighborhood leaders as strong and appreciated (Garbarino & Sherman, 1980; Melton, 1992).

SUMMARY

The foregoing analysis illustrates the main point that child maltreatment is multiply determined by factors operating at multiple levels of analysis that include evolutionary, developmental, situational/contextual, individual, microsocial, macrosocial, and demographic. Maltreatment is the final common outcome of multiple pathways. In any individual child protective service investigation of child maltreatment, it is possible to identify multiple etiological correlates. A different set of correlates, with or without overlapping variables, might not appear in the next investigation. Unique clusters may recur across cases but not in a reliably predictable manner. The multidetermined nature of child maltreatment must be considered by both researchers and clinicians in order for them to better understand and empirically substantiate the transactional processes presumed to contribute to child maltreatment (Belsky, 1993; Condie, 2003). Researchers analyzing risk factors imperfectly distinguish between different forms of maltreatment, because of the frequent comorbidity of different forms of maltreatment and the lack of distinctiveness of any individual or cluster of predictors for any one form of maltreatment (Condie, 2003). A point of emphasis for future research will be to highlight features that distinguish levels of severity and chronicity of child maltreatment. Two physically abusive or neglectful parents might not be equally abusive or neglectful. Descriptive research is needed to better understand what factors contribute to severity and chronicity of child maltreatment and what factors contribute to lessened or diminishing severity and desistance of maltreatment.

Not all abusive or neglectful parents are the same kind of person, and researchers have begun examining the utility of classification schemes. Researchers are beginning to describe typologies of maltreating parents that include combinations of variables at different levels of analysis. The multidetermined nature of child maltreatment may make this undertaking difficult, at least from the perspective of intervention planning. It is a challenge to design interventions that address the needs of a diverse group of parents with diverse contributing influences related to child maltreatment, particularly in the setting of a relatively high rate of child poverty (Korbin et al., 1998). Not every young parent, impoverished parent, single parent, or parent with children having closely spaced births mistreats his or her children. Thus, interventions must address more than impoverishment, fertility, and social support. Program developers have begun to address parental developmental histories, negative emotionality, emotional reactivity, and insecure expectations, but with mixed results (Kohl et al., 2011). Interventions must target multiple factors simultaneously, creating incentives for adolescents to remain in school, reducing school truancy and academic underachievement, addressing neighborhood quality, and increasing school-based case management (Belsky, 1993).

EVALUATION METHODOLOGY

Methodology for evaluations depends on the nature of the referral question. The use of consistent methodology and the use of multimodal assessment procedures enhance the reliability and validity of evaluation results. In child maltreatment cases, flexibility in methodology across referral questions is needed to accommodate the degrees of breadth and depth necessary to answer a given referral question or set of questions. Care and protection evaluation methodologies and reports range from brief consultations to comprehensive descriptions of multiple family members and their interrelationships (Condie, 2003). Although there is no single methodology for care and protection evaluations, the prototypical example includes:

- Obtaining informed consent
- Interviewing one or more parents or caregivers
- Observing the parents or caregivers with the child (when indicated)
- Interviewing the children
- Gathering collateral information and relevant records
- Seeking releases for access to privileged and/or confidential records
- Administration of psychological measures or tools when indicated

Examples of measures include measures of general mental health or adaptive functioning, measures related to specific symptoms or behaviors, or measures that address risk factors and hypothesized correlates of elevated risk of child maltreatment. Specific potentiating factors are included to enhance the predictive validity of risk assessment beyond that of amorphous broad constructs of mental illness, mental deficiency, or organic impairment. When translating information from research to clinical practice, however, one must consider potential compromises to validity and reliability when the relative weight of any risk factor is unknown as it applies to clinical populations. When those factors are included in a risk assessment, there must be a clear link between any factor or combination of factors and child maltreatment. A person can be mentally ill, and a person can engage in child maltreatment. The co-occurrence of those two elements in a single case might or might not be the result of a direct or indirect causal link. Thus, it is important to examine how mental illness manifests itself (e.g., the impact on the individual's interpretive process, attributional style, etc.) and any possible link to elevated risk of child maltreatment.

EVALUATING CAREGIVERS

Using a systematic approach, the caregiver portion of the evaluation satisfies informed consent procedures, introduces the referral questions and evaluation content, and reviews the anticipated scope of the evaluation. The referral question(s) frame the evaluation methodology. Multimodal assessment is conducted to enhance the reliability and validity of the evaluation results. Good methodology allows for

flexibility to accommodate different degrees of breadth or comprehensiveness of referral questions, caregiver variables, and caregiver–child interaction variables.

Informed Consent and Notification of the Limits of Confidentiality. The first step of any evaluation is to obtain informed consent in keeping with prevailing regulations and practice standards. The individual being interviewed must be informed of the limits of confidentiality prior to being interviewed (American Psychological Association [APA], 2013; APA Committee on Professional Practice and Standards, 2011 [the Specialty Guidelines are reprinted as the appendix to this volume with permission of the APA]). If the individual does not comprehend the notification, steps should be taken to determine whether the evaluation ought to proceed. Examples include contacting the referring attorney or notifying the court in the case of a court-ordered evaluation. The explanation should include:

- A clear explanation of the referral question
- The individuals who are a party to the evaluation
- Who will view the report
- The lack of confidentiality
- Who “owns” the report
- Provisions (or lack thereof due to judicial restrictions in some jurisdictions) for release of the report to individuals who are not a party to the legal proceedings
- The difference between medical records and forensic records as defined in state or federal statutes and regulations relevant to both psychological record keeping and care and protection proceedings (Condie, 2003)

When a report describes multiple parties to the case, the evaluator should explain restrictions in releases, namely that one party cannot authorize release of the full report unless all of the other parties also sign releases. Individuals should be told that:

- They are under no obligation to participate in the evaluation.
- They may decline any further questions if they agree to the evaluation but later change their minds.
- They may refrain from answering certain questions if they wish to keep information private.
- Nothing they say remains “off the record.”

Under the circumstance of a court-ordered evaluation, individuals should be advised that the evaluator will prepare a report that includes information gathered from other resources regardless of whether they choose to participate. For all participants in the evaluation, the notification must be given to satisfy ethical and legal requirements.

External Validity. In determining the most appropriate methodology, careful attention should be paid to external validity and its limitations. Contemporary theories and research on parenting behavior continue to focus primarily on optimal or typical parenting rather than the minimal parenting competence needed to satisfy legal standards (Barnum, 1997; Budd, Poindexter, Felix, & Maik-Polan, 2001; Condie, 2003). Few valid indicators of psychological constructs specific to parenting (warmth, nurturance, responsiveness) have been developed for use in forensic assessment (Budd & Holdsworth, 1996). There is little research or clinical consensus as to which psychological constructs, behavioral indices, or functional impairments directly relate to the minimal legal criteria necessary to demonstrate parenting competence (Ayoub & Kinscherff, 2006; Barnum, 1997; Budd et al., 2001; Condie, 2003). Psychological assessment measures, even when specific to parenting behavior, typically were not normed on care and protection samples or other samples of parents thought to be at risk for maltreating their children (Brodzinsky, 1993; Budd & Holdsworth, 1996). Few normative data are relevant to samples of families involved in care and protection matters (Ayoub & Kinscherff, 2006; Barnum, 1997; Condie, 2003). By the same token, there is little basis for conducting care and protection evaluations without turning to psychological assessment measures, except when referral questions touch on issues for which assessment is irrelevant or poorly supported. When appropriately used and interpreted, assessment measures improve the reliability and validity of care and protection evaluations. Although evaluation approaches to care and protection matters take many forms, it is common for psychologists to use some form of psychological assessment in evaluations of parental functioning to enhance the measurement of relevant behaviors and skills, the interfering effects of substance abuse and/or mental illness, cognitive capacities, and general functioning (Ayoub & Kinscherff, 2006; Budd et al., 2001; Condie, 2003). Evaluators should take a multimodal approach to assessment, applying the tradition of seeking converging data (APA Committee on Professional Practice and Standards, 2011; Budd et al., 2001; Condie, 2003). The literature contains suggestions for core features of typical questions asked in care and protection evaluations. Because the band of possible referral questions is not narrow, adherence to a particular methodology in all instances is neither feasible nor indicated. Nonetheless, enough is known about the basic features of commonly recurring referral questions in care and protection matters to develop relatively consistent methodologies.

Some clinics with evaluation contracts face limitations in resources, and thus the scope and feasibility of comprehensive methodology must, unfortunately, take those financial limitations into account. Thus, referral questions are sometimes limited to critical, but highly focused, questions or content areas to accommodate heavy referral loads in the face of scarce resources (Condie, 2003). Evaluators in some settings may pare down evaluation procedures such as multiple sessions or parent-child observations that might improve yield of information but with diminishing returns. For example, although parent-child observations provide useful information about parent-child attachment, they rarely yield critical

data concerning child maltreatment. Although a parent–child observation might contribute to an optimal evaluation, a careful analysis weighing potential information gain against time and labor investment/scheduling issues might lead to a conclusion that it does not meet demands of what is minimally necessary in every evaluation. This particular procedure is highlighted not because it is weaker than other modalities and not because it lacks external validity per se; the same could be said of many other procedures. Evaluators operating with limited budgets often need to make difficult decisions that allow them to meet ethical and legal obligations without producing invalid or unreliable results but also without exceeding budgetary constraints. Researchers and scholars who ignore the reality of limitations in the child protective and court clinic funding mechanisms risk producing methodology recommendations that lack financial feasibility. Sometimes resource limitations spur advances in approaches to court clinic services (Chuang, Moore, Barrett, & Young, 2012).

Clinical Interview. Clinical interview comprehensiveness is determined by the referral question. Some interviews might address an extensive range of historical and current factors, with an in-depth focus on particular spheres of functioning. Other interviews might be specific to one or two areas of functioning (Budd, Connell, & Clark, 2011). Regardless of the comprehensiveness of the interview, it is important to include questions that inquire about both strengths and weaknesses of parenting skills (Condie, 2003). Focusing only on parenting or functioning deficits carries the risk of neglecting important protective factors or compensatory skills, alienating the interviewee, and leading to an imbalance in the analysis of positive and negative factors in the interpretive process. A variety of interview strategies incorporate the necessary content domains needed to respond to referral questions. The evaluator should choose an approach with which he or she is well trained and highly familiar and that is easily adapted to a variety of parenting concerns that would be raised in the context of child protection matters (Budd et al., 2011; Condie, 2003).

Clinical interviews ordinarily progress from emotionally neutral topics to more sensitive areas (Ownby, 1997). In care and protection evaluations, however, it is useful to begin with the parent's perspective of his or her history of involvement with the child protective service system. Interviewees tend to focus on this material at the outset of evaluations regardless of whether the interviewer directs them toward or away from it, and they report frustration if the interview deviates too quickly from the centrality of their concern. There is a cathartic quality to interviewees' first accounts of their involvement with the child protective service system, regardless of whether their description is appropriately detailed or even accurate. Many times they report that the evaluation provided them with their first opportunity to provide their description and perspectives of what happened. This approach eases tension and contributes to rapport building, but finesse is required so that the interviewee does not mistakenly view the evaluator as an ally or advocate (Condie, 2003; Crenshaw & Barnum, 2001).

The first account of the parent's history of child protective service system involvement usually lacks some or many details central to the alleged maltreatment (Chuang et al., 2012; Crenshaw & Barnum, 2001). Follow-up questions are needed in order to gather relevant details. Sometimes details will be gathered in different phases of the interview. Inconsistencies and gaps within the interviewee's accounts, and between interviewee accounts and records, should be noted and addressed. Usually a curious or inquisitive approach is sufficient to elicit greater detail and forthrightness, particularly if follow-up questions are directed toward multidimensional facets or elements of a topic and are designed both to provide richness of texture and to complete any gaps in information (Condie, 2003). Eliciting details through the exploration of multidimensional facets has the benefit of uncovering inconsistencies, and interviewees sometimes will move, either abruptly or gradually, toward candid disclosure of information in the context of textured questioning. By contrast, some evaluators find it useful to use a direct or somewhat confrontational approach. Confrontational approaches must be used with caution, particularly when there is a question of cognitive limitations and suggestibility in the interviewee (Sgroi, 1989). With parents who are not cognitively impaired, it is reasonable to approach a topic more than once if there is inconsistency or if the interviewee avoids responding to questions. Persistence typically yields more information than confrontation. Multiple questions and/or interviews sometimes accomplish multiple goals of facilitating rapport, highlighting contradictory information, and generating of details, resulting in gradual acknowledgment of relevant parenting weaknesses and risk factors (Condie, 2003).

Other interview content depends on the nature of the inquiry. If the referral question contains a request for information about parenting abilities or risk of harm, the interview concerning historical data in most cases includes a broad range of questions about parenting skills and schedules, a variety of factors known to contribute to risk of maltreatment, and a variety of protective factors (Condie, 2003). If information in the records reveals a smaller band of potential risk factors (e.g., known relapsing substance abuse in the relative absence of other indicators of risk), there may be less need for a broadly focused interview except as it applies to known risk factors. When comprehensive data are needed, the evaluator typically moves from neutral to potentially emotionally arousing topics. Typical areas of inquiry include:

- Family of origin history (focusing on quality of relationships, loss, and intrafamilial trauma)
- Important adulthood relationships
- Dating and marital history
- The decision process to bear children
- Educational and occupational history
- Mental health and substance use history

- Medical history (including neurological data)
- Psychosexual history
- Neglect history
- Violence history
- Criminal history (Dixon & Hamilton-Giachritsis, 2009)

The depth of inquiry into any topic area depends in part on a preliminary review of records that guide interview content and on information that the individual might divulge in the course of the interview (Condie, 2003). Depth might also be determined by whether a referral question includes a request for data relevant to treatment progress and prognosis. For example, if substance abuse treatment response and prognosis is a question of interest, the interview content typically would contain questions relevant to:

- History of substance use
- Length and quality of periods of sobriety
- Factors that potentiated and hindered sobriety
- Effectiveness or lack thereof of past rehabilitation efforts
- The individual's current level of participation in and stage of rehabilitation
- Motivating influences for remaining sober
- Options available for resisting urges to drink or use drugs
- Relapse prevention plans
- Other relevant information (Connors, Longabaugh, & Miller, 1996)

The modality of treatment would have bearing on follow-up questions. If a self-help group is the main treatment modality, the evaluator would inquire about these areas:

- Nature and frequency of self-help meetings
- Steps or stages of the self-help process
- Identification of and reliance on sponsors
- Reliability and availability of sponsors or other support persons
- Types of meetings attended
- Implementation of recommendations gleaned from self-help books (Condie, 2003)

Questions about relapse prevention planning would center on:

- Comprehensiveness and thoughtfulness of the plan
- Whether elements of the plan are committed to memory or are readily available in written form
- How often the plan is updated
- Whether the individual's relapse prevention therapist believes the plan is appropriately detailed and feasible (Connors et al., 1996)

Regardless of the treatment modality, the interviewer should have an appropriate level of knowledge about the addictive potential and differential impact of various illicit or licit substances and the impact of those substances on the individual's functioning should he or she relapse. Interview questions should also be tailored to other issues of concern, including possible comorbidities (mental illness and substance use, mental illness and cognitive impairments, etc.) having a direct or indirect impact on parenting capacity (Condie, 2003).

Risk Assessment. Actuarial methods with varying rates of reliability, validity, and clinical utility have been developed for other forms of violence risk (Borum, Otto, & Golding, 1993; Quinsey, Harris, Rice, & Cormier, 1998) but not for child maltreatment risk. Risk assessment typically is included in the interview process. The level of interviewer concordance and the specificity of risk communication depend on the validity of the methodology. Clinicians reach different levels of concordance based on whether they employ probability equations or rely solely on clinical inference (Borum et al., 1993). Risk matrices specifically designed for predictions of child maltreatment have yet to be developed, but cautious use of research guides the process. Specific studies of risk associated with child maltreatment have been based on small sample sizes and a small dimension of variables compared to broader studies of violence risk. Some wide-scope studies are beginning to appear (Appleyard et al., 2010, 2011). It remains to be determined whether the weight and relevance of risk factors differ when comparing physical violence, neglect, and sexual abuse (Ayoub & Kinscherff, 2006).

Interview Content for Risk of Physical Abuse. Physical abuse is a low-frequency act that usually occurs in a private setting. Because of problems with self-report data, it is difficult to demonstrate in applied research whether personality features or behaviors (hostility, criticism, threats) have potent predictive or discriminative validity for physical abuse (Watson & Clark, 1992; Widom, 1989; Zuravin & Greif, 1989). Examples of risk factors that have at least a moderate link to risk of physical abuse (Belsky, 1993; Ertem et al., 2000; Goodman & Tully, 2008; Miller, Cohen, & Wiersema, 1996; Trickett & Kuczynski, 1986) include these:

- Untreated major mood or thought disorders
- Failure to acknowledge a mental illness or the need for treatment
- History of violent outbursts of temper
- Active drug or alcohol addiction
- Childhood history of abuse
- Other adverse childhood experiences (hostile and rejecting family environment, harsh and unfair discipline, placement in foster care or multiple placements, parental discord, institutional upbringing)
- Few ties to neighbors or community

- Violent relationships with intimate partners
- Gross misperceptions about a child or child development
- Gross misperceptions about useful or appropriate discipline strategies
- Unrealistic expectations of children
- Difficulty discerning and responding to cues from the child
- Insecure parent–child attachment
- Role reversals
- Scapegoating the child
- Extreme worry about a child’s well-being
- High levels of parenting stress or social isolation
- Hazardous home environment
- History of violent behavior
- Deliberateness with which the individual harmed the child in the past
- Extent and frequency of harm to the child
- Allowing perpetrator access to the child

Other variables of relevance that sometimes contribute to risk of maltreatment include:

- Adequacy of supervision
- Safety of the home environment
- Age and visibility of the child
- Caregiver age and maturity
- Mental and social development of the child (to better understand variables contributing to a child’s capacity to speak up on his or her own behalf)
- Level of fear the child expresses about the caregiver or the home environment
- Presence of other adequate caregivers
- Level of stress and availability of supports
- Caregiver’s victimization in adulthood by domestic violence

Internal variables include:

- Cognitive appraisals of caregiver–child conflict
- Attributions of blame and responsibility
- Anger management skills
- Recognition of the problem
- Capacity to select suitable substitute caregivers
- Internal response to the child’s misconduct
- Mental appreciation and assignment of family roles (Kim et al., 2009; Kohl et al., 2011; Widom, 1989)

Much attention has been given to substance abuse in maltreating families. For example, Gaudin (1994) reported that it was a factor in 80% to 90% of child maltreatment cases.

Static risk factors are historical factors that cannot be modified. Dynamic or modifiable risk factors that can change over time (Borum et al., 1993) include, among others:

- Caregiver's degree of acknowledgment of the problem
- Intensity of relevant treatable symptoms
- Willingness to take advantage of social support
- Willingness to better understand child development and the needs of children
- Modifiable impinging stress factors
- Cognitive appraisals
- Individual's justification for the maltreatment (Condie, 2003)

Researchers are just beginning to study a number of issues, including:

- How readily these factors are modified
- Which factors tend to be more amenable to intervention
- Whether individual variables contribute to amenability (e.g., would a sociable person isolated because of neighborhood variables rather than individual predilection be more amenable to social skills training)
- Individual potencies of the factors (e.g., the intractable nature of cognitive deficits versus educational background deficits whose outcomes mirror cognitive deficits)
- Potencies of combinations of the factors (e.g., whether certain combinations of risk lead to intractability or certain other combinations lend themselves readily to intervention)
- Desistance of risk in maltreating parents (Appleyard et al., 2010, 2011; Marshall, 2012; Putman-Hornstein et al., 2011)

Interview Content for Risk of Sexual Abuse. Risk assessment of sexual offending incorporates specific factors different from those of other forms of violent offending (Hanson & Thornton, 2000). Although some of the predictors overlap between groups of violent offenders and sexual offenders, risk of child sexual abuse is more directly linked to psychosexual history aberrations than to violence risk. Potent predictors include a reduction in family boundaries, symbiosis of relationships, misdirection of sexuality, cognitive distortions in identification and affiliation, hostility, and aggression (Everson & Faller, 2012; Neutze, Grundmann, Scherner, & Beier, 2012). Although all forms of child maltreatment have an individual component, an interview directed toward sex offending and sexually abusive conduct typically focuses more heavily on individual risk variables than other forms of risk factors. Attention is also paid to the circumstances that make it possible for the maltreatment to occur (e.g., isolation, control). In the clinical interview, neither the sexual nature nor other facets (exploitation, power, humiliation, violence) should be ignored in examining the nature of child sexual abuse.

Psychopathological models of deviant sexual arousal and social-cognitive functioning have been developed to explain sexual offending. Theories tend to be based on samples of convicted offenders and therefore may not fully capture the patterns of functioning seen in undetected offenders who remain in the community or who have never been convicted (Hanson & Thornton, 2000). Factors in explanatory theories include:

- Arrested psychological development
- A sense of inadequacy and immaturity
- Deviant patterns of arousal
- Identification with the aggressor
- Imprinting and conditioning from childhood victimization
- Poor or awkward social skill development
- Socialization that values dominance and power (Cohen, Frenka, Mojtabai, Katsavdakakis, & Galynker, 2007)

Areas of clinical interview inquiry relevant to risk assessment and treatment progress as it applies to specific offending behavior include:

- Offender's recall of details
- Degree of aggression or overt violence in offenses
- Seduction and victim grooming strategies
- Frequency and duration of offenses
- Length and progression of history of sexual offending
- Offense characteristics
- Number of victims in relation to victim access
- Victim selection characteristics
- Preferred victim type
- Victim blame
- Appraisal of victim harm
- Acceptance of personal responsibility for offending behavior
- Degree of arousal and habituation
- Exploitative or addictive behaviors related to the offenses
- Criminal arrests and convictions

Social and cultural factors of victim blame and the reluctance of the legal system to prosecute and punish offenders are hypothesized to contribute to the disinhibition of offenders. Typologies have been developed based on victim gender preference, offender gender, level of fixation for a particular form of offending, degree of regression of the offender, and incest versus extrafamilial victims (Waterman, 1986). There is disagreement over shared features that cut across groups of offenders.

Interview Content for Risk of Neglect. Assessing risk in neglectful parents poses a challenge because of the relative infancy of research relevant to risk of child neglect.

There is enormous heterogeneity among neglectful families (Gaudin & Dubowitz, 1997). Neglect commonly occurs alongside other forms of maltreatment. Researchers have not identified as discrete a set of risk variables as have been identified for physical or sexual abuse. The variability among neglecting families may mask between group differences, correlates or causal factors specific to subtypes of neglecting parents, and the relevance or potency of risk variables across types of neglecting parents (Condie, 2003). Neglect does seem to be associated with these issues:

- Blunted affect
- Apathetic or passive-dependent interaction styles between parents and children
- Nonreciprocal relationships among family members
- Cycling between passive and aggressive behavior
- Critical dispositions (Gaudin & Dubowitz, 1997)
- Chaos and poor planning capabilities
- Impulsive actions
- Conflict-laden relationships between parents
- Social isolation
- Poor quality of social support
- Poor social skills
- Rejection by community members (Mennen, Kim, Sang, & Trickett, 2010)

Protective Factors. Protective factors influence a reduction in or desistance of risk (Lodewijks, de Ruiter, & Doreleijers, 2010; Loeber, Pardini, Stouthamer-Loeber, & Raine, 2007). Protective factors associated with lowered risk of child maltreatment include:

- Social support
- Good parenting role models
- Acknowledgment of parenting problems
- Recognition of the impact of parenting problems on the child's development and functioning
- Adherence to recommended interventions
- Identification of compensatory strategies
- Utilization of community resources
- Positive rehabilitation progress reports
- Stable and consistent visitation with the child
- Recognition of the impact of separation on the child
- Separation from abusive partners
- Strong postreunification after-care plan (Condie, 2003)

Applied clinical and prospective studies of maltreating parents involved with the child protective service system are needed to better understand the relevance and

potencies of individual protective factors or combinations of factors (Colman, Kim, Mitchell-Herzfeld, & Shady, 2009).

Caregiver–Child Relationships. Evaluation of the caregiver–child relationship usually involves a combined approach utilizing parent interview data, child interview data, observational data, record review of historical data, and collateral interviews. Typical inquiries focus on the:

- Strength and quality of the relationship
- Presence and degree of emotional closeness
- Parental perceptions of the child (and child perceptions of the parent)
- Parent’s ability to promote appropriate development in the child
- Parental responsiveness to the child’s needs (Stahl, 1994)

Parents play a role in assisting children through developmental stages. When more than one child is involved in a care and protection petition, the evaluator should individualize the assessment of the parent’s relationship with the child and examine the parent–child relationship and other family dynamics when all children are present (Condie, 2003; Stahl, 1994). Inquiries of a parent might include the:

- Parent’s description of the child
- Awareness of the child’s activities
- Expectations of the child
- How the parent responds when the child seeks attention
- Concern for the child
- Negotiation of parent–child conflict
- Reports of close calls of maltreatment
- Capacity to set a daily routine
- Repertoire of discipline strategies
- Awareness of the child’s friendships and school activities (Budd et al., 2011; Condie, 2003)

Caregiver–Child Visitation. The criteria that visitation agency representatives use to determine the nature and extent of contact between parents and children in foster or residential care are sometimes idiosyncratic. Some agencies form internal standards, but there is little consensus across agencies about criteria for determining frequency of contact, whether contact ought to be supervised, who qualifies as appropriate supervisors, and whether visitation should take place in a natural setting or a visitation center. There is limited information concerning the predictive utility of parent–child visitation reports or observations for determining the quality and safety of supervised or unsupervised visitation contact. Children with histories of intrafamilial trauma report a variety of responses to visitation (Johnston & Goldman, 2010; Pearson & Theonnes, 2000). Evaluators sometimes rely on visitation

supervisors as collateral contacts. It is appropriate to seek visitation observations from collateral contacts who have had an opportunity to view parent–child visits. It is best to seek data that are observational and not conclusory regarding the appropriateness of the visitation plan, reunification, or termination of parental rights. If data gathering is kept at an observational level, it preserves the roles of the visitation supervisor and the evaluator (Pearson & Theonnes, 2000). Parents report mixed experiences at visitation centers. Some parents prefer visitation center supervisors to other supervisors because such supervisors are trained to remain neutral during supervision, because parents feel less awkward in the presence of a professional supervisor compared to other supervisors, and because parents find it is easier to reschedule missed appointments. Concerns about visitation centers include geographic distance from the caregiver’s home or child’s foster home, formality that contributes to the discomfort of being scrutinized, and insufficient staffing. Regardless of the method of supervision, keeping foster children connected to parents through visitation and other forms of contact is essential for successful reunification because it facilitates ongoing family relationships and has the potential to set the child at ease in foster care (Cantos, Gries, & Slis, 1997).

Observations and Home Visits. Observed sessions are a useful aspect of multimodal assessment, but they are readily influenced by social desirability. Most parents are on their best behavior when observed. Observations are illustrative when young children have not yet developed the linguistic capacities to express their level or quality of attachment to parents or when there is obvious conflict between a child and a parent. They may help illustrate how much of the conflict is initiated by the parent and how much it is linked to an adolescent’s struggle for autonomy and independence (Stahl, 1994). Observed sessions can be useful in rare instances in which the parent cannot conduct themselves appropriately even under observation (Condie, 2003; Stahl, 1994). Evaluators have a choice of using structured and unstructured visitation methods. It is difficult to control for reactivity effects when using observational measures due to the lack of blindness of raters to hypotheses concerning why the observation is taking place (Ronay, 2011).

Caregiver Relationships With Helping Sources. Referral questions sometimes address the impact of social isolation or access to social support on parenting. Particularly in cases of neglect, social isolation often is a prominent feature of child maltreatment (Runyan et al., 2005). The caregiver’s capacity to develop relationships with helping sources, positive support individuals, mentors, and nonneglecting role models might be a main goal of an intervention plan. When the caregiver’s relationships with helping sources are examined, inquiry should focus on the breadth and depth of social contacts (Zuravin & Greif, 1989). The reasons for limitations in quality or persistence of constructive relationships should be queried. The quality of relationships whose onset occurred during the pendency of care and protection

oversight should be carefully examined because they could indicate that the caregiver is making a meaningful attempt to develop relationship skills, or they could be transient or superficial relationships meant to mollify child protective service workers (Condie, 2003). Collateral contacts with those helping sources often serve to clarify whether the caregiver is meaningfully forging relationships, learning when and how to take advantage of and rely on helping sources. Similarly, contacts may serve to clarify whether the caregiver is working to curtail or end relationships that pose a maladaptive influence on parenting.

Readiness for Transitions. Referral questions sometimes center on what the parent is doing to prepare for reunification or, conversely, whether the parent has made a competent decision to release his or her child for adoption. Parental self-report data help clarify (1) the parent's own appraisal of skills and limitations in parenting, (2) the degree to which the parent believes he or she has met intervention goals, (3) the degree to which the parent has meaningfully addressed the original problem of maltreatment and factors that led up to it, and (4) the realistic nature of the parent's readiness to regain custody of the child (Ayoub & Kinscherff, 2006). Competence to release a child for adoption requires inquiry into:

- The rationality of the parent's reasons for giving up custody
- Whether his or her decision is voluntary or unduly influenced by other individuals
- The parent's immediate and long-term goals in his or her own life and relationships
- Whether regaining custody is a viable possibility (and, if so, why the parent would choose otherwise)
- Whether a clear process has taken place in which the parent has considered all possible ramifications
- The realistic nature of the parent's wish to have posttermination contact

Courts do not always allow posttermination contact, and when they do, they usually set strict limitations (Condie, 2003).

Psychological Measures and Tools. Psychological measures and tools have the potential to yield useful and relevant data when used appropriately. When measures are used out of context, overinterpreted, or otherwise misinterpreted, assessment data can be a serious drawback to the utility of forensic assessment (Ayoub & Kinscherff, 2006). The utility of any assessment measure or procedure rests on its reliability and validity as it applies to the specific assessment context. When choosing psychological measures, it is best to begin with theoretically or logically derived hypotheses concerning behavioral features or possible causal or explanatory variables that are relevant to the referral question. The evaluator should carefully consider the degree to which the measure demonstrably corresponds to the behavior or set of behaviors

of interest. It is unproductive to use traditional measures indiscriminately, but they can be quite useful if selected carefully based on research or theories supporting a relationship between the indices and specific behaviors of concern in a parent's history or current functioning (Condie, 2003). For example, global indices such as intelligence, diagnostic status, or personality functioning might be used as explanatory data for parenting problems, maltreatment risk, or suitability of interventions. A stronger analysis would integrate indices of judgment and reasoning relevant to the parenting problems (Budd & Holdsworth, 1996). The intellectual, adaptive, and judgment/reasoning indices themselves are not sufficient, in and of themselves, however. They must be meaningfully linked to the maltreating behavior(s) of concern (Condie, 2003).

Psychological assessment data are useful in evaluating a number of issues, including:

- The relationship between current and premorbid functioning
- Specific potential that could contribute to the development of adequate parenting competence
- Specific deficits likely to impair parenting competence
- Global functioning problems that might contribute to parenting deficiencies
- Reasons why a particular parenting intervention might be suitable or not for a given parent
- Problems and deficits that might explain intractability
- Skills and compensatory strategies that could be capitalized on (Ayoub & Kinscherff, 2006; Condie, 2003)

In making a determination of the appropriateness of using psychological assessment measures, it is important to remember that not all parents with cognitive limitations, mental health problems, or a history of using substances engage in child maltreatment. There must be a clear link between those issues and the child maltreatment in order to justify highlighting their hypothesized contributions to the parenting behaviors of concern. For example, limited intellectual functioning of a parent, as assessed by intelligence quotient and scales of adaptive behavior, would be insufficient to explain limitations in parenting capacity unless those indices serve as a framework for understanding specific functioning limitations directly relevant to the parenting behaviors of concern. When relevant, however, assessment measures might prove useful. For example:

- Intellectual and adaptive behavior measures may provide useful information relevant to a parent's specific strengths and weaknesses.
- Achievement and basic skills assessment measures may provide information about why a parent might benefit from or struggle through a particular intervention relevant to parenting.
- Neuropsychological assessment of executive functions may provide a framework for better understanding a parent's capacity for self-control.

- Cognitive and memory assessment may illustrate whether deterioration has resulted from chronic substance abuse (particularly when premorbid indices are available).
- Personality or psychopathology assessment may highlight the ongoing nature or severity of specific facets of interfering mental health or personality impairment on parenting behaviors (Condie, 2003).

Measures should be chosen based on reasonably formed hypotheses about possible causal variables that might contribute to parenting deficiencies or patterns of maltreatment. The need for psychological assessment often is reasonably drawn from the degree and credibility of prior documentation of issues that might be relevant to parenting concerns. Records help highlight the chronic versus episodic nature of problems that interfere with adequate parenting. They help illustrate the plausibility of hypothesized causal links. Diagnostic information alone is not an automatic indicator of the potential for maltreatment, but records help to clarify the extent to which a linkage between symptoms of mental illness or substance abuse and problems in parenting ought to be explored further (Ayoub & Kinscherff, 2006). Assessment data supplement the interpretive process, but they do not prove or disprove allegations of child maltreatment or parenting capacities. Data must be integrated with historical patterns of deficits and potentials and interpreted carefully (Budd & Holdsworth, 1996).

Care and protection cases focus on minimal parenting competence rather than ideal parenting abilities. Because definitions of minimal parenting competence are legally derived and vary from state to state, those definitions may not translate well into theories or techniques that lend themselves to scientific scrutiny. The relevance of any assessment measure, whether global or specific, to parenting behaviors of concern in care and protection matters must be determined on a case-by-case basis, using theory to guide hypotheses and to choose appropriate measures in specific cases. Most experts recommend that the assessment of parenting abilities and capacities be based on a functional set of behaviors in a specific context, with due attention to cultural variations in parenting (Condie, 2003; Weiss & Rosenfeld, 2012). Caution is in order because cultural variations should not be used to divert attention from maltreating behaviors that obviously are serious and deviate from cultural norms. A behavior or style of parenting that is normative should not be used to defend or excuse a parent who takes advantage of that norm in order to seriously harm a child. The evaluator should take steps to understand cultural norms, relying on local experts when necessary, to understand how to interpret measures validly in light of differing standards of behavior (APA Committee on Professional Practice and Standards, 2011).

Records and Collateral Reports. Child protective service records should be reviewed with three goals in mind: (1) documentation of the maltreatment, (2) service plan utility and integrity, and (3) documentation of parental self-reports of maltreatment.

The original allegations of child maltreatment are documented in those records, along with investigation summaries and an indication of why the allegations were supported. Records of allegations and investigations help illustrate the original concerns, whether social services representatives and attorneys have remained focused on the original maltreatment, and the degree to which a parent's report of maltreatment in the context of a clinical interview comports with or deviates from documented information (Condie, 2003). The records illustrate the frequency with which there have been supported allegations of maltreatment. When records are descriptive and accurate, they can be used to gauge parent progress in acknowledging the frequency and severity of maltreatment, a hypothesized necessary step in seeking meaningful treatment (Hardt & Rutter, 2004).

Child protective service records usually contain a service or intervention plan for both parents and children. If only one caregiver is accused of maltreatment, a second caregiver might be included because of concern about tolerance of maltreatment or pathological passivity toward the maltreating caregiver. Another biological parent not living with the family of concern may be identified and included in the intervention phase as well. Children are included in service plans so that interventions can be identified to help them recover from the effects of child maltreatment and remain safe from any further maltreatment. A parent's awareness of the details of the service plan is usually informative. A parent with little awareness of the service plan is unlikely to have taken it seriously whereas a parent who supplements the service plan with other similar services is likely to be serious about change. As in the assessment phase, the issue of social desirability is a necessary component of this phase of the evaluation (Gaudin, 1994).

The service plan sometimes highlights the degree of comportment with identified service goals and the goals of the evaluator's referral question. Agencies sometimes operate with competing goals, providing a service plan that focuses on reunification but looking ahead to a legal process that anticipates a hearing on termination of parental rights. The dual goals are necessary in cases in which statutory provisions highlight reunification, while realistic goals require planning for possible permanent separation of parents and children or even termination of parental rights. Some state statutes call for reunification planning in the early stages of cases and a permanent separation or termination of parental rights hearing after a specified period of time (Condie, 2003). Thus, the intervention planning phase may have two competing goals. A referral question in the early phase of intervention might highlight a parent's amenability to treatment, a referral question in the midphase might highlight parental responses to treatment and whether the quality of the parent-child relationship has improved demonstrably, and a referral question in the latter phase might emphasize a child's attachment to substitute caregivers or the psychological impact of termination of parental rights on a child (Condie, 2003). Sometimes these evaluation goals might be posed concurrently with the same or different evaluators. The service plan may be a launching point to determine if appropriate services have been recommended or neglected in the plan (Gaudin, 1994). The plan

yields information about the identities of service providers who should be included in the list of collateral contacts. Contact with service providers sometimes yields data relevant to whether a parent is accessing services appropriate to his or her level of intellect, whether the parent is benefiting from services, and whether the services are directly or indirectly relevant to the parenting behaviors that led to the allegations of maltreatment. Occasionally service plans neglect specific concerns that ought to be addressed before reunification, because of lack of knowledge about risk factors, funding limitations, or other factors. Thus, the evaluator may make recommendations for other services (Condie, 2003).

Other records that tend to be useful include:

- Records of participation in and response to rehabilitation
- Mental health and substance abuse evaluation and treatment records
- Any records describing premorbid functioning (used to aid in determinations of deterioration or improvement in functioning)
- Educational and vocational records
- Medical records
- Visitation center records
- Police investigation reports relevant to child maltreatment or related offenses
- Frequency of substance-abuse related convictions

Rehabilitation nonattendance should not be taken at face value. Parents sometimes face transportation problems and other legitimate barriers to services (Ayoub & Kinscherff, 2006; Babiker & Herbert, 1998).

In advance of contact, the evaluator should develop the scope of questions appropriate for each collateral contact. Possible sources of information include teachers, child care workers, foster parents, residential placement staff members, extended family members, child maltreatment investigators, pediatricians, mental health and substance abuse service providers, and visitation supervisors (Stahl, 1994). Collecting information relevant to substance abuse rehabilitation for adults or psychotherapy of children sometimes may require a special release or court waiver because of higher privilege protections. Collateral contacts should be informed that the information they provide may appear in the forensic evaluation report. Releases should provide for a one-way release of information from the collateral contact to the evaluator unless a two-way release is indicated for some reason. A two-way release should not serve to compromise the integrity of the evaluation process or the roles even when the collateral sources are professional. The neutrality of the evaluator's position should be preserved (Condie, 2003).

Evaluating Children. The main goals in the initial appointment with a child are to set the child at ease, develop an understanding of the child's linguistic abilities, and provide a notification of the limits of confidentiality suitable to the child's comprehension (Condie & Koocher, 2008). It is helpful to begin with innocuous questions,

but the questions should not inadvertently confuse the child's understanding of the purpose of the evaluation. Similarly, the evaluator should not immediately launch into discourse or questions that will raise the child's anxiety about loyalty bonds with parents. The evaluator must be alert to the possibility that some children will have been notified in advance of the evaluator's role and evaluation goals, either with accurate information or misinformation. Thus, gleaning information from the child about his or her preconceived notions of the evaluation should take place at the outset. Some children may hold clear goals of what they wish to convey to the evaluator. An artful approach is required to determine if information provided by a child has been unduly influenced by other individuals due to recent contact, gifts, promises, or other methods of persuasion (Stahl, 1996).

A child-centered office environment helps set children at ease. Children should feel comfortable without becoming distracted. They should be allowed time to become accustomed to the evaluator and the context. Respect should be given to personal space, boundaries, and bodily integrity. Children should be encouraged to ask questions and seek clarification. It is important to avoid emotionally or morally laden phrases, such as "Bad things that happen to children." Developmentally, children are likely to blame themselves for "bad things," and they are unlikely to desire permanent separation from parents even when those parents have maltreated them (Condie, 2003). From their limited points of reference and experiences, "bad things" might be interpreted quite differently by children, or may pale in comparison to other events or qualities of individuals. There should be an assumption that their egocentric interpretation sometimes precludes comparisons and contrasts. Appropriate care should be used in designing questions that will allow children to voice their concerns without facing fear of moral or other approbations (Waterman, 1986).

Specific standards have been developed in some jurisdictions for audio- or video-recording interviews of children, particularly children whose families are involved in criminal or care and protection proceedings (Saywitz, 1994). Because of concern over the capacities of evaluators to record complete information in written form (Lamb, Orbach, Sternberg, Hershkowitz, & Horowitz, 2000), it is good practice to record interviews in some fashion, taking care to gather special permission in the informed consent process. An explanation of the use of the devices should be given in language the child comprehends. A contingency plan should be available for children who are intimidated by recording devices if it would compromise their willingness to provide relevant information. Recording increases the completeness of information, preserves information that might be used as legal evidence of abuse, promotes the use of proper interview techniques, records nonverbal facets of communication, and precludes or minimizes the need for multiple interviews (Lamb et al., 2000; Saywitz, 1994). Disadvantages include intrusiveness and possible compromises to children's willingness to divulge information, logistical and technological complications, loss of data through equipment malfunctions, focus on technique at the expense of issues of relevance, and release of recordings to

inappropriate sources such as the media (Berliner, 1992). In the absence of electronic recording, detailed written documentation is needed.

Interviewing Children. Whether children should be asked to provide demographic data depends on their age and level of linguistic development. The degree to which narrative accounts of maltreatment or other family interactions should be sought depends on their reporting capacities (Saywitz, 1994). When children cannot credibly report data, other sources of information must be relied on. When they can provide narrative accounts, their accounts should be compared to other reports and checked for consistency (Lamb et al., 2000). Inconsistency may reflect dissimulation, but it can occur for more innocuous reasons, such as a lack of appreciation by the child for salient details and insufficient developmental readiness to report a temporally organized narrative (Saywitz, 1994). Depending on the referral question, relevant content for child interviews may include a description of the child's view of family structure and relationships, other relationships important to the child, historical information (usually relevant only for preadolescents and adolescents), the child's view of his or her treatment needs and treatment progress, and the child's comprehension of the construct of trauma and its relevance or lack thereof to his or her life. Children are unskilled at providing details related to symptoms and behaviors of trauma reactions, chiefly because of their lack of comparative experience base and vocabulary for the terms and behaviors of relevance. Even when provided with symptom checklists, they may shy away from endorsing relevant items because they do not wish to view themselves as impaired. Even the best-designed measures for children contain terms that do not fall neatly within the linguistic capabilities of children (Condie, 2003).

There is no entirely flawless method of determining a child's capacity to provide accurate reports of maltreatment. Evaluators strive to minimize influences that might result in data that lack credibility, but it is important to remember that even the highest professional standards do not require an evaluator to be a good judge of a child's truth-telling capacity (APA Committee on Professional Practice and Standards, 2011). That task is left to the fact finder, and it lies beyond the scope of current scientific research and practice. When estimates of a child's capacity to report trauma are requested, they should be based on the best available empirical data. Examples include:

- Examining the child's account of maltreatment for the development of context
- Use of idiosyncratic words or descriptive phrases
- Inclusion of peripheral or unnecessary information
- Explicit details
- Details that exceed the child's developmental level
- A progression of "grooming" for maltreatment (seduction, isolation, escalation of threats and aggression)
- Other engagement processes

- Strategies designed to discourage the child from reporting maltreatment (secrecy, threats, coercion, pressure, bribes, rewards)
- Affective responses or details congruent with the reported maltreatment
- Consistency of salient details
- A narrative clearly emanating from a child's perspective rather than a rehearsed litany
- Details of attempts to resist or avoid the maltreatment (Heiman, 1992)

There are no pathognomonic signs of maltreatment, nor is there evidence that a particular type of interview response or set of tools or measures will yield data establishing that a child has been maltreated. Referral questions that go directly to this point should be rephrased in a professional consultation and negotiation process before the evaluation proceeds (Condie, 2003). Neither maltreatment nor the identity of a perpetrator can be confirmed or disconfirmed solely by the presence or absence of psychological symptoms or patterns of behavior.

When the child's psychological functioning is part of the referral question, interview data should focus on symptoms and behaviors of relevance to diagnostic criteria for child behavior disorders and trauma reactions (George & Solomon, 1999; Heiman, 1992). Because of the difficulty children have self-reporting data of relevance, it is important to include other sources of observation and information (Condie, 2003; Heiman, 1992). Measuring the impact of child maltreatment does not involve merely rendering a diagnosis. Descriptive information is needed about the impact of trauma on a particular child, the link between maltreatment and the child's reactions (if any), and the child's existing vulnerabilities (Everson & Faller, 2012).

When the child's view of parents, other caregivers, adaptation to placement, and substitute caregivers is central to a referral question, examiners must avoid any attempts to elicit abstract descriptions of relationships. Even when children have the capacities to respond meaningfully to questions about their relationships, their responses might be influenced by loyalty bonds, recent visitation with particular caregivers, and developmental limitations in making comparisons or appreciating potential alternatives to their own experiences (George & Solomon, 1999; Stahl, 1994). Evaluators should be prepared for some inconsistencies because of children's tendency to respond to recent events or points of contact, children's concerns about threats to their stability, distress reactions, conflicted views and ambivalence, and limitations in appreciation of temporal events or the passage of time ("a long time" to a child might be 5 minutes). Sometimes eagerness to reunify with a parent is merely a reflection of a child's indiscriminate attachment behavior (George & Solomon, 1999) or a desire to reunite with school friends (Stahl, 1994).

Psychological Measures. As with adults, assessment measures to be used in the evaluation of children should center on (1) the referral question, (2) the relevance of global and specific indices to the question, (3) theoretically and empirically derived

hypotheses, (4) the validity of the measures in the specific assessment context, and (5) whether the data would add meaningful utility to the evaluation process (Ayoub & Kinscherff, 2006; Barnum, 1997). The developmental readiness of a child for assessment participation is an added consideration. Even when measures are designed for specific age ranges, children sometimes may not be developmentally, cognitively, or linguistically prepared for the process (Condie, 2003). Assessment measures do sometimes yield useful data on the child's capacity to report information of relevance, to benefit from relevant treatment, or to tolerate a foster placement (Everson & Faller, 2012). If adequate pretreatment data are available, it is sometimes possible to measure treatment progress using psychological assessment measures. Assessment measures can highlight these issues in a child:

- Strengths and weakness
- Approach to relationships
- Level of trust in individuals in roles of authority
- Willingness to engage in treatment
- Linguistic capacity to provide a narrative
- Mental health functioning
- Views of helping sources and friendships

Measures sometimes illustrate why a child has had a poor or failed response to a particular treatment approach, why a child might distort reports of relationships or events, or why a child might show a relative lack of resilience in the recovery process (Condie, 2003). As with the evaluation of parents, specific measures relevant to a child's view of parent-child interactions, attachments to parents, and other specific factors should be used and interpreted conservatively unless specific norms are available for the population of interest.

CONCLUSIONS

During any phase of a child protection proceeding, a psychologist may be asked to evaluate different parties for different purposes. As evaluators, psychologists frequently are asked to address these and other issues:

- The impact of child maltreatment
- The risk that it might recur
- How seriously the child's well-being has been affected
- What therapeutic or intervention strategies would be recommended to assist the child and/or family
- Whether parents or other caregivers can be rehabilitated such that risk of maltreatment is reduced
- What the psychological effect on the child would be if the child were returned to parents or other caregivers
- What the psychological effect on the child would be if parental rights were terminated

To understand risk of maltreatment, it is important to understand research on a variety of factors contributing to risk and mediation of risk. Psychologists seek to gather information on:

- Family history
- Personality functioning
- Social and other contextual circumstances
- Developmental needs of the child
- Nature and quality of the parent–child relationship
- Reactions to trauma
- A variety of factors contributing to risk of child maltreatment

They seek to understand risk in the context of sociocultural factors, physical disability, and other extenuating factors of relevance. Evaluation methodology, data interpretation, and procedures for reaching recommendations are derived from codes of ethics, standards of practice, and relevant research literature. Multimodal assessment is the primary buffer against data misinterpretation, overinterpretation, or underinterpretation. Interpreting interview and assessment data may occur in actuarial methods or the context of the examinee’s history. Both approaches facilitate meaningful data interpretation. Risk assessment matrices should include factors identified in empirical studies of risk assessment that are relevant to samples of parents involved in the care and protection system. Analysis of child maltreatment risk should acknowledge appropriate caveats. Further research is needed to better understand the degree of concordance or possible discordance in risk studies relevant to other samples of individuals and those involving risk of child maltreatment.

Although many existing measures and methods are designed to assess the nature and quality of the parent–child relationships, parent–child attachment, and parent–child interactions, their applicability to care and protection cases depends on the availability of relevant supplementary norms. Data interpretation and recommendations made via multimethod approaches that incorporate specific parenting measures should include appropriate cautionary procedures and comments. Similarly, global measures of functioning should be used when judged to be appropriate based on the referral question and other relevant considerations related to reliability and validity of application to care and protection samples. Dissimulation is an issue that is potentially endemic to care and protection evaluations, but methods for detecting dissimulation that are specific to care and protection samples have not been developed. Evaluators should make reasonable efforts to detect dissimulation but without overreliance on measure-specific methods that have no demonstrated validity or reliability in care and protection samples. Methods for minimizing the influence of children’s suggestibility and other impediments to reliability and validity should be used when indicated. Many care and protection cases involve children with cognitive limitations, mental health issues, and other special needs. Assessment methods and procedures should be developed on a case-by-case basis.

Appropriate modifications should be made when needed. Novel procedures should not be used in forensic cases unless they reflect converging professional consensus, research, and scientifically based judgment. The breadth and depth of interview content and indications for the use of forensic assessment measures are drawn from the referral question.

Key approaches to data integration and organization of presentation include (1) providing a specific answer to referral questions (when results are inconclusive, it is best to say so directly), (2) using theory as a template to guide data integration and interpretation, (3) interpreting data in light of the examinee's history, and (4) describing the strengths and limitations of the data. Relevant risk factors should be described in terms of their static and dynamic nature. Mediators and protective factors should be included in any risk analysis. Some risk factors relevant to child maltreatment vary, depending on the type of child maltreatment. Most risk factors are nonspecific. In studies of risk factors, it is difficult to control for concurrent types of maltreatment and their influence on research results.

Recommendations for service plan interventions and modifications sometimes must take statutory provisions about availability of services into account. The statutorily defined need to provide only those services that are available poses a challenge for evaluators asked to make recommendations for optimal intervention approaches. Specific recommendations tend to be more useful than general recommendations. For example, a recommendation for a specific form of intervention for a parent with a specific set of circumstances, symptoms, or problematic behaviors is more useful than a broad recommendation for mental health treatment. Recommendations concerning parental or caregiver amenability to rehabilitation often must be given with statutory time frames for service provision in mind. Statutory time limits for successful rehabilitation pose a challenge for parents who learn at a slow pace, who have transportation or other financial limitations, or who face other challenges to rapid treatment progress. Interpretations and recommendations for children should be made in the context of their levels of developmental maturity, their capacities to benefit from recommended interventions, and any special needs they might have.

REFERENCES

- Affi, T., & MacMillan, H. (2011). Resilience following child maltreatment: A review of protective factors. *Canadian Journal of Psychiatry, 56*, 266–272.
- Alia-Klein, N., Goldstein, R., Tomasi, D., Wiocik, P., Moeller, S., Williams, B., . . . Volkow, N. (2009). Neural mechanisms of anger regulation as a function of genetic risk for violence. *Emotion, 9*, 385–396.
- Alink, L., Cicchetti, D., Kim, J., & Rogosch, F. (2012). Longitudinal associations among child maltreatment, social functioning, and cortisol regulation. *Developmental Psychology, 48*, 224–236.
- American Psychological Association. (2013). Specialty guidelines for forensic psychology. *American Psychologist, 68*(1), 7–19. doi:10.1037/a0029889

- American Psychological Association Committee on Professional Practice and Standards. (2011). *Guidelines for evaluations in child protection matters*. Washington, DC: Author.
- Appleyard, K., Berlin, L., Rosanbalm, K., & Dodge, K. (2011). Preventing early child maltreatment: Implications from a longitudinal study of maternal abuse history, substance use problems, and offspring victimization. *Prevention Science, 12*, 139–149.
- Appleyard, K., Yang, C., & Runyan, D. (2010). Delineating the maladaptive pathways of child maltreatment: A mediated moderation analysis of the roles of self-perception and social support. *Development and Psychopathology, 22*, 337–352.
- Ayoub, C., & Kinscherff, R. (2006). Forensic assessment of parenting in child abuse and neglect cases. In S. Sparta & G. Koocher (Eds.), *Forensic mental health assessment of children and adolescents* (pp. 330–341). New York, NY: Oxford University Press.
- Babiker, G., & Herbert, M. (1998). Critical issues in the assessment of child sexual abuse. *Clinical Child and Family Psychology Review, 1*, 231–252.
- Bae, Hwa-ok, Solomon, P. L., Gelles, R., & White, T. (2010). Effect of child protective services system factors on child maltreatment reporting. *Child Welfare: Journal of Policy, Practice, and Program, 89*, 33–56.
- Baer, J., & Martinez, C. (2006). Child maltreatment and insecure attachment: A meta-analysis. *Journal of Reproductive and Infant Psychology, 24*, 187–197.
- Banyard, V., Williams, L., & Siegel, J. (2003). The impact of complex trauma and depression on parenting: An exploration of mediating risk and protective factors. *Child Maltreatment, 8*, 334.
- Barnum, R. (1997). A suggested framework for forensic consultation in cases of child abuse and neglect. *Journal of the American Academy of Psychiatry and Law, 25*, 581–593.
- Bauer, W. D., & Twentyman, C. T. (1985). Abusing, neglectful, and comparison mothers' responses to child-related and non-child-related stressors. *Journal of Consulting and Clinical Psychology, 53*, 335–343.
- Belsky, J. (1993). Etiology of child maltreatment: A developmental-ecological analysis. *Psychological Bulletin, 114*, 413–434.
- Belsky, J., Conger, R., & Capaldi, D. (2009). The intergenerational transmission of parenting. *Developmental Psychology, 45*, 1201–1204.
- Belsky, J., & Jaffee, S. (2006). The multiple determinants of parenting. In D. Cicchetti & D. Cohen (Eds.), *Developmental psychopathology: Vol. 3. Risk, disorder and adaptation* (2nd ed., pp. 38–85). Hoboken, NJ: Wiley.
- Belsky, J., & Pensky, E. (1988). Developmental history, personality and family relationships: Toward an emergent family system. In R. Hinde & J. Stevenson-Hinde (Eds.), *Relationships within families* (pp. 193–217). Oxford, England: Clarendon Press.
- Berliner, L. (1992). Should investigative interviews of children be videotaped? *Journal of Interpersonal Violence, 7*, 277–288.
- Borum, R., Otto, R., & Golding, S. (1993). Improving clinical judgment and decision making in forensic evaluation. *Journal of Psychiatry & Law, 21*, 35–76.
- Bowlby, J. (1988). *A secure base: Parent-child attachment and healthy human development*. London, England: Routledge.
- Brackenridge, C., Bishopp, D., Moussalli, S., & Tapp, J. (2008). The characteristics of sexual abuse in sport: A multidimensional scaling analysis of events described in media reports. *International Journal of Sport and Exercise Psychology, 6*, 385–406.

- Brandon, P. (2001). State intervention in imperfect families. *Rationality and Society*, 13, 285–303.
- Brent, B., & Corwyn, R. (2002). The effect of abuse in childhood and adolescence on violence among adolescents. *Youth & Society*, 33, 339–365.
- Brodzinsky, D. (1993). On the use and misuse of psychological testing in child custody evaluation. *Professional Psychology: Research and Practice*, 24, 213–219.
- Brooks-Gunn, J., & Duncan, G. (1997). The effects of poverty on children. *Future of Children*, 7, 55–71.
- Budd, K., Connell, M., & Clark, J. (2011). *Evaluation of parenting capacity in child protection*. New York, NY: Oxford University Press.
- Budd, K., & Holdsworth, M. (1996). Issues in clinical assessment of minimal parenting competence. *Journal of Clinical Child Psychology*, 25, 2–14.
- Budd, K., Poindexter, L., Felix, E., & Naik-Polan, A. (2001). Clinical assessment of parents in child protection cases: An empirical analysis. *Law and Human Behavior*, 25, 93–108.
- Bugental, D., Blue, J., & Lewis, J. (1990). Caregiver beliefs and dysphoric affect directed to difficult children. *Developmental Psychology*, 26, 631–638.
- Burgess, R. L., & Draper, P. (1989). The explanation of family violence: The role of biological, behavioral, and cultural selection. In L. Ohlin & M. Tonry (Eds.), *Family violence* (pp. 59–116). Chicago, IL: University of Chicago Press.
- Caliso, J., & Milner, J. (1992). Childhood history of abuse and child abuse screening. *Child Abuse and Neglect*, 16, 647–659.
- Cantos, A., Gries, L., & Slis, V. (1997). Behavioral correlates of parental visiting during family foster care. *Child Welfare*, 76, 309–329.
- Capaldi, D., Pears, K., Patterson, G., & Owen, L. (2003). Continuity of parenting practices across generations in an at-risk sample: A prospective comparison of direct and mediated associations. *Journal of Abnormal Child Psychology*, 31, 127–142.
- Centers for Disease Control and Prevention. (2012). *Cost of child abuse and neglect rival other major public health problems*. Washington, DC: Author.
- Chen, Z., Liu, R., & Kaplan, H. (2008). Mediating mechanisms for the intergenerational transmission of constructive parenting: A prospective longitudinal study. *Journal of Family Issues*, 29, 1574–1599.
- Chuang, E., Moore, K., Barrett, B., & Young, S. (2012). Effect of an integrated family dependency treatment court on child welfare reunification, time to permanency and re-entry rates. *Children and Youth Services Review*, 34(9), 1896–1902.
- Cicchetti, D. (2004). An odyssey of discovery: Lessons learned through three decades of research on child maltreatment. *American Psychologist*, 59, 731–741.
- Cicchetti, D., & Rogosch, F. (2012). Gene x environment interaction and resilience: Effects of child maltreatment and serotonin, corticotropin releasing hormone, dopamine, and oxytocin genes. *Development & Psychopathology*, 24, 411–427.
- Cohen, L., Frenda, S., Mojtabai, R., Katsavdakakis, K., & Galynker, I. (2007). Comparison of sexual offenders against children with sexual offenders against adolescents and adults: Data from the New York State Sex Offender Registry. *Journal of Practical Psychiatry and Behavioral Health*, 13, 373–384.
- Collishaw, S., Pickles, A., Messer, J., Rutter, M., Shearer, C., & Maughan, B. (2007). Resilience to adult psychopathology following childhood maltreatment: Evidence from a community sample. *Child Abuse & Neglect*, 31, 211–229.

- Colman, R., Kim, D., Mitchell-Herzfeld, S., & Shady, T. (2009). Delinquent girls grow up: Young adult offending patterns and their relation to early legal, individual, and family risk. *Journal of Youth and Adolescence, 38*, 355–366.
- Condie, L. (2003). *Parenting evaluations for the court: Care and protection matters. Vol. 18: Perspectives in law and psychology*. New York, NY: Kluwer Academic/Plenum Press.
- Condie, L., & Koocher, G. (2008). Clinical management of children's incomplete comprehension of confidentiality limits. *Journal of Child Custody: Research, Issues, and Practices, 5*, 161–191.
- Conger, R., Neppl, T., Kim, K., & Scaramella, L. (2003). Angry and aggressive behaviour across three generations: A prospective, longitudinal study of parents and children. *Journal of Abnormal Child Psychology, 31*, 143–160.
- Connors, G., Longabaugh, R., & Miller, W. (1996). Looking forward and back to relapse: Implications for research and practice. *Addiction, 91*, S191–S196.
- Corse, S., Schmid, K., & Trickett, P. (1990). Social network characteristics of mothers in abusing and nonabusing families and their relationships to parenting beliefs. *Journal of Community Psychology, 18*, 44–59.
- Coulton, C., Korbin, J., & Su, M. (1999). Neighborhoods and child maltreatment: A multi-level study. *Child Abuse and Neglect, 23*, 1019–1040.
- Crenshaw, W., & Barnum, D. (2001). You can't fight the system: Strategies of family justice in foster care reintegration. *Family Journal: Counseling & Therapy for Couples & Families, 9*, 29–36.
- Crittenden, P. M. (1985). Social networks, quality of child rearing, and child development. *Child Development, 56*, 1299–1313.
- Cyr, C., Bakermans-Kranenburg, M., & van Ijzendoorn, M. (2010). Attachment security and disorganization in maltreating and high-risk families: A series of meta-analyses. *Development and Psychopathology, 22*, 87–108.
- De Bellis, M. (2012). The intergenerational transmission of family violence: The neurobiology of the relationships among child victimization, parental mental health, and addiction. In C. Widom (Ed.), *Trauma, psychopathology, and violence: Causes, consequences, or correlates?* (pp. 79–108). New York, NY: Oxford University Press.
- Dixon, L., & Hamilton-Giachritsis, C. (2009). Patterns of risk and protective factors in the intergenerational cycle of maltreatment. *Journal of Family Violence, 24*, 111–122.
- Dixon, L., Hamilton-Giachritsis, C., Browne, K., & Ostapuk, E. (2007). The co-occurrence of child and intimate partner maltreatment in the family: Characteristics of the violent perpetrators. *Journal of Family Violence, 22*, 675–689.
- Dubowitz, H., Hampton, R., Bithoney, W., & Newberger, E. (1987). Inflicted and non-inflicted injuries: Differences in child and familial characteristics. *American Journal of Orthopsychiatry, 57*, 525–535.
- Egeland, B., Jacobvitz, D., & Papatola, K. (1987). Intergenerational continuity of abuse. In R. Gelles & J. Lancaster (Eds.), *Child abuse and neglect: Biosocial dimensions* (pp. 255–276). Chicago, IL: Aldine.
- Ertem, I. O., Leventhal, J. M., & Dobbs, S. (2000). Intergenerational continuity of child physical abuse: How good is the evidence? *Lancet, 356*, 814–819.
- Everson, M., & Faller, K. (2012). Base rates, multiple indicators, and comprehensive forensic evaluations: Why sexualized behavior still counts in assessments of child sexual abuse

- allegations. *Journal of Child Sexual Abuse: Research, Treatment, & Program Innovations for Victims, Survivors, & Offenders*, 21, 45–71.
- Feshbach, N. (1989). The construct of empathy and the phenomenon of physical maltreatment of children. In D. Cicchetti & V. Carlson (Eds.), *Child maltreatment: Theory and research on the causes and consequences of child abuse and neglect* (pp. 349–373). New York, NY: Cambridge University Press.
- Garbarino, J., & Sherman, D. (1980). High-risk neighborhoods and high-risk families: The human ecology of child maltreatment. *Child Development*, 51, 188–198.
- Gaudin, J. (1994). *Child neglect: A guide for intervention*. Washington, DC: U.S. Department of Health and Human Services.
- Gaudin, J., & Dubowitz, H. (1997). Family functioning in neglectful families: Recent research. In J. Berrick & N. Barth (Eds.), *Child welfare research review* (Vol. 2, pp. 28–36). New York, NY: Columbia University Press.
- Gelfand, D., & Teti, D. (1990). The effects of maternal depression on children. *Clinical Psychology Review*, 10, 329–353.
- George, C., & Solomon, J. (1999). Attachment and caregiving: The caregiving behavioral system. In J. Cassidy & P. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (pp. 649–670). New York, NY: Guilford Press.
- Gibbs, B. (2002). Child maltreatment and negative cognitive styles: A quantitative and qualitative review. *Clinical Psychology Review*, 22, 223–246.
- Goodman, S., & Tully, E. (2008). Children of depressed mothers: Implications for the etiology, treatment, and prevention of depression in children and adolescents. In J. Abela & B. Hankin (Eds.), *Handbook of depression in children and adolescents* (pp. 415–440). New York, NY: Guilford Press.
- Hanson, T., & Thornton, D. (2000). Improving risk assessment for sex offenders: A comparison of three actuarial scales. *Law & Human Behavior*, 24, 119–136.
- Hardt, J., & Rutter, M. (2004). Validity of adult retrospective reports of adverse childhood experiences: Review of the evidence. *Journal of Child Psychology and Psychiatry*, 45, 260–273.
- Heiman, M. (1992). Putting the puzzle together: Validating allegations of child sexual abuse. *Journal of Child Psychology and Psychiatry*, 33, 311–329.
- Herrenkohl, T., & Herrenkohl, R. (2007). Examining the overlap and prediction of multiple forms of child maltreatment, stressors, and socioeconomic status: A longitudinal analysis of youth outcomes. *Journal of Family Violence*, 22, 553–562.
- Herrenkohl, E., Herrenkohl, R., & Toedter, L. (1983). Perspectives on the intergenerational transmission of abuse. In D. Finkelhor, R. Gelles, G. Hotaling, & M. Straus (Eds.), *The dark side of families* (pp. 305–316). Beverly Hills, CA: Sage.
- Herring, D. (2009). Fathers and child maltreatment: A research agenda based on evolutionary theory and behavioral biology research. *Children and Youth Services Review*, 31, 935–945.
- Howes, P., Cicchetti, D., Toth, S., & Rogosch, F. (2000). Affective, organizational, and relational characteristics of maltreating families: A systems perspective. *Journal of Family Psychology*, 14, 95–110.
- Jaudes, P., & Mackey-Bilaver, L. (2008). Do chronic conditions increase young children's risk of being maltreated? *Child Abuse & Neglect*, 32, 671–681.
- Johnston, J., & Goldman, J. (2010). Outcomes of family counseling interventions with children who resist visitation. *Family Court Review*, 48, 112–115.

- Kim, J., Cicchetti, D., Rogosch, F., & Manly, J. (2009). Child maltreatment and trajectories of personality and behavioral functioning: Implications for the development of personality disorder. *Development and Psychopathology, 21*, 889–912.
- Kohl, P., Jonson-Reid, M., & Drake, B. (2011). Maternal mental illness and safety and stability of maltreated children. *Child Abuse & Neglect, 35*, 309–318.
- Korbin, J., Coulton, C., Chard, S., Platt-Houston, C., & Su, M. (1998). Impoverishment and child maltreatment in African American and European American neighborhoods. *Development and Psychopathology, 10*, 215–233.
- Lamb, M., Orbach, Y., Sternberg, K., Hershkowitz, I., & Horowitz, D. (2000). Accuracy of investigators' verbatim notes of their forensic interviews with alleged child abuse victims. *Law and Human Behavior, 24*, 699–708.
- Li, F., Godinet, M., & Arnsberger, P. (2011). Protective factors among families with children at risk of maltreatment: Follow up to early school years. *Children and Youth Services Review, 33*, 139–149.
- Lodewijks, H., de Ruiter, C., & Doreleijers, T. (2010). The impact of protective factors in desistance from violent reoffending: A study in three samples of adolescent offenders. *Journal of Interpersonal Violence, 25*, 568–587.
- Loeber, R., Pardini, D., Stouthamer-Loeber, M., & Raine, A. (2007). *Development and Psychopathology, 19*, 867–887.
- MacKinnon-Lewis, C., Lamb, M., Arbuckle, B., Baradoran, L., & Veiling, B. (1992). The relationship between biased maternal and filial attributions and the aggressiveness of their interactions. *Development and Psychopathology, 4*, 403–415.
- MacMillan, H. (2005). Family violence research: Lessons learned and where from here? *Journal of the American Medical Association, 29*, 618–620.
- Main, M., & Goldwyn, R. (1984). Predicting rejection of her infant from mother's representation of her own experience: Implications for the abused-abusing intergenerational cycle. *Child Abuse & Neglect, 8*, 203–217.
- Main, M., Kaplan, N., & Cassidy, J. (1985). Security in infancy, childhood and adulthood: A move to level of representation. In I. Bretherton & E. Waters (Eds.), *Growing points in attachment theory and research. Monographs for the Society for Research in Child Development, 50* (1–2, Serial No. 209).
- Malkin, C., & Lamb, M. (1994). Child maltreatment: A test of sociobiological theory. *Journal of Comparative Family Studies, 25*, 121–133.
- Marshall, N. (2012). Recognizing and reporting parental psychological maltreatment of children. *Professional Psychology: Research and Practice, 2*, 73–79.
- Melton, G. (1992). It's time for neighborhood research and action. *Child Abuse & Neglect, 16*, 909–913.
- Mennen, F., Kim, K., Sang, J., & Trickett, P. (2010). Child neglect: Definition and identification of youth's experiences in official reports of maltreatment. *Child Abuse & Neglect, 34*, 647–658.
- Miller, T., Cohen, M., & Wiersema, B. (1996). *Victim costs and consequences: A new look*. Washington, DC: National Institute of Justice.
- Neutze, J., Grundmann, D., Scherner, G., & Beier, K. (2012). Undetected and detected child sexual abuse and child pornography offenders. *International Journal of Law and Psychiatry, 35*, 168–175.

- Ownby, R. (1997). *Psychological reports: A guide to report writing in professional psychology* (3rd ed.). New York, NY: Wiley.
- Pearson, J., & Theonnes, N. (2000). Supervised visitation: The families and their experiences. *Family & Conciliation Courts Review, 38*, 123–142.
- Pidgeon, A., & Sanders, M. (2009). Attributions, parental anger and risk of maltreatment. *International Journal of Child Health and Human Development, 2*, 57–69.
- Plomin, R., Nitz, K., & Rowe, D. (1990). Behavioral genetics and aggressive behavior in childhood. In M. Lewis & S. Miller (Eds.), *Handbook of developmental psychopathology* (pp. 119–133). New York, NY: Plenum Press.
- Putman-Hornstein, E., Webster, D., Needed, B., & Magruder, J. (2011). A public health approach to child maltreatment surveillance: Evidence from a data linkage project in the United States. *Child Abuse Review, 20*, 256–273.
- Quinsey, F., Harris, G., Rice, M., & Cormier, C. (1998). Fifteen arguments against actuarial risk appraisal. In V. Quinsey, G. Harris, M. Rice, & C. Cormier (Eds.), *Violent offenders: Appraising and managing risk* (pp. 171–190). Washington, DC: American Psychological Association.
- Quinton, D., Rutter, M., & Liddle, C. (1984). Institutional rearing, parenting difficulties, and marital support. *Psychological Medicine, 14*, 107–124.
- Ronay, K. (2011). The visit: Observing children's experience of visiting a relative in prison. *Infant Observation, 14*, 191–202.
- Runyan, D., Cox, C., Dubowitz, H., Newton, R., Upadhyaya, M., Kotch, J.,... Knight, E. (2005). Describing maltreatment: Do child protective service reports and research definitions agree? *Child Abuse & Neglect, 29*, 461–477.
- Russo, G. (2008). *Child abuse and neglect cost nation over \$100 billion per year*. Washington, DC: Pew Charitable Trusts.
- Saywitz, K. (1994). Effects of a multidisciplinary interview center on the investigation of alleged child sexual abuse. *Violence Update, 3*(3), 6.
- Schulz-Heik, R., Jay Rhee, S., Silvern, L., Haberstick, B., Hopfer, C., Lessem, J., & Hewitt, J. (2010). The association between conduct problems and maltreatment: Testing genetic and environmental mediation. *Behavior Genetics, 40*, 338–348.
- Sgroi, S. (1989). *Vulnerable populations: Sexual abuse treatment for children, adult survivors, offenders, and persons with mental retardation*, Vol. 2. Lexington, MA: Lexington Books/D.C. Heath.
- Simon, R., Lei, M., Stewart, E., Beach, S., Brody, G., Philibert, R., & Gibbons, F. (2012). Social adversity, genetic variation, street code, and aggression: A genetic model of violent behavior. *Youth Violence and Juvenile Justice, 10*, 3–24.
- Simon, R., Whitbeck, L., Conger, R., & Wu, C. (1991). Intergenerational transmission of harsh parenting. *Developmental Psychology, 27*, 159–171.
- Stahl, P. (1994). *Conducting child custody evaluations: A comprehensive guide*. Thousand Oaks, CA: Sage.
- Stahl, P. (1996). Second opinions: An ethical and professional process for reviewing child custody evaluations. *Family and Conciliation Courts Review, 34*(3), 386–395.
- Terry, K., Schuth, K., & Smith, M. (2011). Incidence of clerical sexual abuse over time: Changes in behavior and seminary training between 1950 and 2008. In T. Plante &

- K. McChesney (Eds.), *Sexual abuse in the Catholic Church: A decade of crisis, 2002–2012*. Santa Barbara, CA: Praeger/ABC-CLIO.
- Trickett, P. K., & Kuczynski, L. (1986). Children's misbehaviors and parental discipline strategies in abusive and nonabusive families. *Developmental Psychology, 22*, 115–123.
- U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau (2011). *Child Maltreatment 2010*. Washington, DC: Author.
- Valentine, D. (1990). Double jeopardy: Child maltreatment and mental retardation. *Child & Adolescent Social Work Journal, 7*, 487–499.
- van Ijzendoorn, M. (1995). Adult attachment representation, parental responsiveness, and infant attachment: A meta-analysis on the predictive validity of the Adult Attachment Interview. *Psychological Bulletin, 117*, 387–403.
- Wahler, R., & Dumas, J. (1986). Maintenance factors in coercive mother-child interactions: The compliance and predictability hypotheses. *Journal of Applied Behavior Analysis, 19*, 13–22.
- Wang, C., & Holton, J. (2007). *Total estimated cost of child abuse and neglect in the United States*. Chicago, IL: Prevent Child Abuse America.
- Waterman, J. (1986). Family dynamics of incest with young children. In K. MacFarlane, J. Waterman, S. Conerly, L. Damon, M. Durfee, & S. Long (Eds.), *Sexual abuse of young children* (pp. 204–219). New York, NY: Guilford Press.
- Watson, D., & Clark, L. (1992). On traits and temperament: General and specific factors of emotional experience and their relation to the five-factor model. *Journal of Personality, 60*, 441–476.
- Weiss, R., & Rosenfeld, B. (2012). Navigating cross-cultural issues in forensic assessment: Recommendations for practice. *Professional Psychology: Research and Practice, 43*, 234–240.
- Whipple, E. E., & Webster-Stratton, C. (1991). The role of parental stress in physically abusive families. *Child Abuse and Neglect, 15*, 279–291.
- Widom, C. (1989). Does violence beget violence? A critical examination of the literature. *Psychological Bulletin, 106*, 3–28.
- Wolfe, D. (1985). Child-abusive parents: An empirical review and analysis. *Psychological Bulletin, 97*, 462–482.
- Youngblade, L., & Belsky, J. (1989). Child maltreatment, infant-parent attachment security, and dysfunctional peer relationships in toddlerhood. *Topics in Early Childhood Special Education, 9*, 1–15.
- Zuravin, S. J., & Greif, G. L. (1989). Normative and child-maltreating AFDC mothers. *Social Casework: The Journal of Contemporary Social Work, 74*, 76–84.

PART THREE

APPLYING PSYCHOLOGY
TO CRIMINAL PROCEEDINGS

Assessing Competency to Stand Trial

PATRICIA A. ZAPF, RONALD ROESCH, AND GIANNI PIRELLI

COMPETENCY to stand trial is a concept of jurisprudence allowing the postponement of criminal proceedings for those defendants who are unable to participate in their defense on account of mental disease or intellectual disability. Because trial competency issues are raised substantially more often than the insanity defense, psychologists involved in forensic assessment and consultation are likely to have more experience with competency evaluations than those of criminal responsibility. Estimates are that approximately 60,000 competency evaluations are conducted in the United States annually (Bonnie & Grisso, 2000). This number has increased substantially from estimates in 1997 that placed the annual number of competency evaluations between 25,000 and 39,000 (Hoge et al., 1997). In this chapter, we present an overview of competency laws, research, and methods of assessment, with the aim of providing forensic psychologists with the basic information necessary to conduct competency evaluations. We do not believe, however, that this chapter will sufficiently prepare a novice forensic psychologist to carry out such evaluations. The issues surrounding competency determinations are highly complex; therefore, an evaluator needs not only a high level of clinical knowledge and skill but also a sophisticated knowledge of the legal system.

We urge readers interested in pursuing work in the competency arena to supplement this chapter with additional sources (e.g., Bonnie, 1993; Grisso, 2003; Melton et al., 2007; Stafford, 2003; Winick, 1996; Zapf & Roesch, 2009; Zapf, Viljoen, Whittemore, Poythress, & Roesch, 2002) as well as workshops and other forms of continuing education. The Specialty Guidelines for Forensic Psychology (American Psychological Association [APA], 2013; the Guidelines are reprinted as the appendix to this volume with permission of the APA) also contain discussions relevant to competency evaluations.

DEFINING COMPETENCY

Provisions allowing for a delay of trial because a defendant was incompetent to proceed have long been a part of the judicial system's due process standards. English common law allowed for an arraignment, trial, judgment, or execution of an alleged capital offender to be stayed if he or she "be(came) absolutely mad" (Hale, 1736, cited in Silten & Tullis, 1977, p. 1053). Over time, statutes that have been created in the United States and Canada have further defined and extended the common law practice (see S. Davis, 1994; Rogers & Mitchell, 1991; see Verdun-Jones, 1981; and Webster, Menzies, & Jackson, 1982 for reviews of Canadian competency law and practice). The modern standard in the United States was established in *Dusky v. United States* (1960). Although the wording differs across jurisdictions, all states use a variant of the *Dusky* standard to define competency (Favole, 1983). In *Dusky*, the Supreme Court held:

It is not enough for the district judge to find that "the defendant is oriented to time and place and has some recollection of events," but that the test must be whether he has sufficient present ability to consult with his lawyer with a reasonable degree of rational understanding—and whether he has a rational as well as factual understanding of the proceedings against him. (p. 402)

While the concept of competency to stand trial has been long established in law, ambiguities in the wording of *Dusky* raise a number of questions. What is meant by "sufficient present ability"? How does one determine whether a defendant "has a rational as well as factual understanding"? To be sure, some courts (e.g., *Wieter v. Settle*, 1961) and legislatures (e.g., Utah Code Annotated, §77-15-1 *et seq.*, 1994) have provided some direction to evaluators in the form of articulated *Dusky* standards (discussed next), but the forensic evaluator is left largely unguided except by a common principle, that evaluators cannot reach a finding of incompetency independent of the facts of the legal case (an issue to which we return later).

The problem in defining and assessing competency leads to a broad range of interpretations of the *Dusky* standard. Because the courts and legislatures have given mental health professionals a large share of the responsibility for defining and evaluating competency, it should not be surprising to find that mental status issues, such as presence or absence of psychosis, traditionally have played a dominant role in the findings of evaluators. Historically, in fact, evaluators initially equated psychosis with incompetency (McGarry, 1965; Roesch & Golding, 1980). Furthermore, evaluators in the past rarely took into account the specific demands of a defendant's case.

Practice standards are more clearly delineated at present, and, as a result, practice in this area continues to improve. In the past, evaluators typically were employed in state psychiatric hospital settings (wherein the majority of competency evaluations were conducted) and were not formally trained in the assessment of competency or in matters of the law. As a consequence, the evaluations were based on the same

standard mental status examinations that had been used with other patients in the hospital. Psychological tests were rarely used; if they were employed, they were utilized as a diagnostic tool to determine the presence or absence of psychosis.

Over the past 35 years, these practices have been challenged and improved based on empirical research findings. For example, research has provided evidence that the presence of psychosis was not sufficient by itself for a finding of incompetency (Roesch & Golding, 1980), and modern empirical studies of competency reports have demonstrated that evaluators rarely make that simple conceptual error (Heilbrun & Collins, 1995; Nicholson & Norwood, 2000; Skeem, Golding, Cohn, & Berge, 1998). Nevertheless, although forensic evaluators today typically have more training than those in the past, most states still do not require specific training of mental health professionals who conduct such evaluations (Farkas, DeLeon, & Newman, 1997).

The specific psycholegal abilities required of a defendant are the most important aspect of assessing fitness. The contextual nature of competence has been explored. Some researchers and scholars have argued that competence should be considered within the context in which it is to be used. For instance, the abilities required by the defendant in his or her specific case should be taken into account when assessing competence. This contextual perspective was summarized by Golding and Roesch (1988):

Mere presence of severe disturbance (a psychopathological criterion) is only a threshold issue—it must be further demonstrated that such severe disturbance in *this* defendant, facing *these* charges, in light of existing evidence, anticipating the substantial effort of a *particular* attorney with a *relationship of known characteristics*, results in the defendant being unable to rationally assist the attorney or to comprehend the nature of the proceedings and their likely outcome. (p. 79)

The importance of a contextual determination of specific psycholegal abilities has been repeatedly demonstrated by empirical findings that abilities in one area of functioning are rarely homogenous with those in other areas of functioning (Bonnie, 1992a; Golding & Roesch, 1988; Grisso, Appelbaum, Mulvey, & Fletcher, 1995; Skeem et al., 1998).

Supreme Court decisions in both the United States and Canada, however, have confused this issue by finding that the standard by which competency to be judged is not context-specific. In *Regina v. Whittle* (1994), the Supreme Court of Canada ruled that there is to be only one standard for competency regardless of the specific abilities to be performed by the accused. That Court concluded that there is no difference between the essential abilities needed in making active choices about waiving counsel, making decisions at trial, confessing, or pleading guilty. It ruled that different standards of competency should not be applied for different aspects of criminal proceedings and that the test to be used is one of “limited cognitive capacity” (p. 567) in each of these circumstances. However, unlike in *Godinez v. Moran* (1993), the forensic examiners had actually evaluated Mr. Whittle in these

specific contexts, regardless of whether the standard to be applied was the same or different as a function of the context.

In *Godinez v. Moran* (1993), the United States Supreme Court held similarly that the standard for the various types of competency (e.g., competency to plead guilty, to waive counsel, to stand trial) should be considered the same. Justice Thomas wrote for the majority:

The standard adopted by the Ninth Circuit is whether a defendant who seeks to plead guilty or waive counsel has the capacity for “reasoned choice” among the alternatives available to him. How this standard is different from (much less higher than) the Dusky standard—whether the defendant has a “rational understanding” of the proceedings—is not readily apparent to us. . . . While the decision to plead guilty is undeniably a profound one, it is no more complicated than the sum total of decisions that a defendant may be called upon to make during the course of a trial. . . . Nor do we think that a defendant who waives his right to the assistance of counsel must be more competent than the defendant who does not, since there is no reason to believe that the decision to waive counsel requires an appreciably higher level of mental functioning than the decision to waive other constitutional rights. (p. 2686)

In his dissent, Justice Blackmun noted that the “majority’s analysis is contrary to both common sense and long-standing case law” (p. 2691) and reasoned that competency cannot be considered in a vacuum, separate from its specific legal context. Justice Blackmun argued that “competency for one purpose does not necessarily translate to competency for another purpose” and noted that previous Supreme Court cases “required competency evaluations to be specifically tailored to the context and purpose of a proceeding” (p. 2694). What is missing from the majority’s opinion in *Godinez*, however, is that, unlike the circumstances in *Whittle*, Moran’s competency to waive counsel or plead guilty to death penalty murder charges was never assessed by the forensic examiners, regardless of which standard (reasoned choice or rational understanding) was employed.

The *Godinez* holding has been subsequently criticized by legal scholars (e.g., Perlin, 1996) and courts alike. In the words of the Third Circuit Court of Appeals, “This difficult case presents us with a window through which to view the real-world effects of the Supreme Court’s decision in *Godinez v. Moran*, and it is not a pretty sight” (*Government of the Virgin Islands v. Charles*, 1995, p. 245). The problem is not whether the standards for various psycholegal competencies are higher, different, or the same, but rather, more fundamentally, whether the defendant has been examined with respect to these issues in the first place.

A single standard for competency to stand trial was a basis for finding Colin Ferguson, a man accused of murdering six people and injuring 19 others on the Long Island Rail Road in 1993, competent to waive counsel and represent himself. His lawyers intended to raise the insanity defense, but Ferguson objected and requested that he represent himself. Since he had been found competent to stand trial, the trial court allowed him to represent himself, or proceed *pro se*. The bizarre

nature of the trial, in which Ferguson cross-examined the police who arrested him as well as some victims, led Perlin (1996) to comment that the proceedings were a charade, and he concluded that the *Godinez* decision resulted in a trial in which dignity, a prerequisite of a fair trial, was not preserved.

Subsequently, the United States Supreme Court, in *Indiana v. Edwards* (2008), addressed the issue of whether there should be a different standard for competence to stand trial and competence to represent oneself. Ahmad Edwards, who was diagnosed with schizophrenia, attempted to steal a pair of shoes from a department store in Indiana, and, when security officers tried to apprehend him, he opened fire. He was subsequently charged with attempted murder, battery with a deadly weapon, criminal recklessness, and theft. In 2000, he was found incompetent to stand trial and remanded for treatment. He was restored to competency after 7 months; however, his attorneys requested a second evaluation in 2002, whereby he was found competent. In 2003, his attorneys requested a third evaluation, which resulted in a finding of incompetence. His competence was restored once again and he proceeded with adjudication.

In 2005, when his trial was set to begin, Edwards asked to proceed *pro se*. The trial judge denied the request. Edwards proceeded with appointed counsel and ultimately was convicted of the attempted murder and battery charges. The case carried through the appeals process and eventually was heard by the United States Supreme Court, which held that the Constitution does not prevent states from requiring counsel for those found competent to stand trial but who are not deemed competent to proceed *pro se*. The Court addressed its earlier decision in *Godinez* in the ruling but indicated that it did not provide direction in this case because the defendant's ability to conduct a defense at trial was not an issue at hand in *Godinez* and because the holding allows a state to permit a questionably competent defendant to proceed *pro se*, but it did not indicate whether a state may deny such a defendant.

Standards of competence are one area of inquiry, whereas the conceptualization of competence is another. Some researchers and scholars have provided reconceptualizations of competence to stand trial. Winick (1995) cogently argued that, in some circumstances, it might be in the best interests of the defendant to proceed with a trial even if he or she is incompetent. He postulated that this could take the form of a provisional trial wherein the support of the defense attorney would serve to ensure protection of the defendant. This would allow the defendant to proceed with his or her case while maintaining decorum in the courtroom and respecting the defendant's constitutional rights. Bonnie (1992a) also set forth a reformulation of competence to stand trial. He proposed a distinction between two types of competencies: competence to assist counsel and decisional competence. He argued that defendants found incompetent to assist counsel should be barred from proceeding until they were restored to competence, whereas defendants found decisionally incompetent could proceed in certain cases in which their lawyers were able to present a defense.

Another major change has been the shift in the type of settings in which competency assessments are conducted. Roesch and Golding (1980) argued that inpatient evaluations are unnecessary in all but perhaps a small percentage of cases, because most determinations of competency can easily be made on the basis of brief screening interviews (as discussed later in this chapter in the section on the Fitness Interview Test-Revised). Community-based settings, including jails and mental health centers, are increasingly used to conduct evaluations. In 1994, Grisso, Cocozza, Steadman, Fisher, and Greer published the results of a national survey they had conducted to determine the organization of pretrial forensic evaluation services in the United States. They concluded that "the traditional use of centrally located, inpatient facilities for obtaining pretrial evaluations survives in only a minority of states, having been replaced by other models that employ various types of outpatient approaches" (p. 388).

One compelling reason for this shift is cost. In 1977, Laben, Kashgarian, Nessa, and Spencer estimated that the cost of the community-based evaluations they conducted in Tennessee was one-third of the cost of institution-based evaluations (see also Fitzgerald, Peszke, & Goodwin, 1978). In 1985, Winick estimated that \$185 million was spent annually on competency evaluations; however, more recent estimates are 2 or 3 times higher and probably reach \$1 billion (annually) if the costs associated with the entire competency evaluation and treatment process are considered. For instance, in 2000, Bonnie and Grisso estimated that approximately 60,000 competency evaluations were performed each year. Blending community- and institution-based evaluations, a typical evaluation cost is \$5,000 per defendant; thus, approximately \$300 million may be spent annually on competency evaluations. If 20% of examined defendants are adjudicated incompetent, there are 12,000 restoration commitments yearly. In a typical jurisdiction, the average cost of a year in a forensic institution is \$145,000; thus, a conservative estimate of restoration costs per defendant would be \$36,250 for a 3-month period, bringing the annual restoration total to approximately \$435 million. Therefore, it appears that a conservative estimate of evaluation and restoration costs in the United States is upward of \$700 million annually.

Widespread use of screening instruments would serve to lower these costs, because the majority of individuals for whom competence is clearly not an issue would be screened out. Only those defendants whom the screening instrument has identified as potentially incompetent would then be referred for a more in-depth assessment. Screening instruments can be administered in community-based settings as well as in local jails or courthouses, thereby also serving to eliminate the unnecessary detention of clearly competent individuals.

Base rates for competency referrals (from 2% to 8% of felony arrests) and for incompetency determinations (from 7% to 60% of those defendants referred for competency evaluations) vary widely across jurisdictions and evaluation settings (Pirelli, Gottdiener, & Zapf, 2011; Skeem et al., 1998). This occurs for a number of reasons, including

- variations in examiner training and the use of forensically relevant evaluation procedures (Skeem et al., 1998);
- differences in the application of standards by examiners (Murrie, Boccaccini, Zapf, Warren, & Henderson, 2008);
- the availability of pretrial mental health services;
- the nature of the referral system;
- inadequate treatment services for the chronically mentally ill and criminalization of their conduct; and
- the extent to which judges scrutinize bona fide doubt about a defendant's competency before granting evaluation petitions (Golding, 1992).

Nevertheless, a meta-analysis of 68 studies reported a base rate of incompetence of 27.5% (Pirelli, Gottdiener, & Zapf, 2011). Precise data are not available, but conservatively, half of those found competent presented little or no reason for doubting their competency and could have been detected by adequate screening procedures. This is true in the United States as well as in other countries. Zapf and Roesch (1998) investigated the rate of incompetence in individuals remanded to an inpatient setting for an assessment of fitness to stand trial in Canada. Their results indicated that only 11% of the remands were unfit to stand trial and, further, that with the use of a brief screening interview, 82% of the remands could have been screened out at some earlier time because they were clearly fit to stand trial (Zapf & Roesch, 1997). Many of the assessment procedures we describe later in this chapter are either explicitly designed for screening or could easily be adapted for use in such settings.

A number of instruments designed to assess competence have been developed over the past 50 years. This work was pioneered by McGarry and his colleagues (see Lipsitt, Lelos, & McGarry, 1971; McGarry, 1965) and served as the starting point for a more sophisticated and systematic approach to the assessment of competency. In 1986, Grisso coined the term *forensic assessment instrument* (FAI) to describe instruments that provide the framework for conducting forensic assessments. FAIs are measures designed to evaluate psycholegal abilities rather than the psychological constructs measured by traditional psychological tests. FAIs make forensic assessments more systematic and assist evaluators in collecting important and relevant information by following the decision-making process required under the law. Since the term was coined, a number of assessment instruments have been developed that are designed to work in this way, and it appears that the use of FAIs has been increasing (Borum & Grisso, 1995; Ryba, Cooper, & Zapf, 2003b). This trend is encouraging in that empirical data suggest that trained examiners using FAIs achieve the highest levels of inter-examiner and examiner-adjudication agreement (Golding, Roesch, & Schreiber, 1984; Nicholson & Kugler, 1991; Skeem et al., 1998). Before turning to a review of specific assessment methods, we provide a brief overview of the legal procedures involved in competency-based questions.

OVERVIEW OF PROCEDURES

Laws regarding competency vary from state to state, although most jurisdictions follow procedures similar to the overview we describe in this section. Clinicians should consult the statute for the specific law and procedures applicable in their state.

The issue of competency may be raised at any point in the adjudication process (Golding & Roesch, 1988). If a court determines that a bona fide doubt exists as to a defendant's competency, it must consider this issue formally (*Drope v. Missouri*, 1975; *Pate v. Robinson*, 1966), usually via a forensic evaluation, which can take place, as noted, while the defendant is in the community, in jail, or in another institutional setting.

One legal issue that may concern evaluators is whether information obtained in a competency evaluation can be used against a defendant during the guilt phase of a trial or at sentencing. Although concerns regarding self-incrimination have been raised (Berry, 1973; Pizzi, 1977), all jurisdictions in the United States and Canada provide, either statutorily or through case law, that information obtained in a competency evaluation cannot be used in the context of a defendant's guilt unless the defendant places his or her mental state into evidence at either the trial or the sentencing hearings (*Estelle v. Smith*, 1981; Golding & Roesch, 1988).

Once a competency evaluation has been completed and the written report has been submitted (see Heilbrun, Marczyk, & DeMatteo, 2002; Melton et al., 2007; Skeem et al., 1998; Zapf & Roesch, 2009, for a discussion of the content of these reports), the court may schedule a hearing. If both defense and prosecution accept the findings and recommendations in the report, however, a hearing does not have to take place. As such, it is likely that, in the majority of cases, a formal hearing is not held. If a hearing is held, the evaluators may be asked to testify, but most hearings are quite brief, and usually only the written report of an evaluator is used. The ultimate decision about competency rests with the court, which is not bound by the evaluators' recommendations (e.g., *North Dakota v. Heger*, 1982). In most cases, however, the court accepts the recommendations of the evaluators (Cochrane, Herbel, Reardon, & Lloyd, 2012; Gowensmith, Murrie, & Boccaccini, 2012; Hart & Hare, 1992; Steadman, 1979; Zapf, Hubbard, Cooper, Wheelles, & Ronan, 2004).

At this point, defendants adjudicated competent proceed with their cases. For defendants found incompetent, their trials are postponed until their competency is regained or the charges are dismissed, usually without prejudice. The disposition of incompetent defendants is perhaps the most problematic aspect of this area of law. Until the case of *Jackson v. Indiana* (1972), virtually all states allowed for the automatic and indefinite commitment of incompetent defendants. In *Jackson*, however, the United States Supreme Court held that defendants committed solely on the basis of incompetency "cannot be held more than the reasonable period of time necessary to determine whether there is a substantial probability that he will attain that capacity in the foreseeable future" (p. 738). The Court did not specify how long a period of time would be reasonable, nor did it indicate how progress toward the goal of regaining competency could be assessed.

The *Jackson* decision led to revisions in state statutes that provided alternatives to commitment and limited the length of commitment (Roesch & Golding, 1980). The duration of confinement varies from state to state; some states have specific time limits (e.g., 18 months) while others are based on a proportion of the length of sentence that would have been given if the defendant were convicted.

Medication is the most common form of treatment provided to restore defendants' competence, although some jurisdictions have established psychoeducational programs designed to increase defendants' understanding of the legal process (e.g., Anderson & Hewitt, 2002; Bertman et al., 2003; Pendleton, 1980; Wall, Krupp, & Guilmette, 2003; Webster, Jenson, Stermac, Gardner, & Slomen, 1985) or to confront problems that hinder a defendant's ability to participate in the defense (D. L. Davis, 1985; Siegel & Elwork, 1990). What happens if an incompetent defendant refuses treatment, particularly medications? The United States Supreme Court case provided at least a partial answer to this question. In *Sell v. United States* (2003), the Supreme Court considered the case of Charles Sell, a dentist who was charged with multiple counts of insurance fraud and who was deemed incompetent and committed for treatment at a federal medical center. Sell refused medication, and the treatment staff subsequently sought to have him involuntarily medicated. The medical center's review panels concluded that Sell was mentally ill and dangerous, that medication would be helpful in reducing his symptoms, and that medication would also help restore his competence to stand trial.

Sell appealed, and the federal magistrate who had ordered Sell's commitment supported the government's position that medication was the best treatment alternative to address the issue of his dangerousness and also to restore his competency. The Supreme Court held that antipsychotic drugs could be administered against a defendant's will for the purpose of restoring competency, but only in limited circumstances. The Court noted that this applied only to the issue of competency restoration, and indicated that medication could be justified on other grounds, including dangerousness (see *Riggins v. Nevada*, 1992; *Washington v. Harper*, 1990). Writing for the majority, Justice Breyer identified several factors that must be satisfied before a defendant can be involuntarily medicated. These include a determination that the medication is likely to restore competence but will not result in side effects that might affect a defendant's ability to assist counsel. The court must also find that alternative and less intrusive methods that would achieve the same result are not available.

The *Sell* decision may have limited the use of medication as an option for some incompetent defendants who refuse voluntary treatment. Indeed, Justice Breyer commented that he thought that instances of forced medication might be rare. As a consequence, the *Sell* case may result in a greater emphasis being placed on the development of alternative forms of treatment for restoring competence. Nevertheless, Cochrane and his colleagues (2012) conducted a retrospective records review of all incompetent defendants who were involuntarily treated under *Sell* in the United States federal court system ($n = 132$) between 2003 and 2009. They found

that 79% of those diagnosed with a psychotic disorder responded well to treatment and ultimately were deemed competent to stand trial.

This brief overview of competency procedures is intended to provide a basic understanding of the process. For a more comprehensive discussion of the legal issues involved and a review of empirical research on the various aspects of the competency procedures, the reader is referred to reviews by Pirelli, Gottdiener, and Zapf (2011) and by Zapf and Roesch (2009).

ASSESSING COMPETENCY

Although there has been some confusion over the definition of the term *competency*, there is generally a high level of agreement among evaluators regarding a defendant's competence. Researchers who have investigated interrater reliability generally report that pairs of evaluators agree in up to 80% or more of the cases (Goldstein & Stone, 1977; Poythress & Stock, 1980; Roesch & Golding, 1980; Skeem et al., 1998). Gowensmith, Murrie, and Boccaccini (2012) published findings from a study conducted with a sample of 216 Hawaiian cases and found slightly lower agreement rates; they found 71% agreement among examiners in initial competency evaluations and 61% agreement in cases pertaining to competency restoration. Murrie and colleagues (2008) found great variability in levels of agreement, which they attributed to evaluator, system-, and policy-level factors. When evaluators are highly trained and use semistructured competence assessment instruments, however, high rates of agreement have been reported (Golding et al., 1984; Nicholson & Kugler, 1991; Robinson & Acklin, 2010).

When base rates of findings of competency are considered, these high levels of agreement are less impressive, and they do not suggest that evaluators are necessarily in agreement about the specific criteria involved in a determination of competency. In their recently published meta-analysis, Pirelli, Gottdiener, and Zapf (2011) found the base rate of incompetency to be 27.5% in a total sample of over 20,000 pretrial defendants referred for such evaluations. Thus, examiners, without even directly assessing a group of defendants, could achieve high levels of agreement with an examining clinician simply by deeming all defendants competent (i.e., a base-rate decision). Because in most jurisdictions more than 70% of all referred defendants are adjudicated competent, the psychologist and the examiner would achieve at least modest agreement, even without making any further decisions. (Of note is that studies reporting interrater reliability statistics usually have small samples overall and consequently very few incompetent defendants.) Skeem and her colleagues (1998) demonstrated that examiner agreement on specific psycholegal deficits (as opposed to dichotomous decisions of competency) averaged only 25% across a series of competency domains. It is the more difficult decisions, involving cases where competency is truly a serious question, that are of greatest concern. How reliable are decisions about these cases?

One study used a vignette methodology to determine whether examiners would distinguish between three different standards for competency (a “rational understanding” standard, a “rational manner” standard, and a standard where the word *rational* was not used). Results of this study indicated that examiners were almost equally divided in their opinions regarding the defendant’s competence when applying *Dusky’s* rational understanding standard: 47.6% found the defendant competent and 52.4% found the defendant incompetent (Morris, Haroun, & Naimark, 2004). Similar results were found for the rational manner standard and the standard that did not use the word *rational*. These authors felt that the nearly equally divided response to the same vignette was shocking and concluded that “the defendant’s fate depends only on who performs the evaluation” (p. 237).

High levels of reliability do not, of course, ensure that valid decisions are being made. Two evaluators could agree that the presence of psychosis automatically leads to a finding of incompetency. As long as the evaluators are in agreement about their criteria for determining psychosis, the reliability of their final judgments about competency would be high. As we suggest throughout this chapter, it is quite possible that too many evaluators inappropriately rely on traditional mental status issues without considering the functional aspects of a particular defendant’s case.

The validity of competency judgments is difficult to assess because of the criterion problem. Criterion-related validity typically is assessed by examining concurrent validity and predictive validity (Messick, 1980). Predictive validity is impossible to assess fully in the competency arena, because only defendants who are considered competent are allowed to proceed. It is feasible to look at the predictive validity of decisions about competent defendants, but it is not possible to assess the decisions about incompetent defendants, because they are remanded for treatment and their judicial proceedings are suspended. Concurrent validity is also difficult to investigate, because it is not useful to examine correlations with other measures (e.g., diagnosis, intelligence) if one adopts a functional, case-by-case assessment of a defendant’s competency. For these reasons, there is no “correct” decision against which to compare judgments.

As we have indicated, the courts usually accept mental health professionals’ judgments about competency. Does this mean that the judgments are valid? Not necessarily, because courts often accept the evaluator’s definition of competency and his or her conclusions without review, leading to very high levels of examiner–judge agreement (Cochrane et al., 2012; Gowensmith et al., 2012; Hart & Hare, 1992; Skeem et al., 1998; Zapf et al., 2004)—an issue that has been acknowledged for quite some time (Vann & Morganroth, 1965). Roesch and Golding (1980) argued that the only ultimate way of assessing the validity of decisions about incompetency is to allow defendants who are believed to be incompetent to proceed with a trial anyway. This could be a provisional trial (similar to the Illinois model) in which assessment of a defendant’s competence-related abilities could continue. If a defendant were unable to participate, then the trial could be stopped. If a verdict had already been reached and the defendant was convicted, the verdict could be set aside.

We suspect that, in a significant percentage of trials, alleged incompetent defendants would be able to participate. In addition to the obvious advantages to defendants, the use of a provisional trial could provide valuable information about what should be expected of a defendant in certain judicial proceedings (e.g., the ability to testify, identify witnesses, describe events, evaluate the testimony of other witnesses, etc.). Short of a provisional trial, it may be possible to address the validity issue by having independent experts evaluate the information provided by evaluators and other collateral information sources.

FUNCTIONAL EVALUATION APPROACH

We believe the most reasonable approach to the assessment of competency is based on a functional evaluation of a defendant's ability matched to the contextualized demands of the case. Although an assessment of the mental status of a defendant is important, it is not sufficient as a method of evaluating competency. Rather, the mental status information must be related to the specific demands of the legal case, as has been suggested by legal decisions, such as the ones involving amnesia. As in the case of psychosis, a defendant with amnesia is not per se incompetent to stand trial, as has been held in a number of cases (e.g., *Ritchie v. Indiana*, 1984; *Wilson v. United States*, 1968). In *Missouri v. Davis* (1983), the defendant had memory problems due to brain damage. Nevertheless, the Missouri Supreme Court held that amnesia alone was not a sufficient reason to bar the trial of an otherwise competent defendant. In *Montana v. Austad* (1982), the court held that the bulk of the evidence against the defendant was physical and not affected by amnesia. Finally, in a Maryland decision (*Morrow v. Maryland*, 1982), the court held that, because of the potential for fraud, amnesia does not justify a finding of incompetence. The court also stated that everyone has amnesia to some degree, because the passage of time erodes memory. These decisions are of interest because they support the view that evaluators cannot reach a finding of incompetency independent of the facts of the legal case. Similarly, a defendant may be psychotic and still be found competent to stand trial if the symptoms do not impair the defendant's functional ability to consult with his or her attorney and otherwise rationally participate in the legal process.

Some cases are more complex than others and may, as a result, require different types of psycholegal abilities. Thus, it may be that the same defendant is competent for one type of legal proceeding but not for another. In certain cases, a defendant may be required to testify. In this instance, a defendant who is likely to withdraw in a catatonic-like state may be incompetent to proceed; however, the same defendant may be able to proceed if the attorney intends to enter a plea bargain (the way in which the vast majority of all criminal cases are handled).

The functional approach is illustrated in the famous amnesia case of *Wilson v. United States* (1968). In that decision, the court of appeals held that six factors should

be considered in determining whether a defendant's amnesia impaired the ability to stand trial:

1. The extent to which the amnesia affected the defendant's ability to consult with and assist his lawyer.
2. The extent to which the amnesia affected the defendant's ability to testify in his own behalf.
3. The extent to which the evidence in suit could be extrinsically reconstructed in view of the defendant's amnesia. Such evidence would include evidence relating to the crime itself as well as any reasonable possible alibi.
4. The extent to which the Government assisted the defendant and his counsel in that reconstruction.
5. The strength of the prosecution's case. Most important here will be whether the Government's case is such as to negate all reasonable hypotheses of innocence. If there is any substantial possibility that the accused could, but for his amnesia, establish an alibi or other defense, it should be presumed that he would have been able to do so.
6. Any other facts and circumstances which would indicate whether or not the defendant had a fair trial. (pp. 463–464)

One could substitute any symptom for amnesia in the quote from *Wilson*. If this were done, the evaluation of competency certainly would be one based on a determination of the manner in which a defendant's incapacity may affect the legal proceedings. In fact, some state codes, such as Florida (Florida Rules of Criminal Procedure 3.21(a)(1), see Winick, 1983) and Utah Code (1994), already specify that the evaluators must relate a defendant's mental condition to clearly defined legal factors, such as the defendant's appreciation of the charges, the range and nature of possible penalties, and capacity to disclose to an attorney the pertinent facts surrounding the alleged offense (Winick, 1983). Utah's (1994) statute goes the farthest in this regard, specifying the most comprehensive range of psycholegal abilities to be addressed by evaluators (including the iatrogenic effects of medication and decisional competencies) and also requiring judges to specify which psycholegal abilities are impaired when a defendant is found incompetent.

The assessment of competency requires consideration of both mental status and psycholegal abilities. Unfortunately, research findings have demonstrated that evaluators often do not address an appropriate range of psycholegal abilities and most often do not draw a connection between the defendant's emotional or behavioral impairments and impaired psycholegal capacities (Skeem et al., 1998). We now turn to a review of the history of competency assessment methods.

COMPETENCY ASSESSMENT TOOLS

Prior to the 1960s, there were no standard methods for assessing competency. One of the first was a checklist developed by Robey (1965) that focused on court process-related issues, such as understanding of the legal process. Another early procedure,

developed by Bukatman, Foy, and de Grazia (1971), used a checklist and a set of interview questions. Neither of these early measures was used often (Schreiber, 1978). Undoubtedly, A. Louis McGarry and his colleagues at the Harvard Medical School's Laboratory of Community Psychiatry had the greatest early influence on the formal assessment of competency. McGarry, a psychiatrist, was involved in the development of two measures: the Competency Screening Test and the Competency Assessment Instrument. We next discuss these measures in addition to a number of other measures that have since been developed (see also Ackerman, 1999; Acklin, 2012; Melton et al., 2007; and Zapf & Viljoen, 2003, for a review of competency assessment instruments).

Competency Screening Test. The Competency Screening Test (CST) was created by Lipsitt et al. (1971) as a screening measure to identify clearly competent defendants and minimize the need for lengthy institutional evaluations. Such a screening process was considered important because the vast majority of defendants referred for evaluations are competent. The high rate of competency is that many other factors influence referrals, including the use of the evaluation commitment as a method for denying bail, as a tactical maneuver to delay a trial, as a way of providing a basis for a reduction in charges or sentences, and as a means of getting defendants who are seen as in need of mental health treatment out of jails and into hospitals (Dickey, 1980; Golding, 1992; Menzies, Webster, Butler, & Turner, 1980; Roesch & Golding 1985; Teplin, 1984). Mumley, Tillbrook, and Grisso (2003) referred to this as the "subversion" hypothesis; that is, using competency evaluations in a back-door manner to provide defendants with treatment.

The CST, however, has seldom been used as a screening device because of various validity-related concerns. Specifically, the tool focuses on defendants' knowledge of legal issues to the exclusion of other factors, and its scoring method has been criticized (Brakel, 1974; Roesch & Golding, 1980) because of its idealized perception of the criminal justice system; certain responses actually may reflect a sense of powerlessness in controlling one's outcome in the legal system and may be based on past experiences with that system.

The CST has been examined in a number of studies. Although it has high levels of interrater reliability in terms of scoring the incomplete sentence format (Randolph, Hicks, & Mason, 1981), studies comparing classifications based on CST cutoff scores and hospital evaluation decisions reveal that it has a high false-positive rate (i.e., it tends to identify many individuals as incompetent who are later determined to be competent in institutional evaluations; Lipsitt et al., 1971; Nottingham & Mattson, 1981; Randolph et al., 1981; Shatin, 1979).

The results of these studies lead to a mixed review of the CST. Although it appears that the CST is reliably scored, serious questions can be raised about its usefulness as a screening device, because of the potential for misclassifying potentially competent defendants. At this point, it is not recommended for use as the sole method of screening defendants.

Competency Assessment Instrument. The most important measure developed by McGarry, the Competency Assessment Instrument (CAI), comprises 13 items related to legal issues. It has served as the basis for other subsequently developed forensic assessment instruments. The items include “appraisal of available legal defenses,” “quality of relating to attorney,” and “capacity to disclose pertinent facts.” Each item is scored on a 5-point scale, ranging from “total incapacity” to “no incapacity.” The CAI manual includes clinical examples of levels of incapacity as well as suggested interview questions.

The CAI has been used in a number of jurisdictions, although perhaps more as an interview structuring device rather than the two-stage screening manner (with the CST) as originally intended by McGarry (see Laben et al., 1977; Schreiber, 1978). Unfortunately, few studies report either reliability or validity data. Roesch and Golding (1980) examined the utility of the CAI by comparing 30 interviews conducted by pairs of interviewers who both used the tool; item percentage agreements ranged from 68.8% to 96.7%, with a median of 81.2%. The interviewers were in agreement on the competency status of 29 of the 30 defendants (26 competent, 3 incompetent). The interviewers’ decisions were in concordance with the more lengthy hospital evaluation decisions in 27 of 30 cases, or 90%. In subsequent studies (Golding et al., 1984; others summarized in Nicholson & Kugler, 1991), the CAI has shown high levels of *trained* inter examiner agreement and examiner-outcome agreement. As such, the CAI may be useful as a screening device or as a full-blown interview measure. Its primary disadvantage, relative to the instruments reviewed next, is in the range of psycholegal abilities articulated and its lack of focus on the nexus between psychopathology and psycholegal impairment.

Interdisciplinary Fitness Interview. The Interdisciplinary Fitness Interview (IFI) is designed to assess both the legal and psychopathological aspects of competency (Golding et al., 1984). The original IFI comprised three major sections: (1) legal issues (5 items), (2) psychopathological issues (11 items), and (3) overall evaluation (4 items). The three items in the consensual judgment section reflect post-assessment resolution of differences between judges.

The IFI has been revised (Golding, 1993) to reflect changes in constitutional law and the adoption by many states of “articulated” competency standards (e.g., Utah Code, 1994). In its current form, the Interdisciplinary Fitness Interview–Revised (IFI-R) taps into 31 relatively specific psycholegal abilities organized across 11 global domains. The IFI-R was developed on the original model used by Golding et al. (1984) but was altered to reflect a decade of experience, numerous court opinions, and the accumulated professional literature on competency assessments. For example, it specifically addresses the issue of the iatrogenic effects of psychotropic medications (*Riggins v. Nevada*, 1992), a defendant’s decisional competency to engage in rational choice about trial strategies, proceeding *pro se*, or pleading guilty (see earlier discussion of *Godinez v. Moran*, 1993) and competency to confess. It was developed to mirror Utah’s (1994) articulated competency code, which mandates

that examiners address its 11 global domains. A revised and comprehensive training manual is available (Golding, 1993).

Golding et al. (1984) used the IFI in a study of pretrial defendants in the Boston area who were referred by court clinics to a state mental hospital for competency evaluation. Teams composed of a lawyer and either a psychologist or a social worker interviewed defendants. Although the interviews were conducted jointly, each evaluator independently completed the IFI rating form. The results demonstrated that lawyers and mental health evaluators could make reliable competency judgments together. They were in agreement on 97% of their final determinations of competency. By type of decision, the interviewers found 58 defendants to be competent and 17 incompetent, and they disagreed on only 2 cases.

Fitness Interview Test. The Fitness Interview Test (FIT; Roesch, Webster, & Eaves, 1984) was created in 1984 to structure fitness to stand trial assessments conducted in Canada. It has since been extensively revised, and the current version is referred to as the Fitness Interview Test (revised edition) (FIT-R; Roesch, Zapf, & Eaves, 2006). The FIT-R focuses on the psycholegal abilities of the individual. The scoring system has been changed to a 3-point scale, with a score of 2 meaning definite or serious impairment, 1 meaning possible or mild impairment, and 0 meaning no impairment. The items on the FIT-R were developed to parallel the standards for fitness that were established in section 2 of the 1992 revision of the Criminal Code of Canada, which are quite similar to the standards articulated by the United States Supreme Court in *Dusky v. United States* (1960).

The FIT-R takes approximately 30 minutes to administer and consists of a structured interview that addresses three main areas: (1) ability to understand the nature or object of the proceedings, or factual knowledge of criminal procedure; (2) ability to understand the possible consequences of the proceedings, or the appreciation of personal involvement in and importance of the proceedings; and (3) ability to communicate with counsel or to participate in the defense. Each of these three sections is broken down into specific questions that tap into different areas involved in fitness to stand trial. The first section assesses the defendant's understanding of the arrest process; the nature and severity of current charges; and the role of key players, legal processes, pleas, and court procedure. The second section assesses the defendant's appreciation of the range and nature of possible penalties, appraisal of available legal defenses, and appraisal of likely outcomes. The final section assesses the defendant's capacity to communicate facts to the lawyer, relate to the lawyer, plan legal strategy, engage in his or her own defense, challenge prosecution witnesses, testify relevantly, and manage courtroom behavior.

The FIT-R has demonstrated excellent utility as a screening instrument in research studies. Zapf and Roesch (1997) compared the results of the FIT-R and an institution-based fitness assessment for 57 defendants remanded to an inpatient psychiatric institution for an evaluation of fitness. When specific decision rules were used to classify defendants as either fit or unfit, the FIT-R correctly predicted fitness status

for 49 of the 57 individuals. The remaining 8 individuals were judged to be unfit by the FIT-R and fit as a result of the inpatient assessment. These false-positive errors were expected, as screening instruments should overestimate the rate of unfitness without making any false-negative errors. There was 100% agreement between the FIT-R and the institution-based assessment for those individuals deemed fit to stand trial.

Research also has indicated that the FIT-R has adequate psychometric properties. Viljoen, Roesch, and Zapf (2002) found that the average interrater reliability of the FIT-R for overall determination of fitness was .98. Reliability for the sections was lower and ranged from .54 to .70 for groups of raters. To investigate the predictive validity of the FIT-R, Zapf et al. (Zapf & Roesch, 1997; Zapf, Roesch, & Viljoen, 2001) compared decisions made by the FIT-R to decisions made in an institution-based evaluation of fitness in several samples. In the first sample, the overall rate of agreement between the FIT-R and institution-based judgments was 87%, and the false-negative error rate was 2%. The second sample yielded comparable results. Boddy, Roesch, Zapf, and Eaves (2000) compared defendants who were referred for fitness evaluations, including those who were eventually found unfit, and defendants who were not referred. Performance on the FIT-R effectively distinguished these groups, such that referred defendants demonstrated significantly more impairment. As additional evidence of the construct validity of the FIT-R, Zapf and Roesch (2001) found reasonably high agreement (chance-corrected $\kappa = .51$) between judgments made using the FIT-R and defendant's performance on the MacArthur Competency Assessment Tool—Criminal Adjudication (MacCAT-CA; Poythress et al., 1999). The FIT-R has also been used in research with young offenders (Roesch, 2011; Viljoen, Vincent, & Roesch, 2006; Viljoen, Penner, & Roesch, 2012).

Georgia Court Competency Test. The Georgia Court Competency Test (GCCT) was originally developed by Wildman et al. (1978) and has since gone through a number of revisions (see Bagby, Nicholson, Rogers, & Nussbaum, 1992; Johnson & Mullet, 1987; Nicholson, Briggs, & Robertson, 1988; Wildman, White, & Brandenburg, 1990). The original version consisted of 17 items, and the revised version, referred to as the Mississippi State Hospital Revision (GCCT-MSH), consists of 21 items. The first seven items of the GCCT-MSH require the defendant to visually identify the location of certain participants in the courtroom using a courtroom drawing as a reference. This is followed by questions related to the function of certain individuals in the courtroom, the charges that the defendant is facing, and his or her relationship with the lawyer.

Research on the GCCT-MSH has indicated that this instrument displays strong reliability and validity evidence (Nicholson, Robertson, Johnson, & Jensen, 1988). Three factors identified by Nicholson et al. (1988)—Courtroom Layout, General Legal Knowledge, and Specific Legal Knowledge—were later replicated by Bagby et al. (1992). It was later suggested that this three-factor solution may be appropriate

only for defendants who have been ordered to undergo assessment at the pretrial stage (Ustad, Rogers, Sewell, & Guarnaccia, 1996). These researchers indicated that a two-factor solution (Legal Knowledge and Courtroom Layout) may be more appropriate for defendants who have been adjudicated incompetent and who are undergoing inpatient treatment to restore competence. The major drawback of the GCCT-MSH is that it focuses on foundational competencies and generally ignores the more important decisional competencies stressed in the IFI and FIT approaches (Bonnie, 1992a).

MacArthur Competence Assessment Tool—Criminal Adjudication. The MacArthur Competence Assessment Tool—Criminal Adjudication (MacCAT-CA; Hoge, Bonnie, Poythress, & Monahan, 1999; Poythress et al., 1999) was developed to assess three main subconstructs of the psycholegal abilities: understanding, reasoning, and appreciation. It comprises 22 items that are grouped into three sections. The examiner begins by reading a hypothetical vignette to the defendant to ground the first two sections (16 items). The first section (8 items) assesses the defendant's ability to *understand* information about the legal system and the process. For each item, the defendant is asked a question related to the vignette (e.g., "What is the job of the attorney for the defense?") and is awarded 2 points (items are rated 0, 1, 2) if he or she is able to answer the question in a manner that demonstrates full understanding. If the defendant earns fewer than 2 points, the examiner discloses the answer and asks the defendant to repeat the disclosure in his or her own words. The purpose of the disclosure is to independently assess the defendant's *capacity* to understand and his or her *actual* or preexisting understanding.

The second section (eight items) assesses the defendant's ability to *reason*. The first five items in this section assess the defendant's ability to consider two pieces of factual information and identify the most important or legally relevant piece of information that should be disclosed to a lawyer. The last three items require the defendant to think through mock legal options (relevant to the vignette) and to evaluate them in various ways.

The final section (six items) assesses the defendant's ability to appreciate his or her own legal circumstances and situation. This section departs from the hypothetical vignette format to explore the defendant's beliefs and perceptions about his or her personal role as a defendant and how he or she will be treated during the course of adjudication. These items are scored on the basis of the reasons that the defendant provides for his or her judgment and whether they are plausible or implausible (i.e., grounded in reality or based on delusional beliefs).

It is important to note that the authors of the MacCAT-CA emphasize that this instrument was developed for use as a tool rather than a test of competence and that it constitutes only one component of an overall assessment of competence. Thus, the scores obtained must be interpreted within the context of the specific defendant's case and integrated with all the other clinically relevant factors that may surround the specific circumstances of the case.

The psychometric properties of the MacCAT-CA were examined based on a sample of 729 felony defendants in eight different states (Otto et al., 1998; see also Rogers, Grandjean, Tillbrook, Vitacco, & Sewell, 2001). The results indicated that the MacCAT-CA demonstrated high levels of reliability. For each of the three sections, internal consistency ranged from .81 to .88 ($\alpha = .81$ for Reasoning, .85 for Understanding, .88 for Appreciation), and interrater reliability ranged from very good to excellent (intraclass $R = .75$ for Appreciation, .85 for Reasoning, .90 for Understanding).

Otto and colleagues (1998) reported that additional support for the construct validity of the MacCAT-CA was “found in the pattern of correlations between the MacCAT-CA measures and select clinical variables” (p. 439). MacCAT-CA Understanding, Reasoning, and Appreciation scores correlated (.41, .34, and .14, respectively) with estimated Wechsler Adult Intelligence Scale—Revised (WAIS-R) full scale IQ scores (–.23, –.29, and –.36, respectively), and with Brief Psychiatric Rating Scale (BPRS) Total Scores (these scores correlated more strongly with BPRS Psychoticism and Emotional Withdrawal than Depression and Hostility scales). The three MacCAT-CA scales correlated moderately with clinicians’ global ratings of competency (.36, .42, and .49, respectively).

Evaluation of Competency to Stand Trial—Revised. The Evaluation of Competency to Stand Trial—Revised (ECST-R; Rogers, Tillbrook, & Sewell, 2004) was developed to structure an examiner’s judgments about a defendant’s competence-related abilities and to provide normative data in this regard. The ECST-R represents the first formal competency assessment instrument created specifically to serve, in part, as a screener of feigned incompetency and consists of four scales. Three of these scales were designed to tap into the three *Dusky* prongs: Consult with Counsel (CC), Factual Understanding of the Courtroom Proceedings (FU), and Rational Understanding of the Courtroom Proceedings (RU). The fourth scale, Atypical Presentation (AP), is a response style measure consisting of five subscales: Realistic, Psychotic, Nonpsychotic, Impairment, and Both (Psychotic and Nonpsychotic combined).

The CC scale is composed of 10 questions with five criteria rated on a 5-point scale, wherein 0 = not observed; 1 = questionable clinical significance; 2 = mild impairment; 3 = moderate impairment, unrelated competency (will affect but not impair competency); and 4 = severe impairment, directly related to competency (will substantially impair competency). Some of these items have more than one component to rate (e.g., reasoning and hallucinations), and the highest score given is used toward the total scale score.

The FU scale consists of 15 questions with 13 criteria. With one exception, items are rated on a 5-point scale, wherein 0 = correct, 1 = correct when prompted, 2 = correct with attempts to educate, 3 = wrong despite attempts to educate, and 4 = grossly psychotic and totally unrelated to the question. The exception is item 12a, which is coded either 0 = understands risk of talking to prosecutor or 1 = does not understand risk.

The RU scale consists of 10 questions with six criteria rated on 5-point scales. Most items are rated 0 = not observed; 1 = questionable clinical significance; 2 = mild impairment, does not appear to care about the outcome of the case; 3 = moderate impairment, demonstrates some effort to undermine the case; or 4 = severe impairment, concerted effort to be found guilty or grossly maximize punishment. In the second part, of this scale, these items are rated 0 = not observed; 1 = questionable clinical significance; 2 = mild impairment, unrelated to competency; 3 = moderate impairment, peripherally related competence (will affect but not impair competency); or 4 = severe impairment, directly related to competency (will substantially impair competency). Item 18 relates to the defendant's self-defeating motivations and is rated 0 = none; 1 = slight, no sign of interruption; 2 = mild, likely to need admonitions from judge or defense counsel; 3 = major, likely to interrupt proceedings for nonpsychotic reasons; or 4 = major, likely to interrupt proceedings for psychotic reasons.

The AP scale consists of 28 items rated on a 3-point scale, wherein 0 = no, 1 = sometimes, and 2 = yes. For certain items, if a rating of 2 (yes) is given, impairment is rated either 0 = nonimpaired or 1 = impaired. Some items regard spending additional time with one's attorney, where a yes response is appropriate; other items (e.g., seeing men turning into women) are clearly inappropriate to respond yes to, but if answered in the affirmative, the defendant is asked if such problems would make it difficult to go to court. Scores greater than 6 on this scale indicate feigning.

No overall composite score is generated, but ECST-R scale scores are intended to be interpreted using both normative and case-specific methods. Scale scores are transformed into T scores, and four levels of impairment are moderate (60–69T), severe (70–79T), extreme (80–89T), and very extreme (90T and higher). There is also a measurement of certitude for each scale based on these T scores: preponderant (more likely than not > 50%), probable (84.1% likelihood), very probable (95% likelihood), and definite (98% likelihood). Cut-scores are not intended to detect feigning on the ECST-R; however, the manual provides a 4-step approach to investigate suspected malingering.

Rogers and colleagues (Rogers et al., 2001; Rogers, Jackson, Sewell, Tillbrook, & Martin, 2003) have found high interrater reliability estimates for the ECST-R at both scale (.97–1.00) and item levels (.89–.99), and Rogers et al. (2001) found perfect interrater reliability for ECST-R total scores. Measures of internal consistency ranged from alphas of .83 to .93 across both studies with one exception: In the first study (Rogers et al., 2001), a .72 alpha coefficient was found for the CC scale.

Inventory of Legal Knowledge. The Inventory of Legal Knowledge (ILK; Otto, Musick, & Sherrod, 2010) was developed to assist in the evaluation of a defendant's response style when undergoing evaluations of adjudicative competence. Despite its title, the ILK is not a measure of adjudicative competence; it is to be used solely as a means of evaluating whether a defendant is feigning ignorance of legal knowledge. The ILK

consists of 61 items in a true–false format and takes about 15 minutes to administer. Items ask about a variety of issues relevant to the legal process, including the roles of various courtroom personnel (i.e., judge, defense attorney, witnesses), courtroom procedures, charges, pleas, sanctions, and defendant rights. The ILK uses two strategies to detect feigned deficits in legal knowledge: comparison of obtained scores with scores expected by chance (obtained scores that are significantly lower than chance provides evidence of suppressed legal knowledge) and the floor effect wherein defendants are identified as feigning based on obtained scores that are significantly lower than those obtained by relevant reference samples. The ILK has shown good internal consistency ($\alpha = .88$), with mean item-total correlations of .32 (range = .10–.53); test-retest reliability of .76 with some evidence of a practice effect; and convergent validity with commonly used measures of response style (Otto et al., 2010).

COMPETENCY IN SPECIAL POPULATIONS

In recent years, attention has been paid to two specific populations for whom issues of competency are especially important for various reasons—juveniles and individuals diagnosed with intellectual disabilities or mental retardation (MR). This section examines some of the issues relevant to the assessment of competence for these special populations, including descriptions of instruments that have been developed to assess competence-related abilities within these populations.

Relatively few researchers and scholars have addressed the problems faced by individuals diagnosed with intellectual disabilities within the criminal justice system. Bonnie (1992b) noted that one of the biggest problems for these individuals is underidentification. That is, a considerable number of defendants diagnosed with MR are not referred for psychological evaluations, and a general failure to recognize the magnitude and/or existence of the disabilities of such individuals is a major cause of the low rate of referrals. Although the underidentification hypothesis is widely accepted among researchers and scholars, there are few empirical data to support it (Bonnie, 1992b).

Failure to identify individuals with intellectual disabilities and subsequently refer them for evaluations of competency is often a result of the tendency of these individuals to attempt to hide their limitations. That is, persons diagnosed with MR are often compliant and cooperative with authority figures, such as judges or lawyers, and are likely to pretend to understand their lawyers when, in fact, they may not (Bonnie, 1992b). Research by Everington, Notario-Smull, and Horton (2007) suggests that defendants diagnosed with MR may be able to hide their deficits in certain circumstances. Nevertheless, in many cases, a “cloak of competence” gives these individuals the appearance of normalcy in the competency context (Edgerton, 1993). Legally significant impairments then become visible only when the individual also has a severe mental illness or acts in a strange or disruptive manner (Bonnie, 1992b). It is common for these individuals to proceed to trial

without ever having been identified (Cooper & Grisso, 1997) and, consequently, to proceed through the criminal justice system without understanding “the process or punishment, often unknowingly participating in agreements that can result in grave and long-lasting consequences” (Everington, 1990, p.148). The misidentification of individuals diagnosed with intellectual disabilities can result in a loss of liberty and the violation of the right to a fair trial.

Once individuals diagnosed with MR have been identified, it is important that an examiner who is familiar with the specific issues relevant to the diagnosis conduct the evaluation. Petrella (1992) noted that the probability of an adequate evaluation of issues involving such a diagnosis is highly unlikely for several reasons, including the fact that (1) many evaluators are not qualified to administer intelligence tests, (2) evaluators may have minimal experience with MR and limited exposure to the unique clinical presentation of individuals with such a diagnosis, (3) referrals to professionals and experts in the area of intellectual disability per se seldom occur, and (4) professionals with specific expertise with such persons usually have minimal or no experience with forensic issues.

Research examining the rates of incompetence among defendants with MR has provided conflicting findings. In Missouri, 17% of defendants referred for competence evaluations and diagnosed with MR were adjudicated incompetent whereas in Connecticut and Michigan, rates of incompetence were reported to be 12.5% and 33%, respectively (Petrella, 1992). In addition, the probability of being found incompetent also varies by severity of the defendant’s intellectual limitations. In a Virginia study, 23% of defendants diagnosed with mild MR were deemed incompetent, whereas 68% of those with moderate MR were found incompetent (Petrella, 1992).

In terms of restoration to competency for individuals diagnosed with MR, Bonnie (1992b) hypothesized that the possibility of restoration is highly unlikely. In a sample of 38 incompetent defendants with such a diagnosis in Virginia, only 16% were considered likely to be restored, and it was estimated that restoration was unlikely for approximately 66% of the defendants (Petrella, 1992). In a 1994 study of 271 defendants diagnosed with MR and committed to a special competence restoration program, individuals with higher IQs were more likely to be found competent or restored to competence (Jones, 1994). In addition, the author concluded that the absence of comorbidity (i.e., presence of a psychiatric disorder in addition to the diagnosis of MR) was associated with restorability, as only 3% of those with comorbid disorders were restored to competence as compared to approximately 20% of those without comorbid disorders (Jones, 1994).

Although further research certainly is required in this area, research efforts that have examined issues of intellectual disability within the criminal justice system beginning in the early 1990s can be viewed as a step in the right direction (see Fulero & Everington, 1995). The 1990s also saw an increase in research examining issues of competency within the juvenile population. Competency has become more important for adolescents, because the juvenile justice system has shifted

to a more punishment-oriented model as opposed to a rehabilitative one, and increasing numbers of adolescents are being either waived or transferred to adult court (Salekin, Rogers, & Ustad, 2001).

Research on competency to stand trial in juveniles has examined issues related to the evaluation of developmental maturity (see Ryba, Cooper, & Zapf, 2003a), decision-making abilities (see Grisso et al., 2003), and other factors that might impact adjudicative competence (see Redlich, Silverman, & Steiner, 2003). In addition, comparisons of juveniles' competence-related abilities with those of adults has been the focus of much research and scholarly writing (see Grisso & Schwartz, 2000), and numerous studies have found age and competency to be negatively correlated, such that younger children are more likely to be found incompetent (see Cooper, 1995; Cowden & McKee, 1995; Grisso, 1998, 2003; Redlich et al., 2003). As the focus of this chapter is on competency to stand trial in adults, the reader is referred to additional sources for more in-depth information related to juvenile competency (e.g., Barnum, 2000; Grisso et al., 2003; Kruh & Grisso, 2009; Roesch, 2011; Woolard, 2002).

SPECIALIZED MEASURES OF COMPETENCY

In recent years, there has been a move toward the development of competence assessment instruments for specialized populations of defendants, such as defendants diagnosed with MR or juvenile defendants. Next we briefly describe some of the recent advances in these areas.

Competence Assessment for Standing Trial for Defendants With Mental Retardation. Everington (1990) developed an instrument designed to assess competence with defendants diagnosed with MR called the Competence Assessment for Standing Trial for Defendants with Mental Retardation (CAST-MR; Everington & Luckasson, 1992). The items of the CAST-MR were derived from a review of relevant literature, case law, and existing competency to stand trial assessment instruments (Everington, 1990). The CAST-MR consists of 50 questions, which are administered orally to the defendant. The questions are divided into three sections that address the basic elements of the *Dusky* standard. Section I, Basic Legal Concepts, includes 25 multiple-choice items that address concepts related to the criminal trial process (e.g., the roles of judges, a jury, the prosecutor, and defense attorney) and terms that are critical to the trial process (e.g., felony, plea bargain, and probation). Section II, Skills to Assist Defense, is comprised of 15 multiple-choice items that address the attorney–client relationship. Items on Sections I and II are scored as either correct (1 point) or incorrect (0 points). Section III, Understanding Case Events, consists of 10 open-ended questions designed to assess the defendant's ability to describe the relevant circumstances of his or her offense. Items are scored as 1 point, 1/2 point, or 0 points based on the ability of the individual to relay information regarding his or her case in an accurate and understandable manner (Everington & Luckasson, 1992).

The CAST-MR was developed to assist in the determination of the competency of a defendant diagnosed with MR. The authors of the CAST-MR emphasized its use as only one component of an overall assessment. Results of the CAST-MR should be considered in the context of other relevant information, such as interviews, observations, and social history (Everington & Luckasson, 1992).

The authors of this instrument conducted two validation studies to investigate the psychometric properties of the CAST-MR. The results indicated that the instrument has good reliability and validity. Reliability and validity findings were similar to those found with other competency assessment instruments (Everington, 1990; Everington & Dunn, 1995). Results from the first study demonstrated that the internal consistency of the CAST-MR total score was .93 when estimated by Cronbach's alpha and .92 when estimated by the Kuder Richardson method (Everington, 1990). The results from the second study were consistent with those of the previous study. Internal consistency of the total score using the Kuder Richardson method was estimated between .92 for KR formula 20 and .92 for KR formula 21. These findings indicate that the CAST-MR has a high level of homogeneity (Everington & Dunn, 1995). Test-retest reliability was estimated twice at .89 and .90 (Everington, 1990; Everington & Dunn, 1995). Interrater reliability for Section III was estimated between 80% and 87% (Everington & Dunn, 1995).

Juvenile Measures of Competency. The Juvenile Adjudicative Competence Interview (JACI; Grisso, 2005) is an interview guide for clinicians that provides standardized questions to assist in the evaluation of developmentally sensitive information relevant to competency to stand trial for juveniles. The JACI covers 12 primary areas of evaluation:

1. Understanding and appreciation of the charge and allegations
2. Understanding and appreciation of the purpose of the delinquency hearing
3. Understanding and appreciation of possible pleas
4. Understanding and appreciation of possible penalties
5. Understanding and appreciation of the prosecutor's role
6. Understanding and appreciation of the defense attorney's role
7. Understanding and appreciation of the juvenile probation officer's role
8. Understanding and appreciation of the judge's role
9. Ability to work with and assist counsel
10. Understanding and appreciation of the plea agreement process
11. Decision-making and reasoning abilities relevant to deciding whether to retain counsel, assist counsel, and make plea decisions
12. Ability to participate in and understand the legal proceedings, including the capacity to testify

The JACI guides examiners through the assessment of understanding, appreciation, decision making, and ability to assist counsel for the juvenile and, as items are not scored, requires a qualitative analysis of the information.

The MacArthur Judgment Evaluation (MacJEN; see Grisso et al., 2003; Woolard, 1998) was developed for use in the assessment of immaturity of judgment in juveniles. Similar to the MacCAT-CA, the MacJEN utilizes a vignette-based procedure to address three types of legal decisions typically involved in the criminal process: (1) responding to police interrogations after the commission of a crime, (2) disclosing information to one's defense attorney, and (3) responding to a plea agreement and testifying against other defendants. Respondents are to choose among various options and are asked to recommend a best and worst choice for the person in the vignette. Research studies have been conducted using the MacJEN; however, it is not available for use in clinical settings.

GUIDELINES FOR EVALUATORS

We conclude our chapter with a discussion of several issues to which examiners must pay special attention when conducting an evaluation of competency (APA, 2013; Zapf & Roesch, 2009). Prior to interviewing a defendant, it is good clinical practice to speak with the defense attorneys (or, for court-appointed evaluators, the defense and prosecuting attorneys) to determine as accurately as possible why the fitness issue was raised, what evidence was offered, and which trial and dispositional alternatives are being considered by both parties.

All indications of prior mental health contacts should be pursued *before* the interview takes place so that the examiner has as complete a set of mental health records as possible. Similarly, police reports of the alleged offense are necessary and a criminal history record is helpful, particularly if the defendant has cycled through the criminal justice and mental health systems several times. If the defendant is an inpatient, institutional records and progress notes should be reviewed as well as all available psychological test data. Finally, the examiner should maintain an accurate record of when, where, and how information about the defendant was made available as well as a date and time record of all contacts with the defendant, attorneys, other mental health professionals, and other collateral informants. These records are invaluable at later stages if legal tactics designed to confuse or mislead a witness are attempted.

Having prepared for an examination in this fashion, one can conduct an efficient and comprehensive interview in a relatively brief period of time. Most delays in conducting an evaluation and, therefore, time spent on inpatient status can be largely avoided, and a more relevant examination conducted, if these steps are taken. Prior to the interview, the defendant should be provided with proper informed consent or notification, including being informed about the parameters of confidentiality. The possibility of recording the interview should be considered.

The examiner should be aware of any aspects of the interview and the resulting report that are covered by statute or accepted practice within the jurisdiction. As an example of the former, some states require notifications that inform the defendant of the limitations of confidentiality that may apply. Similarly, other states dictate

the form of the report to the court, and an examiner's report may be excluded if it does not comply with the required format.

In *People v. Harris* (1983), for example, a psychiatrist's report (recommending that the defendant be adjudicated competent) was excluded, and the defendant's subsequent conviction was reversed because the opinion was presented in conclusory terms and failed to provide the clinical facts and reasons upon which it was based, thus precluding the trier of fact from independently assessing the weight to be given such an opinion. The current competency statutes in Illinois, Florida, and Utah are in many ways models of this developing trend, insofar as they require examiners to identify the factual bases for their conclusions and opinions, describe the defendant's mental and physical disabilities and how these may impair his or her ability to understand the proceedings and assist in the defense, discuss the likelihood that the defendant will respond to a specified course of treatment, and explain procedures that would be employed to compensate for the defendant's disabilities, if applicable. We applaud this sort of specification and urge examiners to adopt the practice, even if it is not mandated in their own jurisdiction.

Competency evaluations and the associated reports prepared for the court should be completed in accordance with both the spirit and the letter of the law. The examiner must be thoroughly acquainted with the legal literature and in some sense anticipate developments in one's practice. For example, per *Estelle v. Smith* (1981), it is clearly prohibited to introduce material obtained under court-ordered competency proceedings at a "critical" (guilt or sentencing) stage of trial. Many states mirror this in their statutes but do not regulate the common practice of requesting competency and insanity evaluations at the same time, often resulting in a combined report. We believe this practice is improper, and recommend that separate interviews, with distinct reports, be prepared when possible. The trier of fact is required to separate these issues; however, it is virtually impossible to do so when the reports are combined. A defendant who is clearly psychotic and "legally insane" at the time of an assault may respond rapidly to treatment upon arrest and be nonpsychotic and "legally fit" when actually examined. Caution and fairness dictate keeping the reports separate when possible so that the courts can consider the two issues independently.

CONCLUSIONS

This chapter touches on only a small selection of the vast amount of research and writing on competence to stand trial. The purpose of this chapter was to provide the reader with a brief overview of competency law, research, and assessment. For a comprehensive review of empirical research on competence to stand trial, the reader is referred to Grisso (1992), Cooper and Grisso (1997), and Mumley et al. (2003). These authors reviewed the research on the evaluation of competence in three 5-year intervals (1986–1990, 1991–1995, and 1996–2000). As well, Pirelli, Gottdiener, and Zapf (2011) conducted a meta-analysis using 68 studies and over

20,000 defendants that provides a quantitative review of the comparative research on competence. Pirelli, Zapf, & Gottdiener (2011) set forth a number of guidelines intended to serve as a reference for those conducting research in this area. Finally, Zapf and Roesch (2009) presented a full consideration of best practices in the evaluation of competency to stand trial. These references as well as those listed in the introductory paragraph of this chapter will provide the reader with a more in-depth understanding of competency to stand trial.

REFERENCES

- Ackerman, M. J. (1999). *Essentials of forensic psychological assessment*. New York, NY: Wiley.
- Acklin, M. W. (2012). The forensic clinician's toolbox I: A review of competency to stand trial (CST) instruments. *Journal of Personality Assessment, 94*, 220–222.
- American Psychological Association. (2013). Specialty guidelines for forensic psychology. *American Psychologist, 68*(1), 7–19. doi:10.1037/a0029889
- Anderson, S. D., & Hewitt, J. (2002). The effect of competency restoration training on defendants with mental retardation found not competent to proceed. *Law and Human Behavior, 26*, 343–351. doi:10.1023/A:1015328505884
- Bagby, R. M., Nicholson, R. A., Rogers, R., & Nussbaum, D. (1992). Domains of competency to stand trial: A factor analytic study. *Law and Human Behavior, 16*, 491–507.
- Barnum, R. (2000). Clinical and forensic evaluation of competence to stand trial in juvenile defendants. In T. Grisso & R. G. Schwartz (Eds.), *Youth on trial: A developmental perspective on juvenile justice* (pp. 193–223). Chicago, IL: University of Chicago Press.
- Berry, F. D., Jr., (1973). Self-incrimination and the compulsory mental examination: A proposal. *Arizona Law Review, 15*, 919–950.
- Bertman, L. J., Thompson, J. W., Jr., Waters, W. F., Estupinan-Kane, L., Martin, J. A., & Russell, L. (2003). Effect of an individualized treatment protocol on restoration of competency in pretrial forensic inpatients. *Journal of the American Academy of Psychiatry and Law, 31*, 27–35.
- Boddy, J., Roesch, R., Zapf, P. A., & Eaves, D. (2000, March). *Characteristics of defendants remanded for fitness assessments*. Paper presented at the biennial meeting of the American Psychology–Law Society, New Orleans, LA.
- Bonnie, R. J. (1992a). The competence of criminal defendants: A theoretical reformulation. *Behavioral Sciences and the Law, 10*, 291–316.
- Bonnie, R. J. (1992b). The competence of criminal defendants with mental retardation to participate in their own defense. *Journal of Criminal Law and Criminology, 81*, 419–446.
- Bonnie, R. J. (1993). The competence of criminal defendants: Beyond *Dusky* and *Drope*. *University of Miami Law Review, 47*, 539–601.
- Bonnie, R. J. & Grisso, T. (2000). Adjudicative competence and youthful offenders. In T. Grisso & R. G. Schwartz (Eds.), *Youth on trial: A developmental perspective on criminal justice* (pp. 73–103). Chicago, IL: University of Chicago Press.
- Borum, R., & Grisso, T. (1995). Psychological test use in criminal forensic evaluations. *Professional Psychology: Research and Practice, 26*, 465–473.
- Brakel, S. J. (1974). Presumption, bias, and incompetency in the criminal process. *Wisconsin Law Review, 1974*, 1105–1130.

- Bukatman, B. A., Foy, J. L., & de Grazia, E. (1971). What is competency to stand trial? *American Journal of Psychiatry*, *127*, 1225–1229.
- Cochrane, R. E., Herbel, B. L., Reardon, M. L., & Lloyd, K. P. (2012, July 2). The Sell effect: Involuntary medication treatment is a “clear and convincing” success. *Law and Human Behavior*. Advance online publication. doi:10.1037/lhb0000003
- Cooper, D. K. (1995). Juvenile competency to stand trial: The effects of age and presentation of factual information in the attainment of competency in juveniles. *Dissertation Abstracts International*, *56*(10-B), 5761.
- Cooper, D. K., & Grisso, T. (1997). Five-year research update (1991–1995): Evaluations for competence to stand trial. *Behavioral Sciences and the Law*, *15*, 347–364.
- Cowden, V. L., & McKee, G. R. (1995). Competency to stand trial in juvenile delinquency proceedings—cognitive maturity and the attorney-client relationship. *University of Louisville Journal of Family Law*, *33*, 629–660.
- Criminal Code of Canada, R. S. C., C-46 (1992).
- Davis, D. L. (1985). Treatment planning for the patient who is incompetent to stand trial. *Hospital and Community Psychiatry*, *36*, 268–271.
- Davis, S. (1994). Fitness to stand trial in Canada in light of the recent Criminal Code amendments. *International Journal of Law and Psychiatry*, *17*, 319–329.
- Dickey, W. (1980). Incompetency and the nondangerous mentally ill client. *Criminal Law Bulletin*, *16*, 22–40.
- Drope v. Missouri*, 420 U.S. 162 (1975).
- Dusky v. United States*, 362 U.S. 402 (1960).
- Edgerton, R. B. (1993). *The cloak of competence* (2nd ed.). Berkeley: University of California Press.
- Estelle v. Smith*, 49 U.S.L.W. 4490 (1981).
- Everington, C. T. (1990). The Competence Assessment for Standing Trial for Defendants with Mental Retardation (CAST-MR): A validation study. *Criminal Justice and Behavior*, *17*, 147–168.
- Everington, C., & Dunn, C. (1995). A second validation study of the competence assessment for standing trial for defendants with mental retardation (CAST-MR). *Criminal Justice and Behavior*, *22*, 44–59.
- Everington, C. T., & Luckasson, R. (1992). *Competence assessment for standing trial for defendants with mental retardation*. Worthington, OH: IDS.
- Everington, C., Notario-Smull, H., & Horton, M. L. (2007). Can defendants with mental retardation successfully fake their performance on a test of competence to stand trial? *Behavioral Sciences & the Law*, *25*, 545–560.
- Farkas, G., DeLeon, P., & Newman, R. (1997). Sanity examiner certification: An evolving national agenda. *Professional Psychology: Research & Practice*, *28*, 73–76.
- Favole, R. J. (1983). Mental disability in the American criminal process: A four issue survey. In J. Monahan & H. J. Steadman (Eds.), *Mentally disordered offenders: Perspectives from law and social science* (pp. 247–295). New York, NY: Plenum Press.
- Fitzgerald, J. F., Peszke, M. A., & Goodwin, R. C. (1978). Competency evaluations in Connecticut. *Hospital and Community Psychiatry*, *29*, 450–453.
- Fulero, S. & Everington, C. (1995). Assessing competency to waive Miranda rights in defendants with mental retardation. *Law and Human Behavior*, *19*, 533–543.
- Godinez v. Moran*, 113 St. C. 2680 (1993).

- Golding, S. L. (1992). Studies of incompetent defendants: Research and social policy implications. *Forensic Reports*, 5, 77–83.
- Golding, S. L. (1993). *Interdisciplinary Fitness Interview—Revised: A training manual*. Department of Psychology, University of Utah, Salt Lake City, UT. Available at http://home.comcast.net/~slgolding/publications/ifir_manual.htm
- Golding, S. L., & Roesch, R. (1988). Competency for adjudication: An international analysis. In D. N. Weisstub (Ed.), *Law and mental health: International perspectives* (Vol. 4, pp. 73–109). New York, NY: Pergamon Press.
- Golding, S. L., Roesch, R., & Schreiber, J. (1984). Assessment and conceptualization of competency to stand trial: Preliminary data on the Interdisciplinary Fitness Interview. *Law and Human Behavior*, 8, 321–334.
- Goldstein, R. L., & Stone, M. (1977). When doctors disagree: Differing views on competency. *Bulletin of the American Academy of Psychiatry and the Law*, 5, 90–97.
- Government of Virgin Islands v. Charles, No. 94–7638 (U.S. Ct. App. 1995).
- Gowensmith, W. N., Murrie, D. C., & Boccaccini, M. T. (2012). Field reliability of competence to stand trial opinions: How often do evaluators agree, and what do judges decide when evaluators disagree? *Law and Human Behavior*, 36, 130–139.
- Grisso, T. (1986). *Evaluating competencies: Forensic assessments and instruments*. New York, NY: Plenum Press.
- Grisso, T. (1992). Five-year research update (1986–1990): Evaluations for competence to stand trial. *Behavioral Sciences and the Law*, 10, 353–369.
- Grisso, T. (1998). *Forensic evaluation of juveniles*. Sarasota, FL: Professional Resource Press.
- Grisso, T. (2003). *Evaluating competencies: Forensic assessments and instruments* (2nd ed.). New York, NY: Kluwer Academic/Plenum Press.
- Grisso, T. (2005). *Evaluating juveniles' adjudicative competence: A guide for clinical practice*. Sarasota, FL: Professional Resource Press.
- Grisso, T., Appelbaum, P., Mulvey, E., & Fletcher, K. (1995). The MacArthur treatment competence study II: Measures of abilities related to competence to consent to treatment. *Law & Human Behavior*, 19, 127–148.
- Grisso, T., Cocozza, J. J., Steadman, H. J., Fisher, W. H., & Greer, A. (1994). The organization of pretrial forensic evaluation services: A national profile. *Law and Human Behavior*, 18, 377–393.
- Grisso, T., & Schwartz, R. G. (2000). *Youth on trial: A developmental perspective on juvenile justice*. Chicago, IL: University of Chicago Press.
- Grisso, T., Steinberg, L., Woolard, J., Cauffman, E., Scott, E., Graham, S., Lexcen, F., . . . Schwartz, R. (2003). Juveniles' competence to stand trial: A comparison of adolescents' and adults' capacities as trial defendants. *Law and Human Behavior*, 27, 333–363.
- Hart, S. D., & Hare, R. D. (1992). Predicting fitness for trial: The relative power of demographic, criminal and clinical variables. *Forensic Reports*, 5, 53–65.
- Heilbrun, K., & Collins, S. (1995). Evaluations of trial competency and mental state at time of offense: Report characteristics. *Professional Psychology: Research and Practice*, 26, 61–67.
- Heilbrun, K., Marczyk, G. R., & DeMatteo, D. (2002). *Forensic mental health assessment: A casebook*. New York, NY: Oxford University Press.
- Hoge, S. K., Bonnie, R. J., Poythress, N., & Monahan, J. (1992). Attorney-client decision-making in criminal cases: Client competence and participation as perceived by their attorneys. *Behavioral Sciences and the Law*, 10, 385–394.

- Hoge, S. K., Bonnie, R. J., Poythress, N., & Monahan, J. (1999). *The MacArthur Competence Assessment Tool—Criminal Adjudication*. Odessa, FL: Psychological Assessment Resources.
- Hoge, S. K., Bonnie, R. J., Poythress, N., Monahan, J., Eisenberg, M., & Feucht-Haviar, T. (1997). The MacArthur Adjudicative Competence Study: Development and validation of a research instrument. *Law and Human Behavior, 21*, 141–179.
- Indiana v. Edwards, 554 U.S. 164 (2008).
- Jackson v. Indiana, 406 U.S. 715 (1972).
- Johnson, W. G., & Mullett, N. (1987). Georgia Court Competency Test–R. In M. Hersen & A. S. Bellack (Eds.), *Dictionary of behavioral assessment techniques* (p. 234). New York, NY: Pergamon Press.
- Jones, B. L. (1994). *Factors related to the competency to stand trial of defendants who are mentally retarded*. Unpublished doctoral dissertation, Florida State University, Tallahassee, FL.
- Kruh, I., & Grisso, T. (2009). *Evaluations of juveniles' competence to stand trial*. New York, NY: Oxford University Press.
- Laben, J. K., Kashgarian, M., Nessa, D. B., & Spencer, L. D. (1977). Reform from the inside: Mental health center evaluations of competency to stand trial. *Journal of Clinical Psychology, 5*, 52–62.
- Lipsitt, P. D., Lelos, D., & McGarry, A. L. (1971). Competency for trial: A screening instrument. *American Journal of Psychiatry, 128*, 105–109.
- McGarry, A. L. (1965). Competency for trial and due process via the state hospital. *American Journal of Psychiatry, 122*, 623–631.
- Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C., with Lyons, P. M., Jr., & Otto, R. K. (2007). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (3rd ed.). New York, NY: Guilford Press.
- Menzies, R. J., Webster, C. D., Butler, B. T., & Turner, R. C. (1980). The outcome of forensic psychiatric assessment: A study of remands in six Canadian cities. *Criminal Justice and Behavior, 7*, 471–480.
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist, 35*, 1012–1027.
- Missouri v. Davis, 653 S.W. 2d. 167 (Mo. Sup. Ct. 1983).
- Montana v. Austad, 641 P. 2d. 1373 (Mont. Sup. Ct. 1982).
- Morris, G. H., Haroun, A. M., & Naimark, D. (2004). Assessing competency competently: Toward a rational standard for competency-to-stand-trial assessments. *Journal of the American Academy of Psychiatry and Law, 32*, 231–245.
- Morrow v. Maryland, 443 A. 2d. 108 (Md. Ct. App. 1982).
- Mumley, D. L., Tillbrook, C. E., & Grisso, T. (2003). Five year research update (1996–2000): Evaluations for competence to stand trial (adjudicative competence). *Behavioral Sciences and the Law, 21*, 329–350.
- Murrie, D. C., Boccaccini, M. T., Zapf, P. A., Warren, J. I., & Henderson, C. E. (2008). Clinician variation in findings of competence to stand trial. *Psychology, Public Policy, and Law, 14*, 177–193.
- Nicholson, R. A., Briggs, S. R., & Robertson, H. C. (1988). Instruments for assessing competency to stand trial: How do they work? *Professional Psychology: Research & Practice, 19*, 383–394.
- Nicholson, R. A., Robertson, H. C., Johnson, W. G., & Jensen, G. (1988). A comparison of instruments for assessing competency to stand trial. *Law and Human Behavior, 12*, 313–321.

- Nicholson, R. A., & Kugler, K. E. (1991). Competent and incompetent criminal defendants: A quantitative review of comparative research. *Psychological Bulletin*, *109*, 355–370.
- Nicholson, R., & Norwood, S. (2000). The quality of forensic psychological assessments, reports, and testimony: Acknowledging the gap between promise and practice. *Law and Human Behavior*, *24*, 9–44.
- North Dakota v. Heger, 326 N.W. 2d 855 (1982).
- Nottingham, E. J., & Mattson, R. E. (1981). A validation study of the competency screening test. *Law and Human Behavior*, *5*, 329–336.
- Otto, R. K., Musick, J. E., & Sherrod, C. B. (2010). *Inventory of Legal Knowledge professional manual*. Lutz, FL: PAR.
- Otto, R. K., Poythress, N. G., Nicholson, R. A., Edens, J. F., Monahan, J., Bonnie, R. J., . . . Eisenberg, M. (1998). Psychometric properties of the MacArthur Competence Assessment Tool—Criminal Adjudication (MacCAT-CA). *Psychological Assessment*, *10*, 435–443.
- Pate v. Robinson, 383 U.S. 375 (1966).
- Pendleton, L. (1980). Treatment of persons found incompetent to stand trial. *American Journal of Psychiatry*, *137*, 1098–1100.
- People v. Harris, 133 Ill. App. 3d. 633 (1983).
- Perlin, M. L. (1996). “Dignity was the first to leave”: Godinez v. Moran, Colin Ferguson, and the trial of mentally disabled criminal defendants. *Behavioral Sciences and the Law*, *14*, 61–81.
- Petrella, R. C. (1992). Defendants with mental retardation in the forensic services system. In R. W. Conley, R. Luckasson, & G. N. Bouthilet (Eds.), *The criminal justice system and mental retardation* (pp. 79–96). Baltimore, MD: Brookes.
- Pirelli, G., Gottdiener, W. H., & Zapf, P. A. (2011). A meta-analytic review of competency to stand trial research. *Psychology, Public Policy and Law*, *17*, 1–53.
- Pirelli, G., Zapf, P. A., & Gottdiener, W. H. (2011). Competency to stand trial research: Guidelines and future directions. *Journal of Forensic Psychiatry and Psychology*, *22*, 340–370.
- Pizzi, W. T. (1977). Competency to stand trial in federal courts: Conceptual and constitutional problems. *University of Chicago Law Review*, *45*, 20–71.
- Poythress, N., Nicholson, R., Otto, R. K., Edens, J. F., Bonnie, R. J., Monahan, J., & Hoge, S. K. (1999). *The MacArthur Competence Assessment Tool—Criminal Adjudication: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Poythress, N. G., & Stock, H. V. (1980). Competency to stand trial: A historical review and some new data. *Psychiatry and Law*, *8*, 131–146.
- Randolph, J. J., Hicks, T., & Mason, D. (1981). The Competency Screening Test: A replication and extension. *Criminal Justice and Behavior*, *8*, 471–482.
- Redlich, A., D., Silverman, M., & Steiner, H. (2003). Pre-adjudicative and adjudicative competence in juveniles and young adults. *Behavioral Sciences and the Law*, *21*, 393–410.
- Regina v. Whittle, 2S.C.R. 914 (1994).
- Riggins v. Nevada 112S.Ct. 1810 (1992).
- Ritchie v. Indiana, 468 N.E. 2d. 1369 (Ind. Sup. Ct. 1984).
- Robey, A. (1965). Criteria for competency to stand trial: A checklist for psychiatrists. *American Journal of Psychiatry*, *122*, 616–623.
- Robinson, R., & Acklin, M. W. (2010). Fitness in paradise: Quality of forensic reports submitted to the Hawaii judiciary. *International Journal of Law and Psychiatry*, *33*, 131–137.

- Roesch, R. (2011). Young offenders and legal competencies. In T. Bliesener, A. Beelmann, & M. Stemmler (Eds.), *Antisocial behavior and crime: Contributions of developmental and evaluation research to prevention and intervention* (pp. 275–290). Cambridge, England: Hogrefe.
- Roesch, R., & Golding, S. L. (1980). *Competency to stand trial*. Urbana: University of Illinois Press.
- Roesch, R., & Golding, S. L. (1985). The impact of deinstitutionalization. In D. P. Farrington & J. Gunn (Eds.), *Current research in forensic psychiatry and psychology* (pp. 209–239). New York, NY: Wiley.
- Roesch, R., Webster, C. D., & Eaves, D. (1984). *The Fitness Interview Test: A method for assessing fitness to stand trial*. Toronto, Canada: University of Toronto Centre of Criminology.
- Roesch, R., Zapf, P. A., & Eaves, D. (2006). *Fitness Interview Test: A structured interview for assessing competency to stand trial*. Sarasota, FL: Professional Resource Press.
- Rogers, R., Grandjean, N., Tillbrook, C. E., Vitacco, M. J., & Sewell, K. W. (2001). Recent interview-based measures of competency to stand trial: A critical review augmented with research data. *Behavioral Sciences and the Law*, 19, 503–518.
- Rogers, R., Jackson, R. L., Sewell, K. W., Tillbrook, C. E., & Martin, M. A. (2003). Assessing dimensions of competency to stand trial: Construct validation of the ECST-R. *Assessment*, 10, 344–351.
- Rogers, R. R., & Mitchell, C. N. (1991). *Mental health experts and the criminal courts*. Scarborough, Canada: Carswell.
- Rogers, R., Tillbrook, C. E., & Sewell, K. W. (2004). *Evaluation of Competency to Stand Trial—Revised (ECST-R) and professional manual*. Odessa, FL: Psychological Assessment Resources.
- Ryba, N. L., Cooper, V. G., & Zapf, P. A. (2003a). Assessment of maturity in juvenile competency to stand trial evaluations: A survey of practitioners. *Journal of Forensic Psychology Practice*, 3, 23–45.
- Ryba, N. L., Cooper, V. G., & Zapf, P. A. (2003b). Juvenile competence to stand trial evaluations: A survey of current practices and test usage among psychologists. *Professional Psychology: Research and Practice*, 34, 499–507.
- Salekin, R. T., Rogers, R., & Ustad, K. L. (2001). Juvenile waiver to adult criminal courts: Prototypes for dangerousness, sophistication-maturity, and amenability to treatment. *Psychology, Public Policy, and Law*, 7, 381–408.
- Schreiber, J. (1978). Assessing competency to stand trial: A case study of technology diffusion in four states. *Bulletin of the American Academy of Psychiatry and the Law*, 6, 439–457.
- Sell v. United States, 282 F. 3d 560 (2003).
- Shatin, L. (1979). Brief form of the Competency Screening Test for Mental Competence to Stand Trial. *Journal of Clinical Psychology*, 35, 464–467.
- Siegel, A. M., & Elwork, A. (1990). Treating incompetence to stand trial. *Law and Human Behavior*, 14, 57–65.
- Silten, P. R., & Tullis, R. (1977). Mental competency in criminal proceedings. *Hastings Law Journal*, 28, 1053–1074.
- Skeem, J., Golding, S. L., Cohn, N., & Berge, G. (1998). Logic and reliability of evaluations of competence to stand trial. *Law and Human Behavior*, 22, 519–547.
- Stafford, K. P. (2003). Assessment of competence to stand trial. In A. M. Goldstein (Ed.), *Handbook of psychology: Forensic psychology* (Vol. 11, pp. 359–380). Hoboken, NJ: Wiley.

- Steadman, H. J. (1979). *Beating a rap?: Defendants found incompetent to stand trial*. Chicago, IL: University of Chicago Press.
- Teplin, L. (1984). Criminalizing mental disorder: The comparative arrest rate of the mentally ill. *American Psychologist*, *39*, 794–803.
- Utah Code Annotated §77-15-1 et seq. (1994).
- Ustad, K. L., Rogers, R., Sewell, K. W., & Guarnaccia, C. A. (1996). Restoration of competency to stand trial: Assessment with the Georgia Court Competency Test and the Competency Screening Test. *Law and Human Behavior*, *20*, 131–146.
- Vann, C. R., & Morganroth, F. (1965). The psychiatrist as judge: A second look at the competence to stand trial. *University of Detroit Law Journal*, *43*, 1–12.
- Verdun-Jones, S. N. (1981). The doctrine of fitness to stand trial in Canada: The forked tongue of social control. *International Journal of Law and Psychiatry*, *4*, 363–389.
- Viljoen, J. L., Roesch, R., & Zapf, P. A. (2002). Interrater reliability of the Fitness Interview Test across four professional groups. *Canadian Journal of Psychiatry*, *47*, 945–952.
- Viljoen, J. L., Vincent, G. M., & Roesch, R. (2006). Assessing adolescent defendants' adjudicative competence: Interrater reliability and factor structure of the Fitness Interview Test—Revised. *Criminal Justice and Behavior*, *33*, 449–466.
- Viljoen, J., Penner, E., & Roesch, R. (2012). Competence and criminal responsibility in adolescent defendants: The role of mental illness and adolescent development. In B. C. Feld & D. Bishop (Eds.), *The Oxford handbook on juvenile crime and juvenile justice* (pp. 526–548). New York, NY: Oxford University Press.
- Wall, B. W., Krupp, B. H., & Guilmette, T. (2003). Restoration of competency to stand trial: A training program for persons with mental retardation. *Journal of the American Academy of Psychiatry and Law*, *31*, 189–210.
- Washington v. Harper, 494 U.S. 210 (1990).
- Webster, C. D., Jenson, F. A. S., Stermac, L., Gardner, K., & Slomen, D. (1985). Psychoeducational programmes for forensic psychiatric patients. *Canadian Psychology*, *26*, 50–53.
- Webster, C. D., Menzies, R. J., & Jackson, M. A. (1982). *Clinical assessment before trial*. Toronto, Canada: Butterworths.
- Wieter v. Settle, 193 F. Supp. 318 (W.D. Mo. 1961).
- Wildman, R. W., Batchelor, E. S., Thompson, L., Nelson, F. R., Moore, J. T., Patterson, M. E., & deLaosa, M. (1978). *The Georgia Court Competency Test: An attempt to develop a rapid, quantitative measure of fitness for trial*. (Unpublished manuscript). Forensic Services Division, Central State Hospital, Milledgeville, GA.
- Wildman, R. W., II, White, P. A., & Brandenburg, C. A. (1990). The Georgia Court Competency Test: The base rate problem. *Perceptual and Motor Skills*, *70*, 1055–1058.
- Wilson v. United States, 391 F. 2d. 460 (1968).
- Winick, B. J. (1983). Incompetency to stand trial: Developments in the law. In J. Monahan & H. J. Steadman (Eds.), *Mentally disordered offenders* (pp. 3–38). New York, NY: Plenum Press.
- Winick, B. J. (1985). Restructuring competency to stand trial. *UCLA Law Review*, *32*, 921–985.
- Winick, B. J. (1995). Reforming incompetency to stand trial and plead guilty: A restated proposal and a response to Professor Bonnie. *Journal of Criminal Law and Criminology*, *85*, 571–624.

- Winick, B. J. (1996). Incompetency to proceed in the criminal process: Past, present, and future. In D. B. Wexler & B. J. Winick (Eds.), *Law in a therapeutic key: Developments in therapeutic jurisprudence* (pp. 77–111). Durham, NC: Carolina Academic Press.
- Woolard, J. L. (1998). *Developmental aspects of judgment and competence in legally relevant contexts*. (Unpublished doctoral dissertation). Psychology Department, University of Virginia, Charlottesville.
- Woolard, J. L. (2002). Capacity, competence, and the juvenile defendant: Implications for research and policy. In B. L. Bottoms & M. Bull Kovera (Eds.), *Children, social science, and the law* (pp. 270–298). New York, NY: Cambridge University Press.
- Zapf, P. A., Hubbard, K. L., Cooper, V. G., Wheelles, M. C., & Ronan, K. A. (2004). Have the courts abdicated their responsibility for determination of competency to stand trial to clinicians? *Journal of Forensic Psychology Practice, 4*, 27–44.
- Zapf, P. A., & Roesch, R. (1997). Assessing fitness to stand trial: A comparison of institution-based evaluations and a brief screening interview. *Canadian Journal of Community Mental Health, 16*, 53–66.
- Zapf, P. A., & Roesch, R. (1998). Fitness to stand trial: Characteristics of fitness remands since the 1992 Criminal Code amendments. *Canadian Journal of Psychiatry, 43*, 287–293.
- Zapf, P. A., & Roesch, R. (2001). A comparison of the MacCAT-CA and the FIT for making determinations of competency to stand trial. *International Journal of Law and Psychiatry, 24*, 81–92.
- Zapf, P. A., & Roesch, R. (2009). *Guide to best practices for forensic mental health assessments: Competency to stand trial*. New York, NY: Oxford University Press.
- Zapf, P. A., Roesch, R., & Viljoen, J. L. (2001). Assessing fitness to stand trial: The utility of the Fitness Interview Test (rev. ed.). *Canadian Journal of Psychiatry, 46*, 426–432.
- Zapf, P. A., & Viljoen, J. L. (2003). Issues and considerations regarding the use of assessment instruments in the evaluation of competency to stand trial. *Behavioral Sciences and the Law, 21*, 351–367.
- Zapf, P. A., Viljoen, J. L., Whittemore, K. E., Poythress, N. G., & Roesch, R. (2002). Competency: Past, present, and future. In J. R. P. Ogloff (Ed.), *Taking psychology and law into the twenty-first century* (pp. 171–198). Kluwer Academic/Plenum Press.

Assessing Criminal Responsibility

PATRICIA A. ZAPF, STEPHEN L. GOLDING, RONALD ROESCH, AND GIANNI PIRELLI

MUCH has been written about criminal responsibility and issues of insanity or mental state at the time of the offense. In this chapter, we focus on three major areas:

1. Insanity standards and the construal of criminal responsibility, including a brief review of the historical and jurisprudential roots of culpable *mens rea*, an overview of the evolution of the legal standards for the insanity defense, and an examination of the movement to reform the insanity defense, particularly by adopting guilty-but-mentally-ill verdict options
2. A review of issues related to the assessment of criminal responsibility, including the structure of these evaluations, instruments developed to guide evaluators, the role of psychotic symptoms, particularly delusions, in the evaluation of criminal responsibility, and issues pertaining to the treatment and release of insanity acquittees
3. An overview of the empirical developments regarding criminal responsibility, including research on not guilty by reason of insanity (NGRI) verdicts, judicial instruction, and jury decision making

As it is impossible to provide a comprehensive review in one chapter, the interested reader is also referred to additional resources for an in-depth understanding of this topic area (see especially Borum, 2003; Melton, Petrila, Poythress, & Slobogin, 2007; Packer, 2009; Rogers & Shuman, 2000).

INSANITY STANDARDS AND THE CONSTRUAL OF CRIMINAL RESPONSIBILITY

The complexity of arguments, philosophical debates, opinions, and data on the insanity defense cannot be approached without a personal decision to accept or

reject a rather simple thesis. Belief in this basic thesis is not subject to scientific argument; rather, it is morally axiomatic. That is, one either accepts it as a function of one's fundamental moral, religious, and jurisprudential presuppositions, or one does not. Given the nonprovable nature of this moral thesis, scientific and logical arguments about aspects of the insanity defense and the assessment of mental state are possible, but acceptance or rejection of the argument is not a matter of proof or science. This fundamental belief may be stated in this way:

In cognizing and regulating social interactions in terms of fundamental principles of "fairness" and "justice," we assume that all such social interactions, including the societal judgment of criminal or civil responsibility for certain classes of proscribed behavior, are based upon an ethical calculus that assigns individual blame, culpability, liability, punishability, and moral and criminal responsibility as a function of intentionality and mental capacity. The classical formulation of this moral presupposition is the legal maxim *Actus non facit reum, nisi mens sit rea*, which translates freely into modern English as "An act is not legally cognizable as evil, and hence criminally punishable, unless it is committed by a person who has the capacity to cognize the act as evil and then freely chooses to do it." (Golding & Roesch, 1987, p. 395)

This fundamental belief goes to the heart of the tension in the public's mind, as well as in criminal and civil law, between strict or objective liability, on one hand, and subjective liability, on the other. An examination of the history of the criminal law in Western Judeo-Christian cultures clearly demonstrates the nature of this tension (see Crotty, 1924; LaFave & Scott, 1972). The dilemma is simply this: On one hand, it is clear that, when someone performs a heinous or reprehensible act, he or she is "guilty" in the commonsense meaning of that term (i.e., objective liability). On the other hand, to have a theory of action and responsibility that embodies our cultural sense of "fairness" and "justice," one that reflects our increasing knowledge of psychological processes, in general, and psychotic processes, in particular, we have to consider the conjunction of the proscribed behavior (*actus reus*) and an appropriate degree and type of intentionality and mental capacity (*mens rea*) in ascribing guilty or culpable ownership of an act (i.e., subjective liability).

We will not attempt to address the moral question about the insanity defense—whether it should exist or not. Our belief, along with many scholars who have examined this issue, is that its existence is integral to the fabric of our social structure, which includes, but is not limited to, the structure of our criminal law. Rather, we outline issues relevant to mental health professionals who are called on to evaluate defendants and offer expert opinions on the issue of criminal responsibility.

CONCEPT OF *MENS REA*

It is well established within the historical and jurisprudential literatures that the fundamental concept of *mens rea* within Judeo-Christian cultures has been in

existence since the earliest recordings of Hebrew law (e.g., Gray, 1972; Sayre, 1932; Stroud, 1914). Platt and Diamond (1966), for example, quoted the Babylonian Talmud as observing that “[a] deaf-mute, an idiot and a minor are awkward to deal with, as he who injures them is liable [to pay], whereas if they injure others they are exempt” (note 7, p. 1228). This concept may be traced, in a continuous line of development, through Greek and Roman law wherein the concept of *culpa* (negligence) is distinguished from *dolus* (intentional fraud). Children under the age of 7, for example, were considered *doli incapax*, that is, “not possessed of sufficient discretion and intelligence to distinguish between right and wrong” and hence “incapable of criminal intention or malice” (Black, 1979, p. 433). Children ages 7 to 12 were presumed *doli incapax* unless evidence of capacity to form culpable intention was presented. Interestingly, the pattern of evidence most frequently adduced to infer such intentionality, such as lying about the crime, concealing the body, or other such after-the-fact actions, is still used in modern insanity trials as evidence that the person was capable of the prerequisite intentionality at the time of the crime. The culmination of this doctrine in more “modern” (i.e., since the 13th century) jurisprudence is presented in *Blackstone’s Commentaries* in its classic form:

All the several pleas and excuses, which protect the committer of a forbidden act from the punishment which is otherwise annexed thereto, may be reduced to this single consideration, the want or defect of will. An involuntary act, as it has no claim to merit, so neither can it induce any guilt: the concurrence of the will, when it has its choice either to do or to avoid the fact in question, being the only thing that renders human actions either praiseworthy or culpable. Indeed, to make a complete crime, cognizable by human laws, there must be both a will and an act. . . . For The rule of law as to . . . [lunatics] is, that “*furiosus furore solum punitur*” [“madness alone punishes a madman”]. In criminal cases therefore idiots and lunatics are not chargeable for their own acts, if committed when under these incapacities: no, not even for treason itself. (Blackstone’s *Commentaries*, Book 4, Chapter 2, pp. 20, 24; also cited in *State v. Strasburg*, 1910, pp. 1021–1022)

One can show that the entire structure of the criminal law is built on this principle. Sayre (1932) observed that, whereas one of the earliest legal texts, *Leges Henrici Primi* (*The Laws of Henry I*), alternated between advocating absolute liability—“he who commits evil unknowingly must pay for it knowingly”—and advocating the principle of *mens rea*, it was standard practice for the king to either pardon mentally disordered persons found guilty of “absolute liability crimes” or for other financial arrangements to be made. In fact, in Sayre’s classic review of *mens rea*, it was argued that the tradition of criminal law in England dating back to Henry I originated in theological opposition to secular laws of absolute responsibility. This theological opposition was based on a belief that God could not properly hold an infant, idiot, or lunatic justly responsible. It is interesting to note that the age at which children generally are assumed to be criminally responsible corresponds to the age within all major religions at which they usually pass through a “certification” ritual where

they are deemed morally responsible in the eyes of God. Platt and Diamond (1965, 1966) showed in their historical reviews that the “furiously” insane have been exempted from moral sanction by an extension of the same logic.

Whereas *mens rea* has been historically interpreted in a broad fashion, making it roughly synonymous with “culpable intentionality” (Stroud, 1914, p. 13) or with the general mental and emotional capacity prerequisite to choose freely to commit proscribed acts, the modern trend in criminal law has been to interpret the *mens rea* requirement of criminal conduct more narrowly and equate it with such phrases as proscribed conduct performed *intentionally, recklessly, knowingly, or purposefully*. Discomfort with the insanity defense sometimes has resulted in attempts to either abolish it outright or to change it drastically by restricting the relevance of mental state to such a narrowly defined *mens rea*.

The most comprehensive scholarly review of this narrowing approach was conducted by Wales (1976), who discussed the problem using the well-known metaphor of squeezing a lemon (i.e., a defendant, under this narrow view, would not be guilty of killing his wife, if, while strangling her, he believed he was merely squeezing a lemon). In other words, the prototypic case envisioned as qualifying for exculpation under the narrow view would be delusional mistake of fact. In discussing the legislative history of narrowing attempts, Wales made it clear that the underlying motivation is to eliminate the insanity defense without raising constitutional considerations and to ensure that more “insanity-like” acquittees are first found guilty and then found to be in need of treatment or mentally ill. Wales argued that the cases most likely to be acted on differently are those involving command hallucinations, delusions, and paranoid processes, where it is clear that the defendant acted “knowingly” in the narrow sense of the term, but that the “knowingness” was dependent on delusional, hallucinatory, or otherwise psychotic belief systems.

Beginning with the first Nixon administration, there have been many attempts by federal and state legislatures to accomplish this shift in the focus of the insanity defense.¹ In addition to the narrowing of the concept of *mens rea*, some states have experimented with giving decision makers an in-between verdict, namely guilty but mentally ill (GBMI), which, in reality, is simply a guilty verdict made without the guarantee of differences in sentencing, disposition, or mental health treatment (see Bumby, 1993; Golding, 1992; Golding & Roesch, 1987). Finally, some states continue to experiment with varying levels of abolition. In 2004, Idaho effectively abolished an insanity defense through statutes providing that “[m]ental condition shall not be a defense to any charge of criminal conduct” (Idaho Code, §18–207(1)). A writ of certiorari was recently denied in *Delling v. Idaho* (2012), thus leaving unanswered the question of whether the U.S. Constitution permits a state to abolish the insanity defense in criminal cases. Currently, the modal insanity defense criteria in those 46 states that recognize some version of the insanity defense involve either

1. See Golding and Roesch (1987) for an analysis of the historical origins of this trend, beginning with Queen Victoria.

the traditional American Law Institute (ALI) formulation (with or without the “volitional” prong) or restricted versions of the traditional M’Naghten test (see the next section for a discussion of the historical evolution of various legal standards).

EVOLUTION OF LEGAL STANDARDS FOR CRIMINAL RESPONSIBILITY

M’Naghten’s trial is assumed to be the starting place for the test that:

to establish a defense on the ground of insanity, it must be clearly proved that, at the time of the committing of the act, the party accused was labouring under such a defect of reason, from disease of the mind, as not to know the nature and quality of the act he was doing; or, if he did know it, that he did not know he was doing what was wrong. (M’Naghten’s Case, 1843, p. 722)

However, it is reasonably clear that the knowledge/right-wrong test had already been used implicitly and explicitly in a series of trials in both England and the United States. In fact, there was already considerable discomfort with the perceived “narrow scope” of the rule. Isaac Ray (1802/1838) had already published his *Treatise on the Medical Jurisprudence of Insanity* in which he attacked the narrowness of such formulations as not according with modern knowledge of the forms of mental disorder and their influence on behavior, affect, and cognition. The same debates that occur today over the scope of what should be included under *knowledge*, *appreciation*, and the like were influential in court decisions of the day.

Although the M’Naghten rules were rapidly adopted in the United States, they were almost immediately subjected to challenge on the “narrowness” ground and were modified significantly by some jurisdictions. In 1844, Chief Justice Shaw of the Massachusetts Supreme Court held that, although the “right-wrong” test was proper, a defendant who acted under the influence of an irresistible impulse was not a free agent and, hence, was included under the rule, because he or she could not know right from wrong (*Commonwealth v. Rogers*, 1844). In 1866, Justice Somerville made this logic explicit in the ruling set forth in *Parsons v. State* (1866):

If therefore, it be true, as a matter of fact, that the disease of insanity can . . . so affect the mind as to subvert the freedom of the will, and thereby destroy the power of the victim to choose between right and wrong, although he perceived it—by which we mean the power of volition to adhere in action to the right and abstain from wrong—is such a one criminally responsible for an act done under the influence of such a controlling disease? We clearly think not. (p. 586)

In 1924, Crotty documented that the jurisdictions in the United States were fragmented into four sets of rules: (1) relatively “pure” M’Naghten standard; (2) the M’Naghten standard broadened by interpretation to include irresistible impulse as meeting the test; (3) the M’Naghten standard supplemented by explicit irresistible impulse rules; and (4) the New Hampshire “product” rule, heavily influenced by Isaac Ray and set forth in *State v. Pike* (1869).

Justice Doe, in setting forth New Hampshire's product test, argued that it was a matter of legal fact, to be decided by a judge or jury, whether a defendant suffered from a disease of the mind and whether the proscribed behavior was a product of that disease. As such, he discarded formal rules of specific states of mind and asserted that it was up to the trier of fact to decide "if [the alleged crime] was the offspring or product of mental disease in the defendant, (then) he was not guilty by reason of insanity" (1954, p. 442). In fact, for a long period of time following M'Naghten, there was considerable controversy over insanity rules that surfaced repeatedly.

Charles Guiteau's assassination of President Garfield in 1881 gave rise to a highly controversial trial and execution that took place against the background of a strong concern over "irresistible impulses" and a belief that insanity, especially "moral insanity," was all too easy to feign (Rosenberg, 1968). Rosenberg's scholarly analysis drew out these issues in fine detail, documenting public, legislative, and psychiatric reactions that are reminiscent of the current debate. Judge Cox's highly elaborate instructions to the jury in Guiteau's trial left little doubt that the central issue before the jury concerned whether the alleged moral insanity and irresistible impulse fit into a straightforward interpretation of the right/wrong test outlined in M'Naghten. Whatever the jury may have thought of the battle of the experts and the problems of the insanity defense, however, Guiteau placed himself in fatal jeopardy when he exhibited his uncontested egocentrism and interrupted the prosecutor toward the close of the trial by saying "That is not the issue. The issue is, was my free agency destroyed? I was overpowered. That is what the jury is to pass on" (cited in Rosenberg, 1968, p. 201). His rational comments may have prompted the jury to focus on his current mental state, a problem often confronted by defendants offering an insanity defense.

Controversy surrounding various definitional and procedural aspects of the insanity defense continued over the first half of the 20th century (see, e.g., Ballantine, 1919, and Keedy, 1920, debating a proposal for limiting the insanity defense to the narrower *mens rea* conception, and the materials on early abolition attempts reviewed below in the context of the GBMI option). In 1954, Judge Bazelon of the District of Columbia Court of Appeals attempted to correct numerous deficiencies in the combined right-wrong/irresistible impulse test in *Durham v. United States* (1954), wherein the proposed standard was that an accused be considered not criminally responsible if his unlawful act was the product of mental disease or defect. In *United States v. Brawner* (1972), which ended the D.C. Court of Appeal's experiments with the Durham "product test" and adopted the Model Penal Code recommendations of the American Law Institute (1962), Judge Leventhal carefully reviewed the court's logic in adopting *Durham*. First,

[t]he old right-wrong/irresistible impulse rule for insanity was antiquated, no longer reflecting the community's judgment as to who ought to be held criminally liable for socially destructive acts. We considered the *Durham* rule as restated to have more

fruitful, accurate and considered reflection of the sensibilities of the community as revised and expanded in the light of continued study of abnormal human behavior. (p. 976)

Second, the older test forced expert witnesses to testify in uncomfortably narrow terms of “right/wrong,” making “it impossible to convey to the judge and jury the full range of information material to an assessment of defendant’s responsibility” (p. 976). Although it has been asserted (Goldstein, 1967; Livermore & Meehl, 1967) that the test need not be narrowly cognitive and could include a wider range of affective “knowledge” and “appreciation” if interpreted in the context of a proper jurisprudential and historical perspective, the concern of the *Durham* court was that this was not typical practice and, therefore, needed to be corrected.

Although the *Durham* product test was intended to remedy these problems, it was not perceived as having its intended effect and seemed instead to lead to the undue dominance of experts testifying in conclusory terms. As a result, the majority in *Brawner* adopted the ALI rule and further encouraged judges to adopt instructions that emphasized the importance of nonconclusory testimony and the expert’s role as explaining to the jury the relationship between the defendant’s cognitive, behavioral, and affective functioning and his or her “substantial capacity to appreciate the criminality of his conduct or to conform his conduct to the requirements of the law” (*United States v. Brawner*, 1972, p. 973 restating the ALI Model Penal Code).

Judge Bazelon, in his partial dissent, agreed that the product test needed to be rejected, but he was more pessimistic, viewing the majority’s adoption of the ALI rule as a change that was “primarily one of form rather than of substance” (*United States v. Brawner*, 1972, p. 1010). For Judge Bazelon, the purpose of the reformulation should have been to “ask the psychiatrist [or other mental health professional] a single question: What is the nature of the impairment of the defendant’s mental and emotional processes and behavioral controls?” leaving “for the jury the question of whether that impairment is sufficient to relieve the defendant of responsibility for the particular act charged” (p. 1032). To emphasize this, Judge Bazelon advocated a version of a test first proposed by the British Royal Commission on Capital Punishment in 1953: “A defendant is not responsible if at the time of his unlawful conduct his mental or emotional processes or behavior controls were impaired to such an extent that he cannot justly be held responsible for his act.” Judge Bazelon argued that this “justly responsible” test has the virtue of making overt the underlying moral nature of the insanity defense and placing the “hot potato” aspect of such judgments squarely into the hands of the jury, as representatives of the community. The ALI test in its full form has been adopted only by a few states—Hawaii, Kentucky, Massachusetts, Rhode Island, and Wisconsin—although a number of other states—Arkansas, Connecticut, Maryland, Michigan, Oregon, Vermont, West Virginia, and Wyoming—have adopted a modified version of the ALI standard (Packer, 2009).

Although the full or modified ALI rule has been widely adopted in federal jurisdictions and many states (Keilitz & Fulton, 1983), the movement to reform the insanity defense and limit its perceived abuse has led to an attempt to eliminate the “volitional prong” of the test (i.e., to conform one’s conduct to the requirements of law). Advocates of this alteration have included the American Bar Association (1983) and the American Psychiatric Association (1982); this proposed alteration was adopted into the Federal Code by the United States Congress in the Insanity Defense Reform Act (1984).

The Court of Appeals for the Fifth Circuit has agreed with this abolition of the volitional prong, arguing that the position of the American Psychiatric Association, that the profession did not possess sufficiently accurate scientific bases to measure a person’s capacity for self-control, was persuasive (*United States v. Lyons*, 1984a). In a strongly worded dissent (*United States v. Lyons*, 1984b), it was argued that the “potential threat to society (supposedly) created by the volitional prong” ignored “empirical data that . . . provide little or no support for these fearsome perceptions and in many respects refute them” (p. 995). The dissenting argument cited various studies undercutting the perceptions of the misuse of the insanity defense. The dissenters also could have included Rogers, Bloom, and Manson’s (1984) finding that personality-disordered defendants, the target of the advocates of abolishing the volitional prong, constituted only 18% of a group of successful insanity acquittees. Citing *United States v. Torniero* (1984), where the Second Circuit placed appropriate limits on “creative” uses of the volitional prong for new personality disorders by requiring the defense to show that “respected authorities in the field share the view that the disorder is a disease or defect that could have impaired the defendant’s ability to desist from the offense charged” (p. 730), the dissenters argued that the volitional prong was an essential aspect of the concept of guilt, because this concept “presuppose(s) a morally responsible agent to whom guilt can be attributed. By definition, guilt cannot be attributed to an individual unable to refrain from violating the law” (*United States v. Lyons*, 1984b, p. 1000). Delaware, Illinois, Indiana, and Maine have adopted the ALI cognitive prong only, and New York and the U.S. federal system utilize the cognitive prong only, but as a variant of both the M’Naghten and ALI standards (see Packer, 2009, Appendix A, for a full list of the insanity standards across all 50 states and the federal system).

THE GUILTY BUT MENTALLY ILL VERDICT

As of 2000, 20 states had provisions that allow for a defendant to be found GBMI (see Arrigo, 1996, for a review; see also Borum & Fulero, 1999, for a discussion of various proposed insanity defense reforms), and Utah uses the terminology guilty *and* mentally ill (Packer, 2009). The original GBMI legislation employed in the 20th century was introduced in Michigan in 1975 in the context of *People v. McQuillan* (1974), a case that found Michigan’s automatic commitment of NGRI acquittee’s unconstitutional. The verdict was also adopted in Indiana in 1979 under

similar circumstances. Following Hinckley's attempted assassination of President Reagan, the stage was set for other states to pass GBMI legislation in response to the perceived abuses of the insanity defense. The current GBMI defense was not intended to replace the NGRI defense (except in Utah and Nevada). Whereas both verdicts were introduced to stem the perceived tide of violence committed by offenders who escape "justice," the current form was aimed primarily at jurors, with the hope that it would provide them with a middle-ground decision between guilty and NGRI. It was also motivated by knowledge that an elimination of the insanity verdict itself might be considered unconstitutional, as it had been in *Strasburg* (1910) and *Underwood v. State* (1873).

In *Underwood*, Judge Campbell of the Michigan Supreme Court expressed his sympathies with the abolitionist argument. He acknowledged outrage in response to the "absurd lengths to which the defense of insanity has been allowed to go under the fanciful theories of incompetent and dogmatic witnesses," but he believed that the remedy was to be found elsewhere:

No doubt many criminals have escaped justice by the weight foolishly given by credulous jurors to evidence which their common sense should have disregarded. But the remedy is to be sought by correcting false notions, and not by destroying the safeguards of private liberty. (*State v. Strasburg*, 1910, p. 1028)

The GBMI verdict was intended to make it more difficult to reach a verdict of NGRI (especially in gray-area cases involving severely personality-disordered individuals), with the hope that most jurors would respond to the superficial logic of the verdict ("Okay, he's crazy, but he did it, didn't he?"). Opponents of the GBMI verdict have argued that it should be abolished on the grounds that it confuses and deceives jurors (Melville & Naimark, 2002; see also Palmer, 2000). To address the punitive and abolitionist motivation, defenders of the GBMI legislation added a gloss of rehabilitation by arguing that the new verdict provided an explicit means of recognizing that some of those sent to prison were in need of mental health treatment. Of course, they did not mention that few, if any, new funds were to be appropriated to the prison system to provide more treatment (Beasley, 1983) and that provisions already existed in every state that passed GBMI to laterally transfer a disturbed prisoner to a psychiatric facility for treatment if deemed necessary. In addition, even without the defense, the state is required by the United States Constitution to provide basic services to its inmates, and this legislation guaranteed nothing more. In commenting on this entire enterprise, Professor Richard Bonnie (1983) bluntly wrote, "The guilty but mentally ill verdict should be rejected as nothing more than moral sleight of hand" (p. 194). As Packer (2009) recently noted, none of the GBMI statutes incorporates conditions leading to lesser punishment, thereby fostering misperceptions regarding the meaning of the verdict. Jurors are not entitled to know the nature of the sentences associated with the various forms of mental state defenses, and individuals who are found GBMI are subject to the same penalties as those found guilty.

ASSESSMENT OF MENTAL STATE AT THE TIME OF THE OFFENSE

The evaluation process generally includes three major components or sources from which to elicit data: (1) an interview with the defendant; (2) traditional and/or forensic assessment instruments; and (3) third-party information, including but not limited to collateral reports, witness statements, victim statements, police reports, and records of various sorts (i.e., mental health, treatment, school, medical, crime scene). Each of these three major sources of data is reviewed next; however, the reader is referred to additional sources for more comprehensive and detailed information about the assessment process (see Melton, Petrila, Poythress, & Slobogin, 2007; Packer, 2009; Rogers & Shuman, 2000). In addition, we include a section on the role of delusions in evaluations of criminal responsibility, because the nature and quality of a defendant's thought process is often central in determining the extent of impairment in mental state at the time of the offense (MSO), particularly in contested cases.

INTERVIEW

A comprehensive MSO evaluation can be conceptualized as falling into a series of phases (per Sullivan's [1954] scheme):

1. Formal clinical-legal inception
2. Reconnaissance
3. Detailed inquiry of present mental state
4. Detailed inquiry of MSO
5. Reconciliation with other data sources (including consultation with other professionals that have evaluated the defendant)
6. Termination

Inception. In addition to rapport building, the inception relates to informed consent or notification, such that it requires clearly explaining one's role to the defendant, focusing on why he or she is being evaluated, who will have access to the information gathered, and what limits are placed on the confidentiality of information. These confidentiality rules vary widely across jurisdictions and are strongly influenced by the context of the case; therefore, the examiner must be fully informed in this regard, as a matter of professional competence. In most jurisdictions, once defendants have entered their mental state into the adjudication process by interposing an insanity defense or some other mental state claim, no information revealed to the examiner that can be construed as relevant to that claim is protected. Jurisdictions differ widely, however, as to whether indirect fruits of such evidence are admissible; hence, extreme caution is required when preparing a report. The broadest coverage is found in the federal courts:

No statement made by a defendant in the course of any [forensic] examination . . . with or without the defendant's consent, no testimony by the expert based on the statement,

and no other fruits of the statement may be admitted into evidence against the defendant in any criminal proceeding except on an issue regarding mental condition on which the defendant has introduced [evidence]. (Federal Rules of Criminal Procedure, Rule 12.2 (c) 4, 2011)

As noted earlier, however, jurisdictions vary widely, and the examiner must conform his or her practice to the local rules. It is also good practice to allow the defendant to provide his or her version of the events surrounding the time of the offense (the “filtered version”) before the examiner introduces any contradictory evidence or challenges the defendant on any noted inconsistencies. Near the end of the interview, the examiner may choose to introduce contradictory evidence in an effort to observe the defendant’s reaction, evaluate the defendant’s response style, and determine whether he or she is consciously distorting facts or experiencing bona fide memory-related difficulties.

Reconnaissance. This is a forensically oriented review of the defendant’s history. It is important to obtain information regarding the defendant’s history of psychiatric disturbance, treatments received, and general experience of his or her psychiatric condition. Of particular importance are prior episodes that may have involved criminal charges and/or competency evaluations, civil commitments, or other such dispositions. The pattern of mental state disturbance, its relationship to psychotherapeutic and psychopharmacological treatment, medical conditions (e.g., hypoglycemia, hyperthyroidism), situational stressors, and substance use are particularly important.

Detailed Inquiries—Present Mental State and Mental State at Offense. Typically, it is difficult to separate these concepts, because a disturbed defendant usually will be treated with rapidly acting psychotropic medications. Nevertheless, it is crucial to bear in mind that these mental states, although related, are separable, albeit with great difficulty. Forensic evaluators need to be concerned with “Riggins factors,” (*Riggins v. Nevada*, 1992), that is, alternations in the individual’s mental state as a function of psychotropic medication. Within the context of MSO evaluations, defendants are likely to have been treated with psychotropic medications for a long time period following the alleged incident but prior to the evaluation. This can be problematic as the defendant’s ability to recall and describe his or her prior mental state accurately is potentially impaired.

We advise the use of sections of structured and semistructured interviews to cover the domain of psychopathology in a relatively standardized fashion to improve interexaminer reliability in the elicitation and coding of information. The detailed inquiry with respect to the defendant’s MSO must also focus on the relationship of the psychopathological elements to the criminal conduct charged. Doing this often involves comparing the defendant’s reports of his or her mental state and behaviors with various aspects of crime scene evidence. This part of the interview resembles

a psychological autopsy. The defendant must be asked to reconstruct his or her thoughts, perceptions, experiences, attitudes, and behavior, as well as those in the field of action during the entire legally significant period. Retrospective evaluations present challenges for laypersons, jurors, judges, and examiners alike; therefore, great care must be taken to obtain detailed information and also to avoid, as much as possible, recall-based contamination of the defendant's memories.

Reconciliation and Termination. As emphasized by many advocates and critics of the role of the forensic examiner in the legal process (see Bonnie & Slobogin, 1980; Ferguson & Ogloff, 2011; Melton et al., 2007), the role of the expert is not to present legal conclusions or formal psychopathological diagnoses. Rather, the role of examiner, as expert, is to import state-of-the-art scientific knowledge about the existence of various psychiatric conditions and their relationship to behavioral, perceptual, cognitive, and judgmental capacities into the legal/moral decisional process. Thus, at the reconciliation or termination phase, the examiner should be prepared to integrate the information available at this level and to inform all parties concerned, including, as appropriate, the defendant, defense counsel, prosecutor, and other mental health professionals. One advantage of this openness is that it allows the defendant to produce any additional information that might explain or clarify discrepancies, and it also helps prevent an uninformed battle of the experts. In certain gray-area cases, there may be legitimate disagreements among experts. It assists the trier of fact if the nature of these disagreements, as well as areas of agreement, are drawn as precisely as possible, with each examiner fully aware and able to comment in advance as to the reasons for disagreement. Such pretestimony consultations also tend to produce higher-quality and more informative strategies for direct and cross-examination.

FORENSIC ASSESSMENT INSTRUMENTS

Two specialized forensic assessment instruments have been developed to assist in the evaluation of MSO. Although both of these instruments were developed in the 1980s, standards of forensic practice continue to move in the direction of incorporating forensic assessment instruments in the evaluation of psycholegal issues.

Mental State at the Time of Offense Screening Evaluation. The Mental State at the Time of the Offense Screening Evaluation (MSE; Slobogin, Melton, & Showalter, 1984) is a semistructured interview technique developed to screen out defendants for whom an insanity defense is clearly not applicable. In addition, the MSE can also be used to identify those individuals who are "obviously insane" and therefore do not require a more comprehensive evaluation. The MSE is composed of three sections: (1) historical information, which assesses a defendant's premorbid psychological and cognitive functioning; (2) offense information, which accumulates information

regarding the offense from the defendant and external sources; and (3) present mental status examination.

Although there have not been any published studies of the reliability of the MSE, its validity was evaluated by Slobogin and colleagues (1984). Twenty-four mental health professionals were trained to use the MSE and were then asked to assess 36 cases. They were given only a description of the charge and the preliminary hearing transcript prior to their assessment. Their decisions were then compared to the decisions of an inpatient forensic evaluation team, which included one psychiatrist, one psychologist, and one social worker. Overall, there was satisfactory agreement (72%, or 26 of 36 cases) between trainees and the evaluation team. There was 44% agreement (16 of 36 cases) on the cases that were screened out. Using the decisions made by the evaluation team as the criterion, the decisions made by the trainees were found to have a 0% false-negative rate and a 28% (10 cases) false-positive rate ("screened in" defendants who were screened out by the evaluation team). Compared to the evaluation team's decisions, the trainees' decisions had less agreement with the court's verdict. Of the 10 defendants for whom the evaluation team suspected some "significant mental abnormality," 1 was convicted, 7 had their charges *nolle prossed* (prosecution decided not to pursue the case), and 2 were found insane. Of the 20 defendants whom the trainees suspected had some "significant mental abnormality," 6 were convicted as charged, 4 were convicted of a lesser charge, 6 had charges *nolle prossed*, and 2 were found insane.

The limitations of the MSE have been addressed in the literature (see Poythress, Melton, Pettila, & Slobogin, 2000; Rogers & Shuman, 2000). Given the lack of research on its reliability and the limited validity data available, the MSE should be viewed as a guide for evaluators to ensure that relevant areas of inquiry are addressed. Indeed, evaluators can include the MSE (or the R-CRAS, discussed next) in a comprehensive evaluation that would include multiple sources of data (e.g., psychological tests, third-party information, defendant's interview).

Rogers Criminal Responsibility Assessment Scales. The Rogers Criminal Responsibility Assessment Scales (R-CRAS; Rogers, 1984) were designed to quantify the elements of the ALI criteria for criminal nonresponsibility; however, Rogers noted that it might be applicable to the M'Naughten standard as well (Rogers & Shuman, 2000). Based on a comprehensive evaluation, the examiner rates the defendant with respect to a series of scales grouped into five areas: (1) reliability of report, (2) organicity, (3) psychopathology, (4) cognitive control, and (5) behavioral control. For example, the psychopathology section involves ratings of bizarre behavior, anxiety, amnesia, delusions, hallucinations, depressed or elevated mood, verbal coherence, and affective and thought disorder. In addition, there is a series of more global ratings on final judgments of insanity and impairment. Thus, the R-CRAS is an instrument that reflects the relative importance assigned by examiners to the first-order elements of an insanity decision (e.g., the presence and relevance of psychopathology to MSO). It should be noted that these elements are fairly abstract

psychological and legal terms (e.g., “delusions at the time of alleged crime”) and do not necessarily represent the cues that are actually utilized by professional examiners in making their decisions. This was a major issue of contention among Rogers, Melton, Petrila, Poythress, & Slogobin (1997), and Golding (1992) in the evaluation of this instrument. Rogers believes it is important to quantify the issue; however, it would be unfair to assume that he did not recognize the value of qualitative data (see Rogers & Ewing, 1992). Golding and Melton et al. agreed that quantification is essentially illusory at this stage in the development of evaluations of criminal responsibility. Both groups of authors agree, in large measure, on the domains or conceptual elements to be addressed.

Rogers and Sewell (1999) responded to the criticisms of Melton et al. (1997) by attempting to extend the R-CRAS’s construct validity via the reanalysis of two data sets to address contributions of individual variables to the various components of the decision model. On the basis of discriminant function analysis, the authors concluded that the R-CRAS variables were able to form differentiating patterns (between individuals showing impairment and those not showing impairment) for each of the five components of the decision model. Results indicated average hit rates of 94.3% (ranging from 87.8% for major mental disorder to 97.2% for cognitive control), and the average variance accounted for was 63.7% (ranging from 38.5% for malingering to 79.2% for behavioral control).

Factor analysis of the R-CRAS items has resulted in three factors—bizarre behavior, high activity, and high anxiety—that do not mirror the five scales (see Borum, 2003). Rogers has reported modest interrater reliabilities at the item level (average Kappa 0.58), with lower values (0.49) associated with the product question (i.e., was the loss of control attributable to underlying psychopathological disturbance?), which is one of the most frequent sources of disagreement in contested trials (Rogers & Shuman, 2000; Rogers, Wasyliv, & Cavanaugh, 1984). Final judgments with the R-CRAS have also been associated with reasonable levels of agreement between examiners and triers of fact (96% with respect to sanity with lower levels of agreement on insanity [70%]; Rogers, Cavanaugh, Seman, & Harris, 1984; see Rogers & Shuman, 2000, for a summary). These findings are in general accord with the levels of agreement found between clinicians and courts in other studies of final judgment that use no formalized interviews or rating scales (Golding, 1992). Unfortunately, all studies in this area appear to use criterion-contaminated groups, because the examination process was part of the judicial (criterial) determination.

A number of proposals for semistructured protocols to assist in the evaluation of mental state at time of the offense have been made (see, e.g., Golding & Roesch, 1987; Melton et al., 2007; Ogloff, Roberts, & Roesch, 1993). They share in common an open-ended structure, with special attention to developing multisource data, identification of legally and psychologically relevant dimensions of the criminal responsibility evaluation, and disclosure of the logical links in an evaluator’s reasoning. The need to develop such a comprehensive analysis is clear. These less structured approaches have been shown to be empirically useful in various

contexts (see Melton et al., 2007, for a review), but they have not been studied in the same fashion as the R-CRAS. The R-CRAS and less structured MSO evaluation techniques make their most important contributions by clarifying the underlying bases for professional judgment for the trier of fact (Golding, 1990, 1992), thus potentially highlighting the areas of disagreement so that expert testimony can be of more assistance to judge or juror. There is much reason to believe that forensic examiners reach their generally high level of agreement in “ultimate opinions” by different logical and empirically sustainable routes. The critical issue remains the association between organic or psychopathological disturbance and control/moral judgment capacities; these devices are most useful when they serve the heuristic value of (a) highlighting the aspects of the defendant’s psychological state that are relevant, (b) describing a purported relationship to control and judgment capacities, and (c) organizing known data about the empirical relationships between disorder and psychological capacities in various states and situations.

THIRD-PARTY INFORMATION

It is a commonly accepted standard of practice that forensic evaluators examine the consistency of mental health history and other archival data along with details of the crime scene and witnesses’ accounts of the defendant before, during, and after the alleged incident. This consideration of “consistency” is relevant to issues of malingering and aids in supporting or challenging various psychological interpretations of the defendant’s MSO. The importance of this aspect of a forensic examination at the time of the offense has been addressed in numerous sources (see Melton et al., 2007; Packer, 2009; Rogers & Shuman, 2000). Forensic evaluators need to pay close attention to crime scene data as well as to more traditional sources of third-party information (e.g., mental health and other records, witness statements). All such sources need to be integrated in as straightforward a manner as possible. Where limited scientific data exist to support the inference (e.g., descriptive studies of the characteristics of hallucinations or delusions), they should be referenced, and where otherwise indicated, the evaluator’s logical link analysis should be presented and scrutinized.

As Melton et al. (2007) noted, the role of the forensic evaluator is not to conclusively resolve all conflicting accounts about the case but rather to conduct an evaluation that can be scrutinized in terms of all available evidence, both psychological and behavioral. Modern forensic standards of practice are to address the issue of the consistency of behavioral crime scene evidence with psycholegal formulations and allow the trier of fact to make the determination of their significance, guided by whatever scientific evidence can be directly or indirectly adduced.

ROLE OF DELUSIONS IN ASSESSMENT OF CRIMINAL RESPONSIBILITY

The nature and quality of a defendant’s delusions are central in determining the extent of impairment impacting his or her MSO. Several issues are involved.

In *contested* cases, forensic examiners are particularly apt to encounter defendants with extreme or idiosyncratic beliefs about religion, politics, or personal identity, and questions regarding the delusional basis for those beliefs will arise. Delusionality also enters the adjudicatory process when the issues of intentionality, compulsion, or the reasonableness of the defendant's conduct may be related to a delusion. Finally, delusionality is an important aspect of risk assessment with respect to release decisions.

Although there are surprisingly few studies on the frequency and nature of delusions of NGRI defendants, indirect data, as well as experience, suggest that delusionality is a vital issue (see Litwack, 2003, for a discussion of defendants who refuse to mount an insanity defense on the basis of delusional reasoning). Delusions (especially delusions of reference, persecution, and control, which are particularly relevant to forensic cases) are prevalent among individuals with psychotic disorders (Winters & Neale, 1983). Furthermore, roughly half of defendants who raise the insanity defense and 70% of insanity acquittees have psychotic diagnoses (Steadman & Braff, 1983). Moreover, delusions have been shown to be specifically and substantially related to violence (Taylor et al., 1994).

In an extensive analysis of case records, Häfner and Böker (1982) found that 70% of individuals who were diagnosed with schizophrenia and accused of homicide harbored delusional beliefs about their relationship with their victim. Similarly, in an interview-based study, Taylor (1985) found that 40% of psychotic defendants reported acting *directly* on their delusions during their offenses. The frequency of *nonpathological* but radical religious and/or political beliefs, and the extent of the relationship between these beliefs and violence, have not been systematically studied (Taylor et al., 1994); however, defendants whose criminal acts are related to such fervently held beliefs are clearly plausible candidates for MSO evaluations wherein a primary focus will be the determination of whether these extreme beliefs are delusional.

Distinguishing between radical beliefs and delusions is a difficult, but critical, task in the assessment of criminal responsibility. Respecting the principles of autonomy and self-determination, our legal system holds responsible the extremist who chooses to act on a radical system of beliefs, expressing his or her desires, values, and "personhood" through the crime (see Hermann, 1990). The basic moral logic of the insanity defense, however, excuses the mentally disordered individual who acts on a pathological, uncontrollable belief system that distorts his or her sense of reality, thereby impairing the capacity for rational choice. Arguably, the same logic also applies to delusions that would "justify" the actions.

Notwithstanding the centrality of this issue in assessing criminal responsibility, there are few empirical studies or practice standards available to aid in assessing the delusionality of beliefs. In gray-area cases, or cases in which defendants are not clearly disordered, the classification of beliefs as delusional is presumably a major source of disagreement among examiners. There is no bright line of demarcation between extreme beliefs and delusions (Garety & Hemsley, 1994; Oltmanns, 1988).

Moreover, religious and political belief systems that reference nonphysical entities and events are not scientifically testable; consequently, there is “no full standard of truth independent of what the [defendant] says” (Taylor et al., 1994, p. 167). For these reasons, in this section we review recent progress in defining and assessing delusions and their likely consequences. The purpose of this section is to introduce readers to the complex issues involved in assessing delusions and in conducting informed assessments of defendants’ beliefs in the context of insanity evaluations. Emphasis is placed on issues relevant to distinguishing between extreme religious or political beliefs and delusions.

Defining and Conceptualizing Delusions. In the text revision of the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR;* American Psychiatric Association, 2000), a *delusion* is defined as:

A false belief based on incorrect inference about external reality that is firmly sustained despite what almost everyone else believes and despite what constitutes incontrovertible and obvious proof or evidence to the contrary. The belief is not one ordinarily accepted by other members of the person’s culture or subculture (e.g., it is not an article of religious faith). When a false belief involves a value judgment, it is regarded as a delusion only when the judgment is so extreme as to defy credibility. (p. 821, emphasis added)

Although this definition is adequate for most clinical and clinical-forensic purposes, its shortcomings are readily apparent when one attempts to delineate its boundaries (Sedler, 1995). As noted, there is often no standard of proof by which one can assess the falsity of beliefs in many religious, political, or identity systems. Similarly, it is difficult to evaluate the incredibility or implausibility of beliefs, because clinicians may not agree on the extent to which certain beliefs are bizarre (Flaum, Arndt, & Andreasen, 1991; Oltmanns, 1988; Spitzer, First, Kendler, & Stein, 1993; cf. Mojtabai & Nicholson, 1995). The degree of conviction with which a belief is held also does not clearly distinguish between those that are delusional and nondelusional. Like delusional beliefs, nondelusional but highly valued beliefs are often held with great zeal and intensity, even in the face of contradictory evidence. Moreover, many persons shift between periods in which they experience certainty about their delusions and periods in which they demonstrate partial or full insight (see Harrow, Rattenbury, & Stoll, 1988; Sacks, Carpenter, & Strauss, 1974).

These difficulties are “frequently compounded by ambiguity surrounding the presence or absence of cultural support for the person’s belief” (Oltmanns, 1988, p. 3). It is difficult to determine the extent to which the nature of a belief, its experience, or its expression must deviate from that accepted by a designated subgroup to be classified as delusional. Clearly, examiners must have considerable knowledge of the social, religious, political, and even scientific context of a defendant’s belief to adequately assess its delusionality (see Barnhouse, 1986; Oltmanns, 1988).

Failure to carefully consider patients' cultural and religious backgrounds often results in misdiagnoses (see Lu, Lukoff, & Turner, 1997).

Most current attempts to systematically analyze delusions are based on the seminal work of Jaspers (1963; see Garety & Hemsley, 1994; Mullen, 1985; Sedler, 1995), who provided the most comprehensive, enduring, and clinically useful theory for distinguishing among various categories of delusional and nondelusional beliefs. A simplified summary of this theory focused on differentiating delusional from nondelusional beliefs is presented here (see C. Walker, 1991, for greater detail).

Jaspers (1963) argued that the criteria of level of conviction, imperviousness to counterargument(s), and impossibility or bizarreness were insufficient external criteria that did not capture the essence of delusionality. Rather, "overvalued beliefs," or even "delusion-like ideas," could be distinguished from primary delusions, based on Jaspers's approach, by attention to three fundamental criteria:

1. Primary delusions are distinguished from secondary delusions and beliefs that are merely overvalued based on the extent and nature of their "un-understandability."
2. Primary delusions are "unmediated" by thought, analysis, deduction, or reflection, whereas overvalued ideas and secondary delusions reflect varying degrees of cognitive appraisal and inference.
3. Primary delusions reflect a distinctive change in an individual's personality functioning; that is, they are a distinct change in the totality of the individual's personal meanings and ways of construing the world.

An overvalued idea is an understandable product of cognitive interpretation and can be viewed in terms of an individual's personality, life experiences, and sociocultural background. Overvalued ideas can be understood "as exaggerations, diminutions, or combinations of phenomena which we ourselves experience" (Jaspers, 1963, quoted in C. Walker, 1991, p. 100). In contrast, a secondary delusion ("crazy idea") is understandable only in the sense that it emerges through one's process of reasoning about psychopathological experiences (e.g., based on the quiet voices and buzzing an individual occasionally hears, she arrives at the conclusion that she is a target of government surveillance). A primary delusion is *not* understandable because it originates in a direct, immediate experience of new meaning unmediated by thought and unconnected to the person's fundamental personality (e.g., one sees a "man in a brown coat... he is the dead Archduke," C. Walker, 1991, p. 99). Thus, whereas overvalued ideas have "clear precedent" in an individual's existing personality and meaningful life events, secondary delusions emerge from other psychopathological experiences, and primary delusions fundamentally change an individual's personality or "way of looking at the world" (C. Walker, 1991, p. 101). Although Jaspers's classification has not been subject to much empirical research and can be criticized for relying too heavily on the subjective criterion of "understandability" (Mullen, 1985, p. 335), his theory provides useful guidance in conceptualizing the key distinctions between delusions and overvalued ideas.

ISSUES IN THE TREATMENT AND RELEASE OF INSANITY ACQUITTEES

Inherent in the decision to find a defendant not guilty by reason of insanity is concern about the ultimate “disposition” of such acquittees. Jurors place a great deal of emphasis on this issue in their deliberations (Golding, 1992). A comparison of data on the rates and success of insanity pleas and the nature of insanity dispositions with data on public perceptions about these issues clearly reveals that the public overestimates the frequency and successfulness of the insanity plea and underestimates the nature and length of institutionalization following an insanity “acquittal” (Silver, Cirincione, & Steadman, 1994). In *Shannon v. United States* (1994), the Supreme Court held that NGRI defendants have no right to a jury instruction that explains the post-“acquittal” commitment process, because such an instruction would violate the long-standing principle that a jury must base its verdict on the evidence with which it was presented. A recent case in Maine returned a comparable holding. In *State v. Okie* (2010), the Supreme Judicial Court of Maine ruled that jury instructions needed to be accurate with respect to the law but that it was not appropriate to inform the jury of the consequences of an insanity verdict. The court’s holding was based on legal/historical precedent, which is in line with federal law, and based on the fact that judges, not juries, impose sentences (see Lloyd & Packer, 2011, for more detailed case discussion). Although this may be correct jurisprudential theory, it violates commonsense justice, in that we have strong reason to believe that jurors do pay attention to this issue and that these assumptions are incorrect (see the next section on empirical developments). In a similar regard, in *Foster v. State* (2008), the Georgia Supreme Court unanimously reversed a murder conviction because the trial court did not define “mentally ill” and “mentally retarded” for the jury. The court held that, by not being provided with these definitions, the jury was unable to make an adequate assessment of the conditions and therefore could not properly consider potential verdicts.

Contrary to strong public concerns, NGRI acquittees are not “easily” released. In fact, research indicates that they are more likely to remain institutionalized for longer periods of time than crime-equivalent “guilty” persons (see Miller, 1994; Silver, 1995). The constitutionality of this likelihood has been justified by the U.S. Supreme Court on the grounds that the purpose of their commitment is to treat their dangerousness, not to punish them for a crime for which they were not culpable (see *Foucha v. Louisiana*, 1992; *Jones v. United States*, 1983). A number of factors contribute to this typically lengthy treatment and slow release process.

First, NGRI acquittees are likely to be severely mentally disordered persons for whom currently available treatment, both biochemical and psychosocial, has been ineffective. This is not because they are necessarily bad people or nonadherent to treatment regimens, but because, contrary to medical and pharmaceutical company myth, about one-third of severely mentally disordered individuals fail to show meaningful clinical responses to such treatment (see Relman & Angell, 2002). Such individuals may also become nonadherent as a result of a complex psychological process. That is, unless patients perceive a subjective benefit from treatment, they

are less likely to tolerate unpleasant side effects. Mental health professionals who establish an authoritarian or paternalistic (“You’re sick and you need to take your medicine”) as opposed to a collaborative approach (“Let’s work together to find a treatment strategy where the benefits outweigh the risks and side effects”) are also more likely to contribute to nonadherence (see Appelbaum & Gutheil, 1982). Where such problems in the therapeutic relationship exist, medications are frequently delusionally reinterpreted as the cause for psychotic symptoms. Thus, the typical NGRI admission has a history of unsuccessful treatment, the most recent treatment frequently occurring only months before the index offense (Golding, Eaves, & Kowaz, 1989).

Second, because one’s “dangerousness” (manifested by the *actus reus*) is linked, by virtue of the plea, to his or her mental disorder, the person is unlikely to be released until he or she is deemed psychiatrically stable. Furthermore, as illustrated in *Jones* and *Foucha*, the person can be committed and held for what amounts to an indeterminant length of time (i.e., until he or she is no longer deemed “dangerous”). Stredny, Parker, and Dibble (2012) examined the characteristics of recently hospitalized insanity acquittees in Virginia and found that individuals diagnosed with mood and anxiety disorders were less likely to receive recommendations for inpatient commitment than those diagnosed with psychotic disorders. They also found psychologists and psychiatrists to generally agree on recommendations for the ultimate release of insanity acquittees, with inpatient recommendations more commonly occurring for those with histories of self-injury and/or suicide attempts, family or other psychosocial issues, and a lack of structured activities in the community. Psychologists also tended to recommend commitment when a weapon was used during the commission of the crime, and both groups were more likely to recommend commitment when a person was transferred to the hospital from jail as opposed to from the community (on bond). In another study, Manguno-Mire, Thompson, Bertman-Pate, Burnett, and Thompson (2007) reviewed the records of 91 insanity acquittees in a maximum security forensic hospital in Louisiana and found that the Psychopathy Checklist–Revised (PCL-R) scores and age at which the subject committed the first criminal offense were related to release recommendations. Specifically, those with higher levels of psychopathy and those who engaged in criminal behavior at a younger age were less likely to be recommended for release.

Third, political realities based on the sensational publicity produced by the media, in large part because they perpetuate insanity defense myths in failed NGRI releases (see Silver et al., 1994), make decision makers quite cautious. Finally, few jurisdictions have an articulated prerelease risk assessment, postrelease risk management, and intensive case management system capable of more safely handling the community adjustment and supervision needed for such individuals.

An unstudied aspect of the dispositional issue concerns the ultimate costs and effectiveness of placing prototypical insanity acquittees in forensic treatment contexts versus placement in traditional correctional facilities. Although some data clearly support the monitored release of NGRI acquittees, the trend, in those states

with either GBMI or highly restricted (or nonexistent) insanity defenses, is to place such defendants in correctional environments for the majority of their “sentence” or institutionalization. Wiederanders (1992; Wiederanders & Choate, 1994) and Golding and colleagues (1989) demonstrated that articulated follow-up of insanity acquittees in the community is feasible and worthy of study. What we do not have is informative data on the differences between similar individuals “treated” in correctional versus forensic mental health contexts. We know of no empirical studies of this issue. With respect to mental health economics, the question is whether society eventually pays more or less for treating mentally disordered offenders in prison versus forensic mental health systems. Clearly, the cost per diem while initially incarcerated will favor prison over mental health system dispositions; however, the analysis also needs to include days institutionalized and days in the community at lower cost as well as the likelihood and financial, emotional, and moral costs of recidivism. Again, we know of no direct data, but, on logical grounds, we propose that treated and supervised mentally ill and dangerous offenders ultimately would cost less, financially and emotionally, than incarcerated mentally ill offenders who receive less mental health treatment and supervision.

As is evident from this discussion, clinicians face a multitude of issues with respect to the assessment and disposition of NGRI defendants. We now turn to a review of empirical research developments regarding many of these issues.

EMPIRICAL DEVELOPMENTS REGARDING CRIMINAL RESPONSIBILITY

Research in the area of criminal responsibility has taken a number of forms. We have conceptualized this research as falling into three broad areas:

1. Research on NGRI verdicts, including the frequency of NGRI verdicts, rates of agreement between experts as well as between experts and the courts with respect to NGRI opinions, and the characteristics of NGRI acquittees
2. Research on judicial instruction, including an examination of various legal standards or tests of insanity
3. Jury decision making with respect to insanity, including jurors’ case-relevant attitudes, case construals, and implicit theories or prototypes regarding insanity

Next we present a brief overview of each of these three areas. The reader is referred to other sources for a more detailed and comprehensive review of this literature (see especially Finkel, 1995, 2000; Lymburner & Roesch, 1999; Simon, 1999).

RESEARCH ON NGRI VERDICTS

Research on the empirical realities of the adjudication of criminal responsibility has continued to demonstrate the same basic phenomena since earlier reviews (see Golding, 1992; Pasewark, 1986). The research has focused primarily on describing

the NGRI population in traditional demographic and diagnostic terms and at demythologizing public misconceptions of the insanity defense (see Silver et al., 1994, for a review).

Frequency of NGRI Verdicts. Research indicates that the NGRI defense is seldom raised, averaging less than 1% of total felony indictments (Steadman et al., 1993), and is highly variable in its success rate, with a modal rate of 25% for those who raise the issue. Cirincione, Steadman, and McGreevy (1995) surveyed seven states with respect to the frequency of insanity pleas and the likelihood of insanity acquittals and found an inverse relationship between the two factors. That is, states wherein insanity plea rates were high had a lower insanity acquittal rate. These authors found an average rate of insanity pleas of .85 (less than 1%) per 100 felony indictments and an aggregated success rate of 28%. Silver and colleagues (1994) cited highly variable success rates, ranging from 7% to 87% and averaging 26% (for earlier studies and summaries of success rates, see Janofsky, Vandewalle, & Rappaport, 1989; Pasewark, 1986; Steadman & Braff, 1983).

Cirincione and Jacobs (1999) attempted to collect data on each state with respect to the annual number of insanity acquittals and, despite valiant efforts, were able to obtain data from only 36 states. For these 36 states, the mean number of insanity acquittals was 33.4 per state per year, with a median of 17.7 and a standard deviation of 41.7. California (134.0) and Florida (110.5) had the highest average number of insanity acquittals per year, and New Mexico (0.0) and South Dakota (0.1) had the lowest, with six states reporting no more than one insanity acquittal per year.

Rates of Agreement. There is a high rate of agreement among forensic experts with similar levels of training, experience, and methodology and high levels of agreement between examiner opinions and judge or juror decisions with respect to insanity opinions. Research on the reliability of forensic judgments indicates that the types of cases likely to be contested include issues of an Axis I diagnosis with a comorbid personality disorder diagnosis; highly idiosyncratic and paranoid religious, political, or identity systems; intoxication or failure to take medications; and extremely bizarre conduct (for summaries of the reliability research, see Gowensmith, Murrie, & Boccaccini, 2012; Melton et al., 2007; Rogers & Ewing, 1992). NGRI verdicts typically are achieved as either a stipulation between defense and prosecution or bench trials, and they rarely involve contested battles of experts in front of jurors (Golding, 1992; Melton et al., 2007; Silver et al., 1994).

Research on interexaminer agreement is of limited utility, because it focuses on global agreement. Research on the logic and structure of examiner decision making in competency to stand trial evaluations has demonstrated high levels of global agreement in ultimate conclusions but poor agreement in examiner logic, particularly as it pertains to defendants' specific abilities and incapacities (Skeem, Golding, Cohn, & Berge, 1998). In theory, we would expect the same

in NGRI evaluations. Similarly, we know of no empirical study of the crime characteristics that lead actual judicial decision makers or forensic examiners to conclude that the defendant lacked the legally or morally relevant MSO. This point is particularly crucial, because few examiners, in our experience, pay close attention to collateral reports or crime scene data with respect to its consistency with their inferences about a defendant's mental state (see Melton et al., 2007).

Unfortunately, this type of sociodemographic research does not address the more theoretically interesting question of what types of mental disorder characteristics (beyond psychosis) and what aspects of offense incident characteristics (e.g., planning, intentionality behaviors, reasonableness of motive) influence expert, judge, and juror decision making.

Characteristics of NGRI Defendants. Finkel (1995) and Roberts and Golding (1991) have argued and presented rather convincing data from analog studies that jurors' individual construal of the case and particular defendant characteristics along a set of dimensions are major determinants of mock jurors' decision making. The dimensions along which mock jurors appear to individually construe the defendant include:

- the ability to think and reason rationally and clearly;
- the capacity to perceive and be aware without distortion;
- the capacity to choose courses of action;
- rational motivation for actions; the ability to control thoughts, feelings, and behaviors; and
- responsibility for altering one's mental state by intoxication, noncompliance with medication, or other factors.

It would be important for large-scale research on both examiner judgments and the verdicts reached by judges and the rare trial jury to examine what elements or factors they rely on in reaching their decisions.

Some researchers have focused on the characteristics of defendants who earned NGRI acquittals. These defendants typically have severe mental illnesses, namely psychotic disorders, and extensive mental health histories, often with prior civil commitments or adjudications as incompetent to proceed (see Golding, 1992; Golding et al., 1989; Ogloff, Schweighofer, Turnbull, & Whittemore, 1992; Steadman et al., 1993). In a detailed examination of a large NGRI cohort, Golding and his colleagues (1989) found that 79% had been previously hospitalized (mean = 4.11 times), and 43% of these admissions were for forensic reasons. Over half of the subjects with prior admissions were discharged within 1 year of their index offense, and 45% committed their index offense within 6 months of their last discharge. Miraglia and Hall (2011) examined the characteristics of all NGRI patients committed to a secure facility in New York State between 1980 and 2007 ($n = 440$; 386 cases were included in the analyses). Approximately 2% of females and 14% of males were

rearrested within 2 years of release. The authors went on to state that risk of rearrest was generally limited to the first few years postrelease, such that approximately one-half of the rearrests occurred within the first 2 years of release and nearly two-thirds occurred by year 5. The probability of being rearrested approached zero for those who were not rearrested by their 10th year in the community.

Callahan and Silver (1998a) studied the factors associated with the conditional release of NGRI acquittees in four states (i.e., Connecticut, Maryland, Ohio, New York) and found great variance. In Connecticut, few individuals were likely to be released conditionally, regardless of their characteristics. In New York, however, demographic characteristics were most predictive of conditional release: Females, Whites, and high school graduates were most likely to be conditionally released. In Maryland, clinical prognosis was the most critical variable, with those defendants with schizophrenia being *less* likely than those with other major mental illnesses to be granted conditional release. In Ohio, the nature of the crime was the most significant predictor, with serious offenders less likely to be released. It is somewhat concerning that in New York, the variables most predictive of conditional release were related neither to crime characteristics nor psychiatric variables. In a recent study, Dirks-Linhorst and Kondrat (2012) analyzed 27 years of NGRI acquittals in Missouri. Those who committed homicide were less likely to have obtained conditional release as compared to NGRI acquittees who committed other offenses, which the study authors believed was likely attributable to one or more of these issues: Their symptoms may take longer to manage; the gravity of their offenses results in more caution in release decisions, thereby requiring murder acquittees to demonstrate longer periods of behavioral and psychiatric stability; and/or that admission to a more secure facility likely results in greater lengths of stay.

Revocation rates of conditional release vary widely (e.g., between 35% and 50%, see Callahan & Silver, 1998b; Heilbrun & Griffin, 1993; Wiederanders, Bromley, & Choate, 1997). Variables such as being White, employed, and married are indicative of successful conditional release (see Tellefsen, Cohen, Silver, & Dougherty, 1992). Monson, Gunnin, Fogel, and Kyle (2001) examined the factors related to the revocation of conditional release in a sample of 125 NGRI acquittees and found that minority status, criminal history, and a substance use diagnosis were significantly predictive of revocation of conditional release. Vitacco et al. (2008) investigated factors related to failure to maintain conditional release across all NGRI acquittees in Wisconsin over a 5-year period ($n = 363$). These researchers found that diagnosis of substance abuse, previous revocation of conditional release, and the experience of needing psychiatric hospitalization were all related to revocation of conditional release. In a follow-up study that focused on female insanity acquittees ($n = 76$) conditionally released to the community over a 7-year period in Wisconsin, Vitacco et al. (2011) found that 52 of the females (68.4%) maintained their conditional release, whereas 24 (31.6%) had it revoked due to rule violations or criminal behavior but not violent behavior (despite the fact that approximately 54% of the initial sample was found NGRI subsequent to the commission of a violent offense). Of the 24 females

whose release was revoked, 6 were released again, and 5 of the 6 engaged in behaviors leading to a second revocation, suggesting that prior revocation is a risk factor for future revocation.

RESEARCH ON JUDICIAL INSTRUCTION

Jurors are expected to determine an appropriate verdict by conscientiously applying the law to a fair evaluation of the evidence (*Wainwright v. Witt*, 1985). As suggested, the effects of language differences in legal standards for insanity have been intensely debated for over two centuries. The nature and outcome of these debates, however, have shown either weak or little practical effect on jurors. Research repeatedly demonstrates that mock jurors often do *not* apply judicial instruction on various legal definitions of insanity in rendering verdicts (Finkel, 1995; Simon, 1999). For example, the Insanity Defense Reform Act (IDRA; 1984) was formulated after John Hinckley's acquittal to narrow the language of the ALI standard (1962), with the intention of curbing the number of insanity verdicts (by eliminating the "volitional prong"). In an analog study, Finkel (1989) found *no* verdict differences among mock jurors who were provided IDRA instructions, ALI instructions, or very narrow "wild beast" instructions (*Arnold's Case*, 1724; cited in N. Walker, 1978).

In fact, researchers have found that it is typically inconsequential whether jurors are given *any* test or standard. Mock jurors who receive no insanity definitions or who are told to use their "best lights" judgment to decide a case produce verdict patterns indistinguishable from those of mock jurors who receive various insanity test instructions (Finkel, 2000; Wheatman & Shaffer, 2001; Whittemore & Ogloff, 1995). As Diamond (1997) observed, many of the effects and their strength depend on the way the verdicts are formulated, the verdict alternatives, and the existence of contextual effects, for example. The fact that jurors determine whether a defendant is sane or insane *without* the guidance of specific legal instructions suggests that they rely on their own knowledge about insanity and other cognitive structures to make these decisions. Such does not imply that jurors "nullify" instructions; rather, that their own implicit theories of insanity and responsibility guide their interpretation of the admittedly vague and nonspecific linguistic terms of insanity standards.

Research on the GBMI verdict option indicates that mock jurors who opt for the GBMI verdict option (when given three options: NGRI, GBMI, and guilty) tend to be more moderate in their ratings of deserved blame and punishment as well as in their ratings of a defendant's level of mental disorder, capacity to display rational behavior, and capacity to control psychotic beliefs than mock jurors selecting NGRI or guilty verdicts (Roberts, Sargent, & Chan, 1993). Poulson, Wuensch, and Brondino (1998) found that the addition of the GBMI verdict option resulted in a twofold effect: (1) a reduction in guilty verdicts by about two-thirds and (2) a reduction of NGRI verdicts by about one-half. These authors concluded that the GBMI verdict appears to function as a compromise verdict.

Poulson and his colleagues (Poulson, Braithwaite, Brondino, & Wuensch, 1997; Poulson et al., 1998) found that mock jurors' attitudes were reflected in their verdict selections. For example, jurors who opted for guilty verdicts (as opposed to GBMI or NGRI verdicts) held a crime-control orientation, favorable attitudes toward the death penalty, and unfavorable attitudes toward the insanity defense. The opposite was true of those jurors who opted for an NGRI verdict (over guilty or GBMI).

Research on judicial instructions has highlighted the importance of providing jurors with the opportunity to deliberate. Wheatman and Shaffer (2001) found that dispositional instructions had no effect on the verdict preferences of individual jurors (i.e., individuals who were not given the opportunity to deliberate but instead made verdict decisions on their own and immediately after being presented the trial stimuli); however, after given the opportunity to deliberate as a jury, postdeliberation shifts in mock jurors' initial verdict preferences were evident. *Uninstructed* juries (i.e., those juries provided with no information about the treatment and detainment of individuals acquitted by reason of insanity) were more likely to shift toward a harsher verdict after deliberation, whereas *instructed* juries (i.e., those provided with dispositional information) were more likely to shift toward more lenient verdicts after deliberation. These results underscore the importance not only of dispositional instructions but also the opportunity for jurors to deliberate as juries.

RESEARCH ON JURY AND JUROR DECISION MAKING

Although data on the reliability and validity of well-founded forensic criminal responsibility opinions are encouraging, no modern studies of actual juries or bench trials in terms of the defendant, expert testimony for and against mental state, and case-specific factors are weighed in accepting or rejecting insanity claims. Logically, we can place some weight on surveys of attitudes toward insanity and insanity dispositions and compare them to what is empirically known, most of which is based on jury simulation studies. Problems arise with this particular methodology; however, a careful analysis of analog studies produces a rather consistent set of findings and implications.

Although the legal system implicitly assumes that people are blank slates who can apply the law in a wholly evidence-based fashion, research indicates that people have "knowledge structures" that reflect their life experiences and guide their behavior (Fiske, 1993; Schneider, 1991). These knowledge structures include constructs such as attitudes, schemas, prototypes, and stereotypes, and they appear highly relevant to legal decision making (see Moran, Cutler, & DeLisa, 1994; Pennington & Hastie, 1986; Stalans, 1993). Several sources of research suggest that individual differences in these structures are critical in understanding why jurors reach particular verdicts in insanity defense cases.

Jurors' Case-Relevant Attitudes. The insanity defense is controversial and involves scientific as well as political-moral issues. Although public opinion polls and empirical studies often find support for the basic logic of the insanity defense, they consistently reveal powerful negative attitudes toward it (Borum & Fulero, 1999; Cutler, Moran, & Narby, 1992; Skeem, Eno Louden, & Evans, 2004). For example, Roberts, Golding, and Fincham (1987) found that, although 78% of their subjects believed that severe mental illness suggested impairment in one's capacity to make rational decisions and form criminal intent, 66% believed that insanity should *not* be allowed as a complete criminal defense. Across studies, results reflect a primary concern that the insanity defense is an easily abused loophole in the law that allows many offenders to escape punishment (Silver et al., 1994). Additional concerns include beliefs that insanity is easily feigned and that the public is poorly protected from dangerous criminals who are adjudicated insane (Golding, 1992; see also Perlin, 1994, chap. 5). Similarly, jurors' case-specific negative attitudes toward both mental health experts and individuals with severe mental illness appear to affect their decision making in insanity defense cases (Cutler et al., 1992; Perlin, 1994; Skeem & Golding, 2001).

Many of these concerns reflect inaccurate knowledge about the insanity defense. Such myths are not only prevalent but may also be inflexible. Jeffrey and Pasewark (1984) presented participants with factual statistics on the frequency and success rate of the insanity defense. Approximately *half* of the participants maintained their opinion that the insanity defense was overused and abused despite having seen contradictory evidence. Especially troubling is the robust finding that these prevalent, potentially inflexible, negative attitudes toward the insanity defense considerably influence mock jurors' verdicts in insanity cases (Bailis, Darley, Waxman, & Robinson, 1995; Robinson & Darley, 1995). For example, Roberts and Golding (1991) found that mock jurors' attitudes toward the insanity defense were *more* strongly associated with their verdicts than were the study design variables, which included manipulations of available verdict categories (insanity versus insanity supplemented by guilty but mentally ill) and case facts (the relationship of the defendant's delusion to the crime and the planfulness of the crime). The most determinative dimension underlying these attitudes was mock jurors' beliefs in strict liability (versus a belief that mental state is relevant to a defendant's blameworthiness). Thus, jurors' verdicts may depend more on their attitudes and opinions than on case facts and court instruction.

Despite strong evidence of the biasing effect of negative attitudes toward the insanity defense on verdicts, bias often may go undetected as a result of limitations in current knowledge and legal procedures. Sloat and Frierson (2005) found that only about 4% of their sample ($n = 96$ qualified jurors) could correctly identify the definitions and dispositions of defendants found NGRI and GBMI. Despite the abundance of research on insanity defense attitudes, no well-validated measure of these attitudes has been developed. Skeem and Golding (2001) presented one of the first checklists of jurors' conceptions that, they suggested, could be adapted into a

questionnaire to assess prospective jurors' conceptions of the "typical person who is not responsible for his criminal actions due to mental illness" (p. 607; see also Daftary-Kapur, Groscup, O'Connor, Coffaro, and Galietta's 2011 study wherein they discussed the scale they developed to assess laypersons' knowledge of the insanity defense). Except in cases involving interracial violent crimes, capital punishment, or pretrial publicity, judges have broad discretion in selecting the topics to be addressed during *voir dire* (Johnson & Haney, 1994; Sklansky, 1996). Although insanity defense cases likely invoke equally powerful biases, the case law reflects a trend in which judges refuse to inquire about bias against the insanity defense or even allow the empanelment of jurors who express biases against the defense or against persons with mental illness (Perlin, 1994). In our opinion, the *voir dire* process ideally would be reformed to allow for routine examination of prospective jurors' case-relevant preconceptions and attitudes in insanity defense cases.

Jurors' Case Construals. Additional lines of research suggest that jurors' views are critical. In addition to their case-specific attitudes, jurors' individual ways of interpreting evidence are related to their verdicts. Mock jurors draw *different* inferences about defendants' cognitive and volitional impairments when given *identical* case descriptions (Bailis et al., 1995; Roberts et al., 1993; Whittemore & Ogloff, 1995). These inferences, in turn, strongly predict their verdicts. For example, Roberts and Golding (1991) presented mock jurors with case vignettes in which they manipulated available verdict categories, the relationship of the defendant's paranoid delusion to the crime, and the planfulness of the crime. The attitude-related ways in which mock jurors interpreted the case evidence were the most powerful predictors of verdict choice. For example, individual differences in jurors' *perceptions* of the extent to which a defendant was mentally disordered, capable of rational behavior, capable of acting differently, or capable of understanding the wrongfulness of his or her behavior explained substantially more variance in verdicts than did the objective manipulation of case evidence.

Finkel and Handel (1989), using different methodology, also found that jurors actively construct the meaning of case information in rendering verdicts. They presented mock jurors with four vastly different case vignettes and asked them to render a verdict and explain the reasoning underlying their decisions. Using a rationally derived categorization scheme, they found that mock jurors cited multiple, rational reasons for their decisions in each case (the categorization scheme included, e.g., capacity–incapacity to make responsible choices, unimpaired–impaired awareness and perceptions, and no motive–evil motive for criminal act). The pattern of the cited constructs or reasons systematically *differed* based on the mock jurors' verdicts. In essence, then, jurors construed case information in complex, discriminating ways that were consistent with their verdict choices.

Similarly, Whittemore and Ogloff (1995) found that differences in mock jurors' *perceptions* of a defendant's mental state at the time of the trial predicted their verdicts. Despite manipulation of the defendant's mental state at the time of trial

(i.e., symptom-free, neurotic, or psychotic), mock jurors differed in their perceptions of the extent to which given defendants were mentally disordered. When mock jurors inferred that the defendant was psychotic at the time of the trial, they were more likely to conclude that the defendant was insane (at the time of the offense).

Jurors' Implicit Theories or Prototypes of Insanity. Thus, jurors *construct* the meaning of case information. These "constructions," or interpretations, are more strongly associated with jurors' verdicts than the case as objectively presented and appear unaffected by judicial instruction. Based on these findings, several commentators have argued that jurors render insanity verdicts by carefully resorting to their personal knowledge or implicit theories of insanity (Finkel & Handel, 1989; Roberts & Golding, 1991; Roberts et al., 1987). However, the nature of these theories and the process by which they affect verdicts remains unclear. Researchers who have examined the nature of mock jurors' conceptions of insanity based on jurors' judgments about insanity case vignettes have found somewhat conflicting results with respect to the relative importance of various construal dimensions (see Roberts et al., 1987; cf. Bailis et al., 1995; Finkel & Handel, 1989; Robinson & Darley, 1995). To date, studies that *directly* analyze what people mean by "insane" in the context of a consideration of actual jury decisions are virtually nonexistent: "'What everybody knows' about insanity is perilously uncharted" (Perlin, 1994, p. 294).

Finkel (1995) and colleagues conducted research examining jurors' conceptions of insanity by using a prototype theory of categorization. Finkel and Groscup (1997) found that undergraduate subjects described insanity cases as involving young defendants with a history of strain, mental disorder, violence, and abuse who perpetrate various crimes, including murder, on strangers after various precipitating events (including the loss of loved ones). For successful insanity cases, the defendant's motive is related to a grandiose delusion; for unsuccessful cases, revenge is the motive.

Skeem and Golding (2001) identified three prototypes of insanity that were systematically related to jurors' case-relevant attitudes and demographic characteristics: (1) *severe mental disability*, the prototype representing the majority (47%) of the jurors and characterized by an emphasis on severe, long-standing, functional impairment and intellectual disability that is resistant to treatment; (2) *moral insanity*, the prototype representing about one-third (33%) of the jurors and characterized by an emphasis on traits of psychopathy, psychosis, and violent, unpredictable behavior; and (3) *mental state centered*, the prototype representing about one-fifth (21%) of the jurors and characterized by a narrow focus on issues relevant to the defendant's impaired MSO. These prototypes were related to differences in the ways jurors interpreted case information and rendered verdicts, such that the jurors with mental state-centered prototypes were more likely to render verdicts of NGRI and more likely to perceive defendants as less worthy of punishment, less able to control their beliefs, and more mentally disordered.

Jurors' decisions are not determined solely by subjective factors. Although individual differences in social-moral cognition appear most critical in understanding jurors' verdicts, objective manipulation of case facts does have some impact on jurors' verdicts. For example, the level of a defendant's mental disorder and the planfulness and bizarreness of the crime are associated with jurors' verdicts (Roberts & Golding, 1991; Roberts et al., 1987). As noted earlier, the characteristics of insanity acquittees suggest that juror and judicial decision making is rational and relatively consistent.

SUMMARY AND CONCLUSIONS

In this chapter we provided a broad overview of a number of issues and considerations regarding empirical, legal, and clinical aspects of criminal responsibility. There has been a great deal of discussion and controversy surrounding various legal standards or tests of criminal responsibility, and research indicates that this discussion and controversy may be all for naught, given that judicial instructions do not appear to have a significant impact on juries' verdicts. This being said, however, recent research has begun to address various issues requiring further consideration. Such issues include the role that deliberation plays in jury decision making and verdicts and in the impact of juror prototypes on verdicts and the interpretation of case information. Future research that incorporates samples of jury-eligible adults (as opposed to simple samples of convenience, such as from undergraduate psychology subject pools) will help further this important body of knowledge. Similarly, clinically oriented research will help further develop our assessment techniques and interventions for defendants for whom criminal responsibility is an issue.

REFERENCES

- American Bar Association. (1983). *Recommendations on the insanity defense*. Washington, DC: Author.
- American Law Institute. (1962). *Model penal code*. Philadelphia, PA: Author.
- American Psychiatric Association. (1982). *Statement on the insanity defense*. Washington, DC: Author.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- Appelbaum, P. S., & Gutheil, T. G. (1982). Clinical aspects of treatment refusal. *Comprehensive Psychiatry*, 23, 550–566.
- Arrigo, B. A. (1996). The behavior of law and psychiatry: Rethinking knowledge construction and the guilty-but-mentally-ill verdict. *Criminal Justice and Behavior*, 23, 572–592.
- Bailis, D., Darley, J., Waxman, T., & Robinson, P. (1995). Community standards of criminal liability and the insanity defense. *Law & Human Behavior*, 19, 425–446.
- Ballantine, H. W. (1919). Criminal responsibility of the insane and feeble-minded. *Journal of the American Institute of Criminal Law and Criminology*, 9, 485–499.
- Barnhouse, R. (1986). How to evaluate patients' religious ideation. In L. Robinson (Ed.), *Psychiatry and religion: Overlapping concerns*. Washington, DC: American Psychiatric Press.

- Beasley, W. R. (1983). An overview of Michigan's guilty but mentally ill verdict. *Michigan Bar Journal*, 62, 204–205; 215–217.
- Black, H. C. (1979). *Black's law dictionary* (5th ed.). St. Paul, MN: West.
- Bonnie, R. J. (1983). The moral basis of the insanity defense. *American Bar Association Journal*, 69, 194–197.
- Bonnie, R. J., & Slobogin, C. (1980). The role of mental health professionals in the criminal process: The case for informed speculation. *Virginia Law Review*, 66, 427–522.
- Borum, R. (2003). Not guilty by reason of insanity. In T. Grisso (Ed.), *Evaluating competencies* (2nd ed.). New York, NY: Kluwer/Plenum Press.
- Borum, R., & Fulero, S. (1999). Empirical research on the insanity defense and attempted reforms: Evidence towards informed policy. *Law and Human Behavior*, 23, 375–394.
- Bumby, K. M. (1993). Reviewing the guilty but mentally ill alternative: A case of the blind "pleading" the blind. *Journal of Psychiatry and Law*, 21, 191–220.
- Callahan, L. A., & Silver, E. (1998a). Factors associated with the conditional release of persons acquitted by reason of insanity: A decision tree approach. *Law and Human Behavior*, 22, 147–163.
- Callahan, L. A., & Silver, E. (1998b). Revocation of conditional release: A comparison of individual and program characteristics across four U.S. states. *International Journal of Law and Psychiatry*, 21, 177–186.
- Cirincione, C., & Jacobs, C. (1999). Identifying insanity acquittals: Is it any easier? *Law and Human Behavior*, 23, 487–497.
- Cirincione, C., Steadman, H., & McGreevy, M. (1995). Rates of insanity acquittals and the factors associated with successful insanity pleas. *Bulletin of the American Academy of Psychiatry and Law*, 23, 399–409.
- Commonwealth v. Rogers, 7 Metc. (Mass.) 500 (1844).
- Crotty, H. D. (1924). The history of insanity as a defence to crime in English criminal law. *California Law Review*, 12, 105–123.
- Cutler, B., Moran, G., & Narby, D. (1992). Jury selection in insanity defense cases. *Journal of Research in Personality*, 26, 165–182.
- Daftary-Kapur, T., Groscup, J. L., O'Connor, M., Coffaro, F., & Galietta, M. (2011). Measuring knowledge of the insanity defense: Scale construction and validation. *Behavioral Sciences & the Law*, 29, 40–63.
- Delling v. Idaho, 568 U.S. 133 S. Ct. 504 (2012).
- Diamond, S. S. (1997). Illuminations and shadows from jury simulations. *Law and Human Behavior*, 21, 561–572.
- Dirks-Linhorst, P. A., & Kondrat, D. (2012). Tough on crime or beating the system: An evaluation of Missouri Department of Mental Health's not guilty by reason of insanity murder acquittees. *Homicide Studies: An Interdisciplinary & International Journal*, 16, 129–150.
- Durham v. United States, 214 F. 2d 862 (D.C. Cir., 1954).
- Federal Rules of Criminal Procedure, Rule 12.2 (c) 4 (2011).
- Ferguson, M., & Ogloff, J. R. P. (2011). Criminal responsibility evaluations: Role of psychologists in assessment. *Psychiatry, Psychology and Law*, 18, 79–94.
- Finkel, N. J. (1989). The Insanity Defense Reform Act of 1984: Much ado about nothing. *Behavioral Sciences and the Law*, 7, 403–419.
- Finkel, N. J. (1995). *Commonsense justice: Jurors' notions of the law*. Cambridge, MA: Harvard University Press.

- Finkel, N. J. (2000). Commonsense justice and jury instructions: Instructive and reciprocating connections. *Psychology, Public Policy, and Law*, 6, 591–628.
- Finkel, N. J., & Groscup, J. (1997). Crime prototypes, objective versus subjective culpability, and a commonsense balance. *Law and Human Behavior*, 21, 209–230.
- Finkel, N. J., & Handel, S. (1989). How jurors construe “insanity.” *Law and Human Behavior*, 13, 41–59.
- Fiske, S. T. (1993). Social cognition and social perception. *Annual Review of Psychology*, 44, 155–194.
- Flaum, M., Arndt, S., & Andreasen, N. (1991). The reliability of “bizarre” delusions. *Comprehensive Psychiatry*, 32, 59–65.
- Foster v. State, 646 S.E.2d 838 (Ga. 2008).
- Foucha v. Louisiana 504 U.S. 71 (1992).
- Garety, P., & Hemsley, M. (1994). *Delusions: Investigations into the psychology of delusional reasoning*. New York, NY: Oxford University Press.
- Golding, S. L. (1990). Mental health professionals and the courts: The ethics of expertise. *International Journal of Law and Psychiatry*, 13, 281–307.
- Golding, S. L. (1992). The adjudication of criminal responsibility: A review of theory and research. In D. K. Kagehiro & W. Laufer (Eds.), *Handbook of psychology and law* (pp. 230–250). New York, NY: Springer-Verlag.
- Golding, S. L., Eaves, D., & Kowaz, A. (1989). The assessment, treatment and community outcome of insanity acquittees: Forensic history and response to treatment. *International Journal of Law and Psychiatry*, 12, 149–179.
- Golding, S. L., & Roesch, R. (1987). The assessment of criminal responsibility: A historical approach to a current controversy. In I. B. Weiner & A. K. Hess (Eds.), *The handbook of forensic psychology* (pp. 395–436). New York, NY: Wiley.
- Goldstein, A. S. (1967). *The insanity defense*. New Haven, CT: Yale University Press.
- Gowensmith, W. N., Murrie, D. C., & Boccaccini, M. T. (2012, July 9). How reliable are forensic evaluations of insanity? *Law and Human Behavior*. doi:10.1037/lhb0000001
- Gray, S. (1972). The insanity defense: Historical development and contemporary relevance. *American Criminal Law Review*, 10, 559–583.
- Häfner, H., & Böker, W. (1982). *Crimes of violence by mentally abnormal offenders*. Trans. H. Marshall. Cambridge, England: Cambridge University Press.
- Harrow, M., Rattenbury, F., & Stoll, F. (1988). Schizophrenic delusions: An analysis of their persistence, of related premorbid ideas, and of three major dimensions. In T. Oltmanns & B. Maher (Eds.), *Delusional Beliefs*. New York, NY: Wiley.
- Heilbrun, K., & Griffin, P. A. (1993). Community-based forensic treatment of insanity acquittees. *International Journal of Law and Psychiatry*, 16, 133–150.
- Hermann, D. (1990). Autonomy, self-determination, the right of involuntarily committed persons to refuse treatment, and the use of substituted judgment in medication decisions involving incompetent persons. *International Journal of Law and Psychiatry*, 13, 361–385.
- Idaho Code § 18-207(1) (2004).
- Insanity Defense Reform Act of 1984, Public Law 98-473, Sections 401–406 (1984).
- Janofsky, J. S., Vandewalle, M. B., & Rappaport, J. R. (1989) Defendants pleading insanity: An analysis of outcome. *Bulletin of the American Academy of Psychiatry & Law*, 17, 203–211.
- Jaspers, K. (1963). *General psychopathology*. Trans. J. Hoenig & M. Hamilton. Manchester, England: Manchester University Press.

- Jeffrey, R., & Pasewark, R. (1984). Altering opinions about the insanity plea. *Journal of Psychiatry & Law, 11*, 29–40.
- Johnson, C., & Haney, C. (1994). Felony voir dire: An exploratory study of its content and effect. *Law & Human Behavior, 18*, 487–506.
- Jones v. United States, 103 S. Ct. 3043 (1983).
- Keedy, E. R. (1920). Criminal responsibility of the insane—A reply to Professor Ballantine. *Journal of American Institute of Criminal Law and Criminology, 10*, 14–34.
- Keilitz, I., & Fulton, J. P. (1983). *The insanity defense and its alternatives: A guide for policymakers*. Williamsburg, VA: National Center for State Courts.
- LaFave, W., & Scott, A. (1972). *Handbook on criminal law*. St. Paul, MN: West.
- Litwack, T. R. (2003). The competency of criminal defendants to refuse, for delusional reasons, a viable insanity defense recommended by counsel. *Behavioral Sciences and the Law, 21*, 135–156.
- Livermore, J. M., & Meehl, P. E. (1967). The virtues of M'Naghten. *Minnesota Law Review, 51*, 789–856.
- Lloyd, K. P., & Packer, I. K. (2011). Jury instructions in a case with a defense of not guilty by reason of insanity. *Journal of the American Academy of Psychiatry and the Law, 39*(3), 424–425.
- Lu, F., Lukoff, D., & Turner, R. (1997). Religious or spiritual problems. In T. A. Widiger, A. J. Frances, H. A. Pincus, M. First, R. Ross, & W. Davis (Eds.), *DSM-IV sourcebook*. (Vol. 3, pp. 1001–1016). Washington, DC: American Psychiatric Association.
- Lymburner, J. A., & Roesch, R. (1999). The insanity defense: Five years of research (1993–1997). *International Journal of Law and Psychiatry, 22*, 213–240.
- Manguno-Mire, G. M., Thompson, J. W., Jr., Bertman-Pate, L. J., Burnett, D. R., & Thompson, H. W. (2007). Are release recommendations for NGRI acquittees informed by relevant data? *Behavioral Sciences & the Law, 25*(1), 43–55.
- Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C. (1997). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (2nd ed.). New York, NY: Guilford Press.
- Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C. (2007). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (3rd ed.). New York, NY: Guilford Press.
- Melville, J. D., & Naimark, D. (2002). Punishing the insane: The verdict of guilty but mentally ill. *Journal of the American Academy of Psychiatry and Law, 30*, 553–555.
- Miller, R. (1994). Criminal responsibility. In R. Rosner (Ed.), *Principles and practice of forensic psychiatry* (pp. 198–215). New York, NY: Chapman & Hall.
- Miraglia, R., & Hall, D. (2011). The effect of length of hospitalization on re-arrest among insanity plea acquittees. *Journal of the American Academy of Psychiatry and the Law, 39*(4), 524–534.
- M'Naghten's Case, 8 Eng. Rep. 718 (1843).
- Mojtabai, R., & Nicholson, R. (1995). Interrater reliability of ratings of delusions and bizarre delusions. *American Journal of Psychiatry, 152*, 1804–1806.
- Monson, C. M., Gunnin, D. D., Fogel, M. H., & Kyle, L. L. (2001). Stopping (or slowing) the revolving door: Factors related to NGRI acquittees' maintenance of a conditional release. *Law and Human Behavior, 25*, 257–267.

- Moran, G., Cutler, B., & DeLisa, A. (1994). Attitudes toward tort reform, scientific jury selection, and juror bias: Verdict inclination in criminal and civil trials. *Law & Psychology Review, 18*, 309–328.
- Mullen, P. (1985). The mental state and states of mind. In P. Hill, R. Murray, & A. Thorley (Eds.), *Essentials of postgraduate psychiatry* (2nd ed., p. 335). London, England: Grune & Stratton.
- Ogloff, J. R., Roberts, C. F., & Roesch, R. (1993). The insanity defense: Legal standards and clinical assessment. *Applied and Preventative Psychology, 2*, 163–178.
- Ogloff, J., Schweighofer, A., Turnbull, S., & Whittemore, K. (1992). Empirical research regarding the insanity defense: How much do we really know? In J. Ogloff (Ed.), *Law and psychology: The broadening of the discipline* (pp. 171–210). Durham, NC: Carolina Academic Press.
- Oltmanns, T. (1988). Approaches to the definition and study of delusions. In T. Oltmanns & B. Maher (Eds.), *Delusional beliefs* (pp. 3–11). New York, NY: Wiley.
- Packer, I. K. (2009). *Best practices in forensic mental health assessment: Evaluation of criminal responsibility*. New York, NY: Oxford University Press.
- Palmer, C. A. (2000). The guilty but mentally ill verdict: A review and conceptual analysis of intent and impact. *Journal of the American Academy of Psychiatry and the Law, 28*, 47–54.
- Parsons v. State, 81 Ala. 577, 2 So. 854 (1866).
- Pasewark, R. (1986). A review of research on the insanity defense. In S. A. Shah (Ed.), *The law and mental health: Annals of the American Academy of Political and Social Science, 484*, 100–114.
- Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology, 51*, 242–258.
- People v. McQuillan, 221 N.W. 2d 569 (Supreme Court of Michigan, 1974).
- Perlin, M. (1994). *The jurisprudence of the insanity defense*. Durham, NC: Carolina Academic Press.
- Platt, A. M., & Diamond, B. L. (1965). The origins and development of the “wild beast” concept of mental illness and its relation to theories of criminal responsibility. *Journal of the History of the Behavioral Sciences, 1*, 355–367.
- Platt, A. M., & Diamond, B. L. (1966). The origins of the “right and wrong” test of criminal responsibility and its subsequent development in the United States: An historical survey. *California Law Review, 54*, 1227–1259.
- Poulson, R. L., Braithwaite, R. L., Brondino, M. J., & Wuensch, K. L. (1997). Mock jurors’ insanity defense verdict selections: The role of evidence, attitudes, and verdict options. *Social Behavior and Personality, 12*, 743–758.
- Poulson, R. L., Wuensch, K. L., & Brondino, M. J. (1998). Factors that discriminate among mock jurors’ verdict selections: Impact of the guilty but mentally ill verdict option. *Criminal Justice and Behavior, 25*, 366–381.
- Poythress, N., Melton, G. B., Pettila, J., & Slobogin, C. (2000). Commentary on “The Mental State at the Time of the Offense Measure.” *Journal of the American Academy of Psychiatry & the Law, 28*, 29–32.
- Ray, I. (1962). *A treatise on the medical jurisprudence on insanity*. Cambridge, MA: Belknap Press. (Originally published 1838)
- Relman, A. S., & Angell, M. (2002). Resolved: Psychosocial interventions improve clinical outcomes in organic disease (Con). *Psychosomatic Medicine, 64*, 558–563.

- Riggins v. Nevada, 112 S. Ct. 1810 (1992).
- Roberts, C. F., & Golding, S. (1991). The social construction of criminal responsibility and insanity. *Law and Human Behavior, 15*, 349–376.
- Roberts, C. F., Golding, G., & Fincham, F. (1987). Implicit theories of criminal responsibility: Decision making and the insanity defense. *Law and Human Behavior, 11*, 207–232.
- Roberts, C. F., Sargent, E. L., & Chan, A. S. (1993). Verdict selection processes in insanity defense cases: Juror construals and the effects of the guilty but mentally ill instructions. *Law and Human Behavior, 17*, 261–275.
- Robinson, P. & Darley, D. (1995). *Justice, liability and blame: Community views and the criminal law*. San Francisco, CA: Westview Press.
- Rogers, R. (1984). *Rogers Criminal Responsibility Assessment Scales (R-CRAS) and test manual*. Odessa, FL: Psychological Assessment Resources.
- Rogers, J. L., Bloom, J. D., & Manson, S. M. (1984). Insanity defenses: Contested or conceded? *American Journal of Psychiatry, 141*, 885–888.
- Rogers, R., Cavanaugh, J. L., Seman, W., & Harris, M. (1984). Legal outcome and clinical findings: A study of insanity evaluations. *Bulletin of the American Academy of Psychiatry and the Law, 12*, 75–83.
- Rogers, R., & Ewing, C. P. (1992). The measurement of insanity: Debating the merits of the R-CRAS and its alternatives. *International Journal of Law and Psychiatry, 15*, 113–123.
- Rogers, R., & Sewell, K. W. (1999). The R-CRAS and insanity evaluations: A re-examination of construct validity. *Behavioral Sciences and the Law, 17*, 181–194.
- Rogers, R., & Shuman, D. W. (2000). *Conducting insanity evaluations*. New York, NY: Guilford Press.
- Rogers, R., Wasyliw, O. E., & Cavanaugh, J. L. (1984). Evaluating insanity: A study of construct validity. *Law and Human Behavior, 8*, 293–303.
- Rosenberg, C. E. (1968). *The trial of the assassin Guiteau: Psychiatry and law in the Gilded Age*. Chicago, IL: University of Chicago Press.
- Sacks, M., Carpenter, W., & Strauss, J. (1974). Recovery from delusions: Three phases documented by patient's interpretation of research procedures. *Archives of General Psychiatry, 30*, 117–120.
- Sayre, F. B. (1932). Mens rea. *Harvard Law Review, 45*, 974–1026.
- Schneider, D. (1991). Social cognition. *Annual Review of Psychology, 42*, 527–561.
- Sedler, M. (1995). Understanding delusions. *Psychiatric Clinics of North America, 18*, 251–262.
- Shannon v. United States. 114 S. Ct. 2419 (1994).
- Silver, E. (1995). Punishment or treatment?: Comparing the lengths of confinement of successful and unsuccessful insanity defendants. *Law and Human Behavior, 19*, 375–388.
- Silver, E., Cirincione, C., & Steadman, H. J. (1994). Demythologizing inaccurate perceptions of the insanity defense. *Law and Human Behavior, 18*, 63–70.
- Simon, R. J. (1999). *The jury and the defense of insanity*. New Brunswick, NJ: Transaction.
- Skeem, J. L., Eno Loudon, J., & Evans, J. (2004). Venireperson's attitudes toward the insanity defense: Developing, refining, and validating a scale. *Law and Human Behavior, 28*, 623–648.
- Skeem, J. L., & Golding, S. L. (2001). Describing jurors' personal conceptions of insanity and their relationship to case judgments. *Psychology, Public Policy, and Law, 7*, 561–621.
- Skeem, J. L., Golding, S. L., Cohn, N. B., & Berge, G. (1998). Logic and reliability of evaluations of competence to stand trial. *Law and Human Behavior, 22*, 519–547.

- Sklansky, J. (1996). Right to a jury trial. [Special issue: 25th Annual Review of Criminal Procedure]. *Georgetown Law Journal*, 84, 1139–1160.
- Sloat, L. M., & Frierson, R. L. (2005). Juror knowledge and attitudes regarding mental illness verdicts. *Journal of the American Academy of Psychiatry and the Law*, 33(2), 208–213.
- Slobogin, C., Melton, G. B., & Showalter, C. R. (1984). The feasibility of a brief evaluation of mental state at the time of offense. *Law and Human Behavior*, 8, 305–321.
- Spitzer, R., First, M., Kendler, K., & Stein, D. (1993). The reliability of three definitions of bizarre delusions. *American Journal of Psychiatry*, 150, 880–884.
- Stalans, L. (1993). Citizens' crime stereotypes, biased recall and punishment preferences in abstract cases: The educative role of interpersonal sources. *Law and Human Behavior*, 17, 451–470.
- State v. Okie, 987 A.2d 495 (Me. 2010).
- State v. Pike, 49 N.H. 399 (1869).
- State v. Strasburg, 110 P. 1020 (Supreme Court of Washington, 1910).
- Steadman, H. J., & Braff, J. J. (1983). Defendants found not guilty by reason of insanity. In J. Monahan & H. Steadman (Eds.), *Mentally disordered offenders: Perspectives from law and social science* (pp. 109–129). New York, NY: Plenum Press.
- Steadman, H. J., McGreevy, M. A., Morrissey, J., Callahan, L. A., Robins, P. C., & Cirincione, C. (1993). *Before and after Hinckley: Evaluating insanity defense reform*. New York, NY: Guilford Press.
- Stredny, R. V., Parker, A. L. S., & Dibble, A. E. (2012). Evaluator agreement in placement recommendations for insanity acquittees. *Behavioral Sciences & the Law*, 30(3), 297–307.
- Stroud, D. A. (1914). *Mens rea or imputability under the laws of England*. London, England: Sweet & Maxwell.
- Sullivan, H. S. (1954). *The psychiatric interview*. Oxford, England: Norton.
- Taylor, P. (1985). Motives for offending among violent and psychotic men. *British Journal of Psychiatry*, 147, 491–498.
- Taylor, P., Garety, P., Buchanan, A., Reed, A., Wessely, S., Ray, ... Grubin, D. (1994). Delusions and violence. In J. Monahan and H. Steadman (Eds.), *Violence and mental disorder: Developments in risk assessment* (pp. 161–182). Chicago, IL: University of Chicago Press.
- Tellefsen, C., Cohen, M. I., Silver, S. B., & Dougherty, C. (1992). Predicting success on conditional release for insanity acquittees: Regionalized versus nonregionalized hospital patients. *Bulletin of the American Academy of Psychiatry and the Law*, 20, 87–100.
- Underwood v. State, 32 Michigan 1 (Supreme Court of Michigan, 1873).
- United States v. Brawner, 471 F. 2d 969 (D.C. Cir., 1972).
- United States v. Lyons, 731 F. 2d 243 (5th Cir., 1984a).
- United States v. Lyons, 739 F. 2d 994 (5th Cir., 1984b).
- United States v. Torniero, 735 F. 2d 725 (2nd Cir., 1984).
- Vitacco, M. J., Erickson, S. K., Kurus, S., Apple, B. N., Lamberti, J. S., & Gasser, D. (2011). Evaluating conditional release in female insanity acquittees: A risk management perspective. *Psychological Services*, 8, 332–342.
- Vitacco, M. J., Van Rybroek, G. J., Erickson, S. K., Rogstad, J. E., Tripp, A., Harris, L., & Miller, R. (2008). Developing services for insanity acquittees conditionally released into the community: Maximizing success and minimizing recidivism. *Psychological Services*, 5, 118–125.

- Wainwright v. Witt, 105 S.Ct. 844 (1985).
- Wales, H. W. (1976). An analysis of the proposal to "abolish" the insanity defense in S. 1: Squeezing a lemon. *University of Pennsylvania Law Review*, 124, 687–712.
- Walker, C. (1991). Delusion: What did Jaspers really say? *British Journal of Psychiatry*, 159, 94–103.
- Walker, N. (1978). *Crime and insanity in England, Volume I: The historical perspective*. Edinburgh, Scotland: University of Edinburgh Press.
- Wheatman, S. R., & Shaffer, D. R. (2001). On finding for defendants who plead insanity: The crucial impact of dispositional instructions and opportunity to deliberate. *Law and Human Behavior*, 25, 167–183.
- Whittemore, K., & Ogloff, J. (1995). Factors that influence jury decision making: Disposition instructions and mental state at the time of the trial. *Law and Human Behavior*, 19, 283–303.
- Wiederanders, M. R. (1992). Recidivism of disordered offenders who were conditionally vs. unconditionally released. *Behavioral Sciences and the Law*, 10, 141–148.
- Wiederanders, M. R., Bromley, D. L., & Choate, P. A. (1997). Forensic conditional release programs and outcomes in three states. *International Journal of Law and Psychiatry*, 20, 249–257.
- Wiederanders, M. R., & Choate, P. A. (1994). Beyond recidivism: Measuring community adjustments of conditionally released insanity acquittees. *Psychological Assessment*, 6, 61–66.
- Winters, K. C., & Neale, J. M. (1983). Delusions and delusional thinking in psychotics: A review of the literature. *Clinical Psychology Review*, 3, 227–253.

Specific Intent and Diminished Capacity

CHARLES R. CLARK

THE defense of diminished capacity, when it has been allowed in U.S. jurisdictions, never led forensic psychologists to adopt common or consensual approaches to assessment to an extent that occurred with the insanity defense in its various permutations and with other psycholegal issues such as competency. It often is uncertain just what questions the forensic examiner should address under the rubric of diminished capacity and what sort of clinical assessment data is needed to formulate opinions on the question. In large part, the lack of any standardized forensic assessment methodology in this area reflects the rarity of the diminished capacity defense and thus the call for expert testimony; more important, it reflects muddled and changing legal conceptualizations of the defense.

Although thoughtful ways to infuse more research-based psychology into legal conceptions of intent have been proposed (e.g., Barratt & Felthous, 2003; Denno, 2003; Malle & Nelson, 2003), in the absence of consistent guidance from the law itself, forensic examiners have been left to cobble together approaches to diminished capacity assessment—extrapolating from approaches used in assessment of criminal responsibility and related issues—that did not always address the legal questions in a relevant or coherent manner. The history of diminished capacity, more than that of any other psycholegal issue, is a reminder of the difficulties and pitfalls in adapting psychological formulations and assessment to any legal questions—in trying to speak about legal issues in psychological terms. In asking mental health professionals to offer opinions about a defendant's ability to form the intent necessary for conviction, the law proceeded on the assumption that its legal constructs have genuine psychological content that is open to clinical investigation. That this assumption is unwarranted in many of the cases in which diminished capacity is raised is evident from examining the development of this defense and the practical issues that arise in forensic evaluations of defendants. The defense may always have made more sense from a legal than from a psychological perspective (Clark, 1982).

Essentially supported by expert witnesses working on their own assumptions about the law, the use of diminished capacity at trial could lead to outcomes regarded by many as strange, even grotesque. That, or what might be called its impracticality, ultimately led to the repudiation of diminished capacity in various jurisdictions in the state in which it was largely forged as a modern approach to weighing intent—California. Where it has not simply been abolished, diminished capacity has been diminished as a defense, with a reassertion by legislatures and courts of older concepts of legal intent integral to American criminal law. It survives in those jurisdictions as a seldom-tried approach to reduce criminal liability, with predictably meager results, that is hemmed in by restrictive understandings of the basic terms involved and by a consequent restriction of the scope of expert testimony. Paradoxically, this stricter approach to gauging intent also survives in four states—Idaho, Kansas, Montana, and Utah—as a substitute for more traditional tests of legal insanity.

To understand how diminished capacity reached its current position in the arsenal of criminal defenses and what remains to be said by forensic examiners when the defense is raised, it is necessary to review the history of the approach and of the law's formulation of concepts of intent. A walk through the history and a look at the sorts of cases in which the defense has been raised may serve as a caution to limit the readiness of mental health professionals to weigh in on questions of intent and the capacity to form intent. Even in its more circumscribed state, diminished capacity continues to be a legitimate vehicle for introducing psychological considerations into trial proceedings, but it also continues to be one in which there is a danger that the expert witness will go beyond proper applications of clinical science in trying to resolve contradictions in the law regarding intent.

A claim of diminished capacity may be raised in cases in which it is unlikely that any court or jury would agree that the defendant was simply too disordered by mental disease or defect to be held accountable for a crime. Rather than insanity—a lack of criminal responsibility—the claim may be that the defendant suffered from some disordered state of mind that is less than psychotic in severity—perhaps one that was transitory, perhaps one that was induced by drugs or alcohol. On that account, the defendant, although guilty of something, is deemed as having been unable to form the intent to commit the major crime being charged. But this is not an issue of partial guilt. When diminished capacity is claimed, the question is not the extent to which the defendant should be considered guilty but instead precisely what crime the defendant is guilty of having committed. Diminished capacity is conceptually distinct from insanity in not addressing culpability—legal responsibility for illegal acts—but basic guilt instead.

ACTUS REUS AND MENS REA

In Anglo-American law, two elements must be present for a finding of guilt: a wrongful deed, or *actus reus*, and a wrongful purpose or criminal intent, *mens rea*.

Both of these elements, along with all the elements of the crime charged, must be established beyond a reasonable doubt in order to convict (*In re Winship*, 1970).

The person who unwillingly performs an illegal act, whether consciously or unconsciously, is not guilty of an offense. Although volition may be viewed as a mental construct—and it is explicitly included in certain tests of legal insanity—volition typically has not been considered relevant to the question of *mens rea*. The store clerk forced at gunpoint to tie up other employees and open a safe for a robber is not guilty, just as a medical patient delirious with fever is not guilty for striking and injuring a nurse. *Actus reus*—the designation of an act as wrongful—although always open to question and never assumed in a criminal prosecution, is seldom open to a psychological defense. Except in cases in which the insanity defense is used—and it has always been controversial there—questions of voluntariness ordinarily pertain to the issue of *actus reus*, not *mens rea* (Morse, 1999). The so-called automaton defense, for instance, in which a defendant asserts a lack of volition and often a lack of consciousness as well (and which may well exist as a pure legal fiction rather than as a legal description of a real clinical phenomenon), attacks the prosecution claim that there was an *actus reus* (Melton et al., 2007). It is in respect to the question of whether a defendant could have formed *mens rea* that forensic examiners are most likely to be asked to provide an expert opinion.

Insanity, in jurisdictions where it is still available as a defense, is the principal approach that challenges the prosecution's claim that the defendant, who typically has not contested the occurrence of *actus reus*, behaved with *mens rea*. Insanity negates *mens rea*—in the sense of mental guilt—altogether, in a global way, not in terms of particular, specific elements that may make up a crime, for instance, the intent to steal or to kill. The individual who meets the test of insanity, regardless of how insanity is defined, is not guilty, not because there was no *actus reus* or illegal act committed but because the individual lacked mental guilt or culpability.

Although insanity may be considered a *mens rea* defense in this way, some authorities hold that sanity is not a true element of the offense, at least not one that needs to be proven by the prosecution (and in the federal and many state jurisdictions, the burden of proving insanity is borne by defense). In this view, defendants are legally insane and not culpable even if they formed the intent ordinarily required for conviction (Steadman et al., 1993).

The meaning of *mens rea* has changed over time and contexts, from a term encompassing the concept of blameworthiness to one that more simply denotes the intent to cause a defined act (Steadman et al., 1993). In contrast to the wholesale negation of mental guilt and therefore criminal responsibility by the insanity defense, diminished capacity considers *mens rea* in terms of the presence or absence of discrete elements of intent. In theory, if not in practice, diminished capacity is not exculpatory and does not lead to outright acquittal of all wrongdoing but rather to a finding that a person was not capable of forming intent for the particular crime charged. What would follow, in theory, is a conviction of some lesser offense included in the

crime originally charged. Additionally limiting the scope of diminished capacity is the fact that not all crimes are susceptible to this partial defense.

GENERAL AND SPECIFIC INTENT

Central to the diminished capacity approach are the legal distinctions between crimes involving general intent and those involving specific intent. A diminished capacity defense typically is available only for specific intent crimes. Confusingly, diminished capacity often is referred to as a *mens rea* defense, along with insanity, although, strictly speaking, *mens rea* itself includes general intent as well as the more global moral blameworthiness involved in the concept of criminal responsibility or insanity. To add to the confusion, there is no universal agreement as to which crimes involve specific intent as opposed to general intent. The development of the law in this regard has been viewed as more a matter of expediency than the logical result of legal theory (Dix, 1971).

Finally, the essential logic behind diminished capacity or *mens rea* defenses is virtually identical to that of the so-called intoxication defense, involving voluntary intoxication, which also employs the distinction between general and specific intent. Functionally, diminished capacity and intoxication defenses may be indistinguishable. In one investigation of sequential pretrial referral for criminal responsibility and insanity examinations to Michigan's centralized diagnostic facility, the Center for Forensic Psychiatry, during the heyday of the diminished capacity defense in that state, it was found that most (77%) of those defendants for whom diminished capacity was raised as a defense reported substance abuse at the time of the offense. By contrast, among those defendants for whom insanity but not diminished capacity was raised, only 39% reported intoxication at the time of the offense (Clark, 1988). Because of the essential similarity of its conceptual foundations to diminished capacity, the intoxication defense, as it involves a claim of incapacity to form intent, need not be treated separately. For a broader discussion of other ways in which intoxication or addiction may affect the grading of responsibility or findings of guilt, see Melton et al. (2007).

Distinctions between general and specific intent follow a ranking of criminal culpability that is perhaps best illustrated by the American Law Institute (ALI) Model Penal Code (MPC; 1962), which included the recommendation that a defendant's level of culpability should be measured by an examination of mental state with respect to all elements of the offense (p. 24). The drafters proposed that liability be assigned depending on whether the offender acted purposely, knowingly, recklessly, or negligently. *Purposeful* intent, the highest level of criminal intent, occurs when the offender has the conscious object of committing the act or causing a particular result. Intent that is *knowing* involves awareness by the perpetrator of the nature of the criminal conduct or its circumstances. A *reckless* intent occurs when the offender consciously disregards the substantial and unjustifiable risks involved in the conduct, and *negligent* intent exists when the actor should have been aware of

the risk. The term *willfulness*, which appears in a number of statutory definitions of crimes, was proposed as an intent element corresponding to knowing intent. In the MPC hierarchy of intent, specific intent, which involves the imputation of a positive subjective mental state—unlike negligence and possibly recklessness as well—most closely corresponds to purposeful and knowing intent (Melton et al., 2007). The drafters of the MPC proposed that negligence, in any case, roughly corresponds to the common-law requirement of general intent (p. 23).

In theory, specific intent may be regarded as a higher-order or more seriously criminal purpose than general intent. Another way to understand the distinction is that general intent crimes are seen as those involving the simple intent to commit the illegal act itself, although that illegal act may be quite serious, such as rape. Persons convicted of general intent crimes should have been conscious of their actions and their predictable results (Melton et al., 2007). By contrast, specific intent crimes usually are seen as requiring the intent to achieve some additional result beyond the consequences of the general intent crime itself.

Thus, in some jurisdictions, a person who breaks into a house may be guilty of a general intent crime only (perhaps only a misdemeanor, such as trespass), whereas breaking and entering with the further intent of committing larceny always would be viewed as a specific intent offense. Similarly, a sexual assault may be a general intent crime in some jurisdictions, whereas breaking and entering with intent to rape, or assault with intent to rape, would be specific intent offenses. In the solitary case of murder in the highest degree, specific intent may consist of premeditation and deliberation—supposedly greater intent elements more demanding than intent to kill with malice aforethought.

With the exception of first-degree murder, a specific intent offense may entail the actual accomplishment of some effect beyond that of the general intent crime, but it need not. The would-be rapist may be stopped before he actually perpetrates the rape, for instance. *Actus reus* in that case would refer to the deed done in pursuit of the specifically intended goal. Whether specific intent results in a completed act or not, the specific intent crime may be regarded as more serious than an included or subordinate general intent offense—for example, in regard to the possible penalty—due to the enormity of the act intended.

In some jurisdictions, any crime that by statute incorporates an explicit intent, such as assault with intent to commit murder, is on that account a specific intent offense. Crimes of larceny typically are viewed as specific intent offenses regardless of their circumstances or the nature and value of what is stolen. Larceny involves not merely the unlawful taking of another's property but taking it with the further intent of carrying it away, converting it to one's own use, or otherwise depriving the owner of it (Black, 1979).

In any jurisdiction in which diminished capacity or *mens rea* defenses have been permitted, the highest degree of homicide—usually designated first-degree murder—is considered to be a specific intent crime. In fact, as seen in the discussion of legal developments in this area, much of the focus of case law with respect

to diminished capacity has been on murder. In this instance, the specific intent consists of premeditation and deliberation. Those elements, operationally unitary in the sense that both need to be present for a first-degree murder to be committed, are subject to somewhat different constructions in various jurisdictions. Generally, however, they refer respectively to plotting, contriving, planning, or thinking about the killing beforehand and weighing and examining the reasons for and against a contemplated act or course of conduct, acts, or means (Black, 1979).

With short-lived exceptions only in California, what is usually called second-degree murder, which requires malice aforethought but not premeditation and deliberation, has been regarded as a general intent offense that is not susceptible to a diminished capacity defense. Despite its intuitive connotations of evil intent, malice survives in law as a mental element that is simply a term of art, a shorthand designation for one of a number of mental elements that would satisfy the requirements for a murder conviction, as opposed to a conviction of the lesser offenses of manslaughter or negligent homicide (Morse, 1979, 1984). The malice required for a second-degree murder conviction could involve the intent to kill, but it also could call for no more than the intent to cause great bodily harm, the willful and wanton disregard of risk to life, or simply the commission of another felony during which a homicide occurs (LaFave & Scott, 1972).

To recapitulate, specific intent offenses that potentially are open to a diminished capacity or *mens rea* defense are somewhat arbitrarily determined but include at least premeditated and deliberated murder, larceny, and those offenses defined to explicitly incorporate some further intent.

ELABORATION OF DIMINISHED CAPACITY DOCTRINE IN CALIFORNIA

The development of *mens rea* doctrine, prior to the development of modern psychology with its complex hypotheses explaining human behavior and motivation, may be viewed by mental health professionals today as never incorporating anything more than unsophisticated, commonsense notions of intent. For years, the law in this area did not invite deeper explorations of the dynamics of intentional conduct, such as its developmental origins or its expression in unconscious drive or, from a behavioral perspective, the individual's learning history. In gauging whether a person was capable of criminal intent, the law did not reflect any particular concern about factors now considered by psychologists as inextricably involved in intentional conduct, such as mood and its biological concomitants, cognitive expectations, learned attitudes, or maturation—especially in cortical development. Arguably, much of what psychologists would see as pertinent to thinking about intent could be viewed as aspects of volition, but volition or what is thought of as free will—the objective or subjective freedom of the individual to create intention and purpose—was not recognized as germane to *mens rea*. Aside from those jurisdictions that employed an insanity test with a volitional prong—some variation of the irresistible impulse test—considerations of volition were judged relevant, if at

all, only to the question of whether the *actus reus* requirement was met (Bonnie & Slobogin, 1980). Conduct that was compelled in some manner might deserve to be excused in some instances, but not because intentionality, conceived of as a purely cognitive process, had been absent or impossible. In one view, the lack of volition in committing a crime negates not *mens rea* but rather *actus reus*: There is no criminal act that requires further (mental) explanation (Dix, 1971; Erlinder, 1983). Accordingly, there would be no need for expert witnesses if volition was all that was at issue.

All of this changed over the course of about a quarter century in California. The changes wrought there eventually produced a backlash of outrage engendered by the deeply unpopular results in some notorious cases. However, the effects of those changes continued to resonate in other jurisdictions.

As was the case in a handful of other states and the federal jurisdictions, California dealt with the dilemma of protecting the accused's right to avoid self-incrimination when pleading insanity—an affirmative defense requiring the commission of a criminal act—by bifurcating the trial into a guilt phase and an insanity phase. In this way, a determination could be made that the defendant was indeed the one who committed the act before it would be determined if the defendant was culpable or criminally responsible. It might be thought that a determination of guilt intrinsically involves a determination that not only *actus reus* but also *mens rea* is proven. Because insanity, in theory, negates *mens rea*, it is natural to wonder how evidence of mental abnormality could ever be excluded from the guilt phase of a trial. If guilt as well as criminal responsibility is mental as well as physical, expert testimony about mental disorder might require presentation at the guilt phase. This was the conclusion reached by the California Supreme Court in *People v. Wells* (1949), and it made possible a series of decisions expanding the opportunities for expert testimony about diminished capacity.

Wesley Wells was serving a sentence in a California penitentiary when he allegedly assaulted prison guards, throwing a heavy crockery cuspidor and injuring one guard severely. The offense with which Wells was charged was potentially a capital offense. Accordingly, at the guilt phase of his trial Wells attempted to introduce evidence on the question of whether he had entertained malice. Defense experts intended to testify that Wells was experiencing tension resulting from fears for his own safety at the time he assaulted the guards with the cuspidor. In line with bifurcation procedures, this testimony was ruled by the trial court as inadmissible. On appeal of Wells's conviction, the California Supreme Court held that the evidence of Wells's claimed abnormality was indeed material to the question of his guilt, and it had been an error to exclude this testimony from the guilt phase.

By breaching the separation between guilt and culpability, the *Wells* decision influenced the course of diminished capacity law in California. The real impetus to introduce testimony about mental state into determination of guilt, rather than reserve it solely for determination of insanity, was the perceived inadequacy of the insanity test itself (Morse, 1979). At the time, California, like many states, was using a variation of the century-old M'Naghten Rule (M'Naghten's Case, 1843).

To establish insanity, it had to be proven that, at the time of the act, as a result of a defect of reason, from disease of the mind, the defendant did not know the nature and quality of the act he was doing, or if he did know it, that he did not know he was doing what was wrong. This seemingly all-or-nothing, black-or-white, and wholly cognitive test of insanity seemed to make no allowance for other important considerations, such as volition or the ability to resist committing an act even disturbed individuals might recognize as wrong but that their mental disorder impelled them to do.

Some felt that the M'Naghten standard for insanity was too narrow a test to provide a just and humane result in many cases involving obviously disordered individuals. The expansion of diminished capacity into areas it had never gone before seemed likely to compensate for the shortcomings of the insanity test. As the California Supreme Court later acknowledged in *People v. Henderson* (1963), diminished capacity became the means by which the courts could ameliorate the harshness of the M'Naghten standard that had been imposed by the legislature; it came to be seen as the ameliorative defense. Its first application came in the case of *People v. Gorshen* (1959).

Nicholas Gorshen was a Russian immigrant and a longshoreman. He reportedly went to work intoxicated one day and was sent home by his foreman, after fighting with him. He got a gun, went back to the docks, and shot his foreman dead in front of several witnesses. At his trial for first-degree murder, a psychiatrist, Dr. Bernard Diamond, testified that Gorshen had been impaired by schizophrenia for years and, when his foreman told him to leave work, his precarious psychological equilibrium was threatened. He viewed the demand as an attack on his manhood, and he was compelled to retaliate against the source of the threat. Dr. Diamond did not dispute Gorshen's admission that he had consciously intended to shoot his foreman, but he testified that, in any case, Gorshen did not possess the mental state required for malice aforethought or anything implying intention, premeditation, or deliberation.

No testimony was presented rebutting Dr. Diamond's diagnosis of schizophrenia, but Gorshen was found sane by the judge at a bench trial under California's M'Naghten-based insanity standard and convicted of first-degree murder. On appeal, the California Supreme Court upheld Gorshen's conviction but added that the expert testimony that he would have been incapable of the malice required for murder—testimony considered but rejected by the trial court—had been properly admitted. By its ruling, the court endorsed what amounted to the redefinition of malice that had been offered by the expert witness, one that involved a volitional component (Dix, 1971) that previously had been absent not only from the concept of malice but from the insanity test. The California Supreme Court held that malice exists when an individual commits an act intentionally, of his or her own free will, rather than as a result of an abnormal compulsion. The way seemed clear to provide greater consideration of subjective psychological factors in guilt determinations. If traditional *mens rea* concepts such as insanity and malice did

not permit relevant expert testimony on mental abnormality, those intent elements might be so construed as to permit a judge or jury to weigh the defendant's ability to morally assess or control his or her conduct (Morse, 1979). This process was clearly at work in the next significant case, *People v. Wolff* (1964).

Dennis Wolff, charged with murder, may have been more obviously disturbed than Gorshen, yet like him, he did not seem to qualify for an insanity verdict under the M'Naghten standard. Wolff was only 15 years old when he allegedly developed a plan to kidnap girls and bring them home for sex. For the plan to work, he needed to get his mother out of the way. After one failed attempt, he succeeded in this by beating her to death with an ax handle. Despite expert testimony that he had schizophrenia and was legally insane, Wolff was convicted of first-degree murder.

As in the Gorshen case 5 years earlier, the California Supreme Court in *Wolff* upheld the jury's finding of sanity, but it boldly held that Wolff had not been capable of the mental process needed to commit first-degree murder, namely premeditation. It was plain that Wolff had carefully prepared the homicide and had given considerable thought to it. However, the court held that more was required for premeditation than Wolff had been capable of, given his level of impairment. The defendant must have been able, the court ruled, to "maturely and meaningfully reflect upon the gravity of his contemplated act" (*People v. Wolff*, 1964, p. 821). Accordingly, the court found that Wolff could not have been guilty of more than second-degree murder. In essence, it had declared that it was not enough that a defendant engaged in what might resemble commonsense notions of premeditation and deliberation. Rather, the "quantum of his moral turpitude and depravity" (*People v. Wolff*, 1964, p. 822) needed to be ascertained. At issue was not the apparent fact of intent but the quality of whatever intent had been formed. Although *mens rea* might otherwise be apparent, it could be negated by a finding that the defendant was morally incapable of true intent. This naturally begged the question of just what the real element looked like, but it opened the door to an expanded consideration of expert testimony on the issue.

A growing trend seemed to validate the hope expressed by Dr. Diamond (1961), the defense expert in the Gorshen case, that diminished capacity might become a vehicle for the introduction of a richer, expanded view of the role that mental health issues play in criminal behavior. With this new approach, he wrote, society could no longer evade its obligations to provide defendants therapeutic help, because they would now officially be labeled as sick, and the courts would have publicly acknowledged the need for treatment. Indeed, as time went on and new developments took place, diminished capacity would be regarded as a finely tuned instrument for the defense. Defense attorneys could imagine how even such factors as the stress of being a prisoner could be introduced to mitigate guilt (Marx, 1977).

In the decade following *Wolff*, an expanded view of intent elements continued to hold sway in California. *People v. Conley* (1966) saw the *Wolff* logic regarding premeditation applied to the question of malice aforethought. William Conley allegedly had been drinking heavily for a number of days when he killed his

estranged lover and her husband. He later claimed that he had no intention to kill anyone and had no memory of having done so. Expert testimony was heard that the amount of alcohol Conley had consumed would have impaired the judgment of an ordinary person. A psychologist testified that Conley was in a dissociative state at the time of the homicides and could not function normally. Arguing that diminished capacity could negate the malice aforethought required for murder, the defense requested that the jury also be instructed that it could find Conley guilty of voluntary manslaughter; this motion was denied, and the jury returned a two-count conviction on charges of first-degree murder. On appeal, the California Supreme Court reversed Conley's convictions because of the trial court's denial of the motion to instruct the jury on manslaughter as well as murder. What amounted to a redefinition of malice that had started in *Wells* and continued in *Gorshen* was elaborated on. For malice to have been present, the court held, the person must have been capable of comprehending his duty to govern his actions in accord with the duty imposed by law. This refinement or elaboration of malice stood in contrast to more traditional formulations. Although still cast in terms of *mens rea*, malice under *Conley* became a mini-insanity defense, a cognitive-affective version of the M'Naghten standard (Morse, 1979). In that way, diminished capacity did ameliorate the perceived harshness and strictures of the legislated insanity test.

In 1973, in *People v. Cantrell*, the California Supreme Court ruled that irresistible impulse—not recognized in California as part of its test of legal insanity—could not serve as a complete defense to a crime. For just that reason, however, the court held that a defendant claiming diminished capacity must be permitted at the guilt phase of trial to show by competent evidence that the act alleged was a result of irresistible impulse caused by mental disease. Such testimony, the court stated, could be brought to bear on issues of intent to kill as well as malice aforethought.

The recognition of a role for volitional considerations that was evident in *Conley* and *Cantrell* was further reinforced in *People v. Poddar* (1974). Prosenjit Poddar may be familiar to many as the killer of Tatiana Tarasoff and thus a subject of the leading case on the duty of mental health professionals to protect third parties endangered by their patients, *Tarasoff v. Regents of the University of California* (1976). Poddar was an Indian naval architecture student who reportedly was rejected by Tatiana Tarasoff, a fellow student to whom he had formed an attachment. After a period of despondency and emotional distress during which he reportedly disclosed to a university psychologist that he intended to kill his girlfriend when she returned from abroad, Poddar fatally stabbed Tarasoff after shooting at her with a pellet gun. At trial, defense experts described Poddar as schizophrenic, although the prosecution rebutted this testimony. The jury failed to find Poddar insane and convicted him of first-degree murder. Citing its rulings in *Conley* and *Cantrell*, however, the California Supreme Court overturned the conviction and ruled that Poddar could be convicted at most of voluntary manslaughter unless it could be established that he was both aware of his duty to act within the law and was capable of doing so.

Sweeping aside what had been a bar against using a *mens rea* or diminished capacity argument as a complete defense to a crime, the California Supreme Court ruled in *People v. Wetmore* (1978) that evidence of mental disorder could be used at the guilt phase of trial to negate any mental element, even if outright acquittal might result. Wetmore, a chronic psychiatric patient, was charged with burglary after he broke into an apartment based on his claimed belief that the apartment was his. As in many other cases of burglary—or breaking and entering with larcenous intent—if the specific intent to commit larceny can be negated, there might be no lesser included felony of which the defendant could be convicted.

DENOUEMENT: THE END OF DIMINISHED CAPACITY IN CALIFORNIA

In retrospect, such significant changes in the way mental guilt was assessed could not have gone unchallenged forever. California was a battleground on which advances in the use of diminished capacity—advances that influenced practice in other jurisdictions—were followed by a retrenchment that proved equally influential. The logical results of an expanded or enriched diminished capacity doctrine by the California Supreme Court galvanized opposition and contributed to its demise. As with the closely contemporaneous events regarding the insanity defense following John Hinckley's 1982 acquittal by reason of insanity for his attempted assassination of President Ronald Reagan, larger social trends were involved in the changes that reversed diminished capacity law.

Certainly the expanded view of diminished capacity had its critics even before worse came to worst. Dix (1971) detailed a variety of objections to it. Among them were fears that the successful use of diminished capacity would yield shorter prison terms for convicted criminals, particularly killers, and that some individuals would win outright acquittal by means of this defense, with no protection for society, even the psychiatric hospitalization mandated for those found legally insane. Dix cited fears by some that the issues involved in the new diminished capacity defense were too complex for juries to understand and that by default the question would be turned over to expert witnesses whose testimony would be admitted despite its unreliability. Echoing this concern, Morse (1979) asserted that California's diminished capacity approach had not provided any clear standards for judges and juries to apply. It was unclear from the *Wolff* decision, for instance, just how to determine whether a murderer's premeditation and deliberation were "mature" and "meaningful."

The expanding and increasingly crucial role of expert witnesses in determinations of diminished capacity was at the crux of the problem seen by critics of the expanded diminished capacity defense. Dix (1971) complained that the admission of expert testimony on this issue was an error. The *Wells* decision, he wrote, leapt from a finding that evidence of psychological abnormality had some logical relevance to the question of guilt to a conclusion that such evidence was therefore admissible. A critical analysis of the way psychiatric testimony was actually used in these cases,

Dix offered, indicated that the testimony never actually addressed the question of whether some state of mind required for the imposition of criminal liability was absent but instead supported an entirely different claim, namely, that both the defendant's actions and his state of mind were the result of unconscious influences. In the view of Dix and others, citing psychological abnormality as a way of disproving intent was creating a legal fiction that simply permitted a psychological explanation for the behavior to be offered in court. Dix saw this as placing the expert witness in an unfair position, playing a ritualistic role in the proceedings to mouth magic words that would mitigate the defendant's guilt.

Arenella (1977) viewed the California developments similarly. The *Wolff* decision, he wrote, shifted the focus away from whether the defendant entertained the requisite intent to how and why the defendant entertained it. In this manner, diminished capacity, ostensibly an investigation of a defendant's capacity for intent, had become a disguised version of diminished responsibility, an entirely different concept. Where diminished responsibility has been used, as in some European jurisdictions, mental abnormality simply mitigates or reduces the level of culpability of a criminal, but it does not do so by seeking to disprove the defendant's ability to form some element of intent. If diminished capacity were employed honestly, as the *mens rea* doctrine it purported to be, Arenella argued, it would involve only evidence that actually had a bearing on whether requisite intent was missing, and it would rarely serve any purpose not already served by the insanity defense.

Morse (1979, 1984) agreed that diminished capacity had come to be treated as diminished responsibility. Indeed, he pointed out, there is no easy way for mental health testimony that is strictly confined to questions of *mens rea* to rule out or negate the capacity of the defendant to form those intent elements. Morse (1984) saw no danger in a strictly applied *mens rea* approach to diminished capacity, which he regarded as unlikely to benefit the defendant in any case. Anticipating rulings by federal courts in the coming decade, Morse asserted that a strict *mens rea* defense, challenging the prosecution's claim that all requisite intent elements were present at the time of a crime, is constitutionally protected.

It is not apparent, of course, that the essential rationale behind the expansion of diminished capacity doctrine in California was the provision of greater or more comprehensive psychiatric or psychological explanations for criminal behavior—as if the law wanted to hear more from experts. Nor was the increased involvement of mental health experts in court proceedings necessarily seen as desirable, although that was what happened during this time with respect to a variety of legal issues. The apparent rationale for the expanded diminished capacity approach was the perception that existing statutes inadequately considered the influence of mental abnormality and that the harshness of the M'Naghten insanity standard, in particular, needed to be ameliorated.

If the California diminished capacity approach was based in this perceived need, it was deprived of its rationale when the California Supreme Court, in *People v. Drew* (1974), decided that it could, in effect, bypass the legislature and ruled that

the ALI MPC insanity test (ALI, 1962) could serve as the insanity test in California. Always before, despite its concern that the M'Naghten standard was inadequate, the court had been unwilling to invade the province of the legislature and set up a judicial test of insanity; its rulings from *Gorshen* going forward avoided any need to do so. The ALI insanity standard differs from M'Naghten in two ways: It provides a volitional as well as a cognitive test of culpability, and its advisedly ambiguous use of the terms *substantial* and *appreciate* grants considerable discretion in determining whether a particular defendant was insane. Under the ALI standard, *Gorshen*, Wolff, Poddar, and even Conley might have been adjudicated not guilty by reason of insanity, obviating the need for tortured redefinitions of *mens rea* requirements.

With the *Drew* decision, the continued relevance of the diminished capacity ameliorative approach was called into question (Waddell, 1979). But that by itself did not bring about the demise of the approach. It was a defense that had outlived its original mission and had taken on a life of its own. Then, with the Dan White murder trial, the diminished capacity defense went from being well known to being notorious and deeply troubling to many in the legal community and public. It cannot be a coincidence that this reaction took place during a period of general questioning of the expanded use of the insanity defense itself, questioning that, following the Hinckley trial in 1982, culminated in an unprecedented wave of insanity reform legislation in the U.S. Congress and in state legislatures around the country.

Daniel White was 32 years old when on November 27, 1978, he shot and killed San Francisco mayor George Moscone and a city supervisor, Harvey Milk, a champion of gay rights. As later reported by a defense psychiatric expert (Blinder, 1981–1982), White was one of eight children of a firefighter and was himself a firefighter and a former police officer when he decided to run for one of the city supervisor positions in 1977. After a hard-fought campaign, he was elected to represent his working-class district but was stymied by personal financial problems. After resigning in early November 1978, he reconsidered and asked Mayor Moscone to reappoint him to the supervisor post. Reportedly, the mayor first promised White his support and then withdrew it. On the morning of November 27, White went to City Hall with a loaded handgun and 10 extra rounds of ammunition. He avoided security personnel and metal detectors by entering City Hall through a ground-floor window. He confronted the mayor and shot him four times before reloading his gun. He then encountered Harvey Milk, whom he believed was involved in the mayor's decision not to reappoint him. He shot Milk five times, then left the building, retreated to a church, and ultimately surrendered to the police.

Many questioned the remarkable outcome of Dan White's murder trial and how it could be that the killing of two prominent political figures resulted in a jury verdict of only voluntary manslaughter. There were suggestions (Szasz, 1981–1982) of a politically motivated collusion between the defense and the surprisingly ineffectual prosecution as well as the influence of antihomosexual prejudice. In any event, the ostensible justification for the jury finding seemed to be provided by expert testimony on diminished capacity.

The press seized on a point in the trial when a defense expert, Dr. Martin Blinder, testified that White was a manic-depressive whose depressive episodes, one of which he said led to the homicides, were exacerbated by his bingeing on “junk food—Twinkies, cupcakes, and Cokes” (Szasz, 1981–1982). Reporters immediately ridiculed diminished capacity as the “Twinkie defense,” although the actual effect of that part of the testimony on the jury verdict is uncertain. The jury heard from several defense experts that White had been incapable of forming the intent elements required for first-degree murder or even second-degree murder. Despite evidence of preparations by White to commit murder, Dr. Donald Lunde, echoing the language of the *Wolff* decision 15 years earlier, testified that not only had White not premeditated and deliberated the killings, but, because of his mental condition—severe depression and a compulsive personality—he was not capable of any sort of mature and meaningful reflection (Szasz, 1981–1982). Blinder asserted that premeditation and deliberation both require “reasonably clear thinking” and that, at the time of the offense, White “no longer had his wits about him” (Blinder, 1981–1982).

As a footnote to this case, Dan White was sentenced to the maximum term of confinement for voluntary manslaughter and was released on parole when he became eligible in 1984. He quietly returned to San Francisco, where he committed suicide in October 1985.

INSANITY DEFENSE REFORM AND THE FATE OF DIMINISHED CAPACITY

The controversy raised by the Dan White trial contributed to efforts to abolish the diminished capacity defense in California. The verdict was widely viewed as an outrage—not only by the gay community, which erupted in vocal protest and riot—and an offense to common sense. Closely contemporaneous with the White trial, the California Supreme Court, in *People v. Wetmore* (1978), had conceded that diminished capacity could conceivably result in outright acquittal when the crime charged, such as Wetmore’s burglary, permitted no conviction for any lesser included felony. The court in *Wetmore* indicated the need for legislative clarification of diminished capacity procedures (Morse & Cohen, 1982). What the legislature ultimately delivered was nothing short of a repudiation of the entire diminished capacity approach pioneered by the court in the previous three decades.

California Senate Bill 54 (1981a) reversed judicial redefinitions of intent elements and in effect codified the older, once-superseded traditional forms of premeditation, deliberation, and malice. The legislation provided that it would no longer be necessary to prove that a defendant “maturely and meaningfully” reflected on the depravity of an act to find that a killing was done with premeditation and deliberation. It further provided that it would not be necessary to a finding of malice to determine that the defendant was aware of an obligation to act within the body of laws governing society.

In a related development, California Senate Bill 590 (1981b) aimed to restrict the scope of expert testimony in criminal cases by providing that psychiatrists and psychologists would not be presumed to be able to determine sanity or insanity. Finally, a ballot initiative, Proposition 8, entitled "The Victims' Bill of Rights," which was approved by California voters in 1982, "abolished" diminished capacity as a defense at trial (Kraus, 1983) and replaced the California Supreme Court-imposed ALI-based insanity test with a modified but more restrictive M'Naghten test (Steadman et al., 1993).

The Hinckley trial, which resulted in the acquittal by reason of insanity of a would-be assassin of a U.S. president from California, may well be the defining moment of the last century's treatment of mental disorder at the time of an offense. It is uncannily similar in its impact to Daniel M'Naghten's acquittal 140 years earlier, after his attempt to assassinate the British prime minister. In both instances, an acquittal at trial resulted in public outrage and restrictive reform. The Dan White trial, in a smaller way, was a defining moment in the brief and more parochial history of diminished capacity in California. It seemed in many ways to exemplify all that was wrong with the overly broad consideration of mental abnormality in adjudicating guilt for criminal acts. Still, it is apparent that, at least with respect to diminished capacity and reform measures, larger social forces were at work. Efforts to rescind what was seen as an overly generous treatment of mentally disordered offenders in California were already under way by the time White committed his homicides, just as similar efforts were under way in other states before Hinckley shot President Reagan.

A get-tough attitude in California had already yielded changes in sentencing procedures and provisions for the commitment of insanity acquittees by 1978, and these had begun to yield results. Indeed, a downturn in the number of insanity pleas in California preceded the return of the M'Naghten standard to that state and reflected the changes already implemented that made insanity a less attractive defense option (Steadman et al., 1993). That the shift in California from the ALI to the M'Naghten insanity standard did not produce any further demonstrable effects in either the rate of insanity pleas or the rate of successful insanity pleas (Steadman et al., 1993) suggests that any notion that the strictures of M'Naghten needed to be ameliorated, either by an expanded diminished capacity option or a liberalized insanity test, may have been mistaken in the first place.

In any event, diminished capacity was tarred with the same brush as was insanity during the wave of reform legislation that took place in the late 1970s and early 1980s. Aside from the decisive repudiation in California of its signature approach to diminished capacity, what may have been the strongest blow to an expanded treatment of diminished capacity as diminished responsibility was delivered by the federal Insanity Defense Reform Act (IDRA; 1984), which followed the Hinckley case. Along with its elimination of the volitional prong of the ALI insanity test then in use in federal courts, its shifting of the burden of proof from prosecution to defense, and its requirement that insanity be proven by clear and convincing

evidence, the IDRA sought to eliminate the diminished capacity defense. After delineating the conditions under which a severe mental disease or defect could result in the affirmative defense of insanity, Congress determined that “mental disease or defect does not otherwise constitute a defense.”

In conjunction with the passage of the IDRA in 1984, the Federal Rules of Evidence (FRE) were revised to restrict the scope of expert testimony in criminal proceedings. The revised rule stated: “No expert witness testifying with respect to the mental state or condition of a defendant in a criminal case may state an opinion or inference as to whether the defendant did or did not have the mental state or condition constituting an element of the crime charged or of a defense thereto.” Rather, “such ultimate issues are matters for the trier of fact alone” (FRE 704(b)). This restriction was intended to limit testimony by expert witnesses to the presentation and explanation of their diagnoses, such as whether the defendant had a severe mental disease or defect and what the characteristics of such a condition may have been (Senate Report No. 225, 1984).

Given subsequent federal court rulings, it appears that Congress actually failed to do more than eliminate the possibility of an affirmative California-type diminished capacity defense. As Morse (1984) pointed out at the time, the *mens rea* diminished capacity partial defense, as opposed to the affirmative California defense, is constitutionally protected from legislative reform. But if Congress did nothing more than that, it may have done enough. The expanded diminished capacity defense, already sinking under its own weight in California, was clearly precluded in the federal jurisdictions, even while federal appeals courts in most cases affirmed Morse’s view that a strict *mens rea* inquiry, and the defendant’s right to dispute prosecution’s claim that an act was done with criminal intent, could never be eliminated.

In two instances (*U.S. v. White*, 1985; *U.S. v. Hood*, 1988), federal circuit courts of appeal, the First and the Fourth, interpreted IDRA as precluding specific mental health testimony directed solely at negating an element of intent and establishing that a defendant lacked requisite *mens rea* rather than that the defendant was insane. All of the district and appellate courts that considered the matter, however, drew the distinction between a strict *mens rea* approach to diminished capacity and the affirmative California diminished-capacity-as-diminished-responsibility approach, which all federal courts regarded as having been abolished by IDRA. All of them indicated that IDRA did not establish a general rule that evidence of a mental disorder is always inadmissible except in relation to insanity.

In *U.S. v. Frisbee* (1985), a U.S. district court in California ruled that a defendant could submit expert testimony in support of his contention that he did not have the specific intent required for first-degree murder. The defendant reportedly was a chronic alcohol abuser who suffered from periodic blackouts or seizures and from amnesia. However, the court also held that the experts could not opine as to whether the defendant did or did not form the specific intent to kill. The court reasoned that the IDRA provision that mental disease or defect does not constitute a defense

apart from insanity was not intended to limit the admissibility of evidence negating specific intent. Rather, it held that this provision was aimed at eliminating any affirmative defense other than insanity in which mental abnormality is offered in exculpation, as an excuse for an offense. The court drew the distinction between the defense of diminished capacity, which is properly aimed at negating the presence of intent elements, and a diminished responsibility approach, which does not negate elements or deny guilt, but instead simply mitigates guilt.

Citing *Frisbee* and the Senate Report on the change of the FRE, the District of Columbia District Court, in *U.S. v. Gold* (1987), found that the reform measures did not preclude defense-offered testimony on the capacity of the defendant to formulate specific intent. In a similar way, the Third Circuit Court of Appeals, in *U.S. v. Pohlott* (1987), overturned a district court ruling that prevented a defendant from introducing any evidence of mental abnormality. It held that both the wording and the legislative history of IDRA “leave no doubt that Congress intended . . . to bar only alternative ‘affirmative defenses’ that ‘excuse’ misconduct but not evidence that disproves an element of the crime itself” (827 F.2d 889) and that admitting mental health evidence to negate *mens rea* does not constitute a defense but could only negate an element of the offense and is therefore not barred.

The court drew the distinction between diminished capacity—focused on the presence of intent elements—and diminished responsibility, which it identified with the California case law. It treated *mens rea* as an element that is generally established satisfactorily by any showing of purposeful activity regardless of the activity’s psychological origins. By contrast, the court held that testimony hoping to explain the behavior might mislead a jury. The *Pohlott* court, while affirming the right of the defendant to present expert testimony on the question of *mens rea*—testimony on both the presence or absence of intent and the defendant’s capacity to formulate intent—barred testimony concerning the defendant’s unconscious motivation in attempting to hire someone to murder his wife. The court regarded that testimony as not strictly pertinent to the question of *mens rea* but rather to the question of whether the intent was formed. Regarding evidence bearing on how meaningful the defendant’s understanding was of his or her actions and their consequences, the court wrote, “We often act intending to accomplish the immediate goal of our activity, while not fully appreciating the consequences of our acts. But purposeful activity is all the law requires” (*U.S. v. Pohlott*, 1987).

The Ninth Circuit Court of Appeals reached a similar conclusion the following year in *U.S. v. Twine* (1988). Agreeing with the courts in *Frisbee*, *Gold*, and *Pohlott*, it held that IDRA did not abolish the diminished capacity defense as such. However, after finding that the district court judge had considered and was unpersuaded by the defendant’s diminished capacity defense, it affirmed Twine’s conviction for making telephone and mail threats to kidnap and injure. Similarly, the Eighth Circuit, in *U.S. v. Bartlett* (1988), held that IDRA did not render inadmissible expert testimony in support of the contention that the defendant was incapable of forming the requisite intent.

In 1989, the Ninth Circuit, following its decision the previous year in *Twine*, ruled in *U.S. v. Brown* that evidence of mental defect offered to show that the defendant lacked the specific intent to commit first-degree murder could not be excluded on the basis of IDRA. The Sixth Circuit in *U.S. v. Newman* (1989) ruled similarly in a case focusing on a defendant's claim that alcoholism produced a chronic brain syndrome that precluded the formulation of requisite *mens rea* for a crime of transporting stolen property. While upholding the defendant's conviction, the court held that both insanity and diminished capacity were permissible defenses under IDRA. The following year, the Eleventh Circuit Court of Appeals made the same finding in vacating a conviction for drug offenses in *U.S. v. Cameron* (1990). It held that, despite a "semantic war of labels" (907 F.2d 1051), both Congress and the courts had recognized a distinction between evidence of psychological impairment that supports an affirmative defense and evidence that simply negates an element of the offense charged. Testimony that helps the trier of fact determine the defendant's specific state of mind with regard to actions at the time the alleged offense was committed does not constitute an affirmative defense, the court held, but goes instead to whether the prosecution has met its burden of proving all the essential elements of the crime.

Continuing this chain of findings, the District of Columbia Circuit Court reached a similar conclusion in *U.S. v. Childress* (1995), a case in which a psychologist had been prevented from testifying at trial that the defendant, facing drug conspiracy charges, was in essence mentally retarded. The court held that, although the trial court would need to determine whether the intended expert testimony was sufficiently grounded in science to warrant its use in the courtroom, psychological testimony not intended to establish insanity would be admissible, providing it was aimed not at making an insanity-like excuse for illegal conduct but at addressing whether the defendant had entertained the specific intent alleged, namely, in this case, conspiratorial understanding and purpose.

In 1997, the First Circuit Court of Appeals revisited this matter in *U.S. v. Schneider* and reached the same conclusion as other circuit courts had. The defendant in that case had been convicted of multiple counts of mail and wire fraud. Testimony was rejected from a medical doctor and a psychiatrist that would have presented the defendant's judgment as having been impaired by misprescribed and overprescribed medications, chemical dependency, depression, and probably mania. The proposed testimony, which would not have gone to demonstrate insanity, was rejected by the trial judge as misleading and irrelevant to the task of negating intent. The First Circuit agreed with the trial court that the testimony would have been of only limited relevance and could easily mislead the jury into thinking that the defendant's condition amounted to insanity or lessened the offense. The court held that a defendant seeking to present evidence of a mental condition short of legal insanity must show that the evidence is relevant to the requisite intent that is at issue, that its probative value is not substantially outweighed by confusion or delay,

and that, if the evidence comes in by way of expert testimony, it is scientifically reliable and helpful to the jury.

As indicated by *Childress* and *Schneider*, it is one thing for the courts to acknowledge in principle the legal viability of expert testimony on the question of whether the defendant was able to form specific intent or actually did so and something else to find a sound scientific basis for any such testimony in the case at hand. Indeed, a consideration of actual cases indicates that it is difficult to marshal the psychological evidence needed to support an opinion that the defendant had diminished capacity. As Morse pointed out (1984), if the *mens rea* approach is applied strictly and is not simply the vehicle for a proposed excuse, it would be unlikely to be of any benefit to the defendant who seeks to use it. If the assumption of sanity is hard to rebut, it is far more difficult to establish that the intent for a crime that appears to have been committed could not have been formed or that, if it could have been formed, it actually was not.

It may have been some sense of the great practical difficulty associated with a strictly applied *mens rea* defense that encouraged a handful of states—to date five—in the years surrounding the Hinckley verdict and IDRA to substitute it for insanity, a defense they essentially abolished. Although each of these states ostensibly retained some opportunity for a defendant to introduce mental health testimony, if strictly applied, those provisions could be expected to avail the defendant little, even in cases of gross mental disorder or disability. The United States Supreme Court ruled that the Constitution does not require that an insanity defense be made available (see *Powell v. Texas*, 1968; *Clark v. Arizona*, 2006), but a substitute defense focusing on *mens rea* elements may be criticized as not fairly responding to the influence that mental disorder actually has on criminal responsibility (Morse, 1999).

Montana was the first state to abolish the insanity defense, in 1979, providing instead for consideration by the court of “whether the defendant had, at the time that the offense was committed, a particular state of mind that is an essential element of the offense” and that a defendant’s inability “to appreciate the criminality of . . . behavior or to conform the . . . behavior to the requirements of the law” would be considered only at sentencing (Montana Code §46-14-101).

Idaho abolished its insanity test in 1982, declaring that a “mental condition shall not be a defense to any charge of criminal conduct” while providing that “nothing herein is intended to prevent the admission of expert evidence on the issue of any state of mind which is an element of the offense” (Idaho Statutes §18-207).

In 1983, Utah abolished its insanity defense, providing only that “it is a defense to a prosecution . . . that the defendant, as a result of mental illness, lacked the mental state required as an element of the offense charged” and that “mental illness is not otherwise a defense, but may be evidence of special mitigation reducing the level of a criminal homicide or attempted criminal homicide” (Utah Code §76-2-305).

In 1995, Nevada abolished the insanity defense and substituted a plea of “guilty but mentally ill,” with a provision only that insanity or intoxication could be taken

into consideration whenever purpose, motive, or intent is a necessary element of a criminal offense (Nevada Revised Statutes, since amended, §§174.035 and 193.220).

Though not a mountain state like these four, Kansas also abolished insanity in 1995, providing instead that the only mental state defense would be that “the defendant, as a result of mental disease or defect, lacked the mental state required as an element of the offense charged” (Kansas Statutes §22-321).

Concern has been voiced by some in the legal community regarding the constitutionality of the abolition of the insanity defense in these states and the impracticality and injustice of what amounts to a substitution of a strict *mens rea* defense (see, e.g., Gardner, 1993; LeBlanc, 2007; Nevins-Saunders, 2012; Phillips & Woodman, 2008; Stimpson, 1993). However, with one exception—in Nevada (*Finger v. State*, 2001)—the abolitionist state statutes have been upheld on appeal: in Montana (*State v. Korell*, 1984); in Idaho (*State v. Searcy*, 1990; *State v. Delling*, 2011); in Utah (*State v. Herrera*, 1995); and in Kansas (*State v. Jorrick*, 2000; *State v. Albright*, 2002; *State v. Bethel*, 2003). Especially after the decision of the U.S. Supreme Court in *Clark v. Arizona* (2006), which affirmed that Arizona’s truncation of the M’Naghten insanity standard was constitutional, there appeared to be no reason for critics of the insanity abolition/*mens rea* defense substitution, in the four states that have enacted and upheld it, to expect federal relief. This was further confirmed in 2012 (*Delling v. Idaho*, cert. denied).

To summarize, the general trend in the United States in the last quarter of the 20th century of restricting the use of mental state evidence in criminal cases—with respect to both diminished capacity and insanity—led to the curtailment of testimony that serves to explain an offender’s mental state, leaving intact in some places, as the only option, a nominal ability to introduce evidence to negate intent. In none of the statutory changes or case law concerning the newly restricted diminished capacity or *mens rea* approach, however, is there any guidance for the forensic examiner confronted with these questions—or for the judges and juries that must determine guilt or innocence.

FORENSIC EVALUATIONS OF DIMINISHED CAPACITY AND *MENS REA*

To an extent exceeding cases in which simple insanity is at issue, the forensic examiner who is asked to offer an opinion about a defendant’s capacity to formulate the specific intent for an alleged offense is hard put to offer anything that is truly relevant and material. Examiners could explore how defendants’ mental abnormality—if there is any, or if any can plausibly be raised—affected the defendants’ awareness of what they were doing; their legal, ethical, and moral evaluation of it; their anticipation of its effects or consequences; and the restricted range of alternatives they had to behaving as they did. Examiners can, in other words, use the diminished capacity issue as an analog of the insanity defense, substituting some relatively minor mental abnormality for the severe mental disease or defect required for insanity and thus making a case for exculpation by explanation.

If no settled mental abnormality is apparent, the expert could conceivably advert to intoxication or even to “stress,” whatever that may be. None of these possibilities is recommended. It is apparent from the history of diminished capacity and *mens rea* law that such approaches to demonstrating an inability to form requisite intent are now strictly irrelevant to the issues as the law has defined them in many jurisdictions. Since the death of diminished-capacity-as-diminished-responsibility in California, there appears to be no jurisdiction that specifically invites an expert to hazard a guess as to what psychological elements true criminal intent ought to involve or what character flaws, attitudes, and emotional or chemical states could spoil what would otherwise constitute requisite *mens rea*.

Those forensic examiners who feel obliged to confine themselves to speaking of intent strictly in the terms laid down by the law usually will find themselves with little to say. In large part, this is because the question of the capacity for intent is only theoretically, and not actually, separate from the question of whether intent was formed. Consideration of psychological factors ordinarily cannot lead to a conclusion that a defendant lacked the capacity to formulate intent for an offense the defendant appears to have committed. In many instances, unless the expert undertakes to redefine what “real” specific intent is—precisely what the California Supreme Court did in holding, in *Wolff*, that premeditation and deliberation in murder must involve “mature and meaningful” reflection on the enormity of the offense contemplated—there is not much left to be said.

There is a fundamental difference between a question of whether intent was formed or could have been formed, on one hand, and a question of the quality and characteristics of whatever intent was formed, on the other. The law as it has developed is concerned at most only with the first question, and the expert attempting to address that question is not particularly helped by exploring the second question. The discovery that a mother was depressed and that the intent she formed to smother her child and relieve him or her of what she viewed as unbearable sorrows may explain how and why the killing took place, but it does not answer the question of whether she premeditated and deliberated a murder. From the point of view of the law’s minimal and wholly pragmatic construction of intent, the answer to that question—one solely reserved for the trier of fact in any case—will not require any sophisticated analysis of the defendant’s psychology or psychopathology. In most instances, if the perpetrator engaged in planning or preparation, especially if he or she announced an intention to kill, it will be impossible from a strict behavioral science viewpoint to refute a prosecution assertion that the person not only had the capacity to premeditate and deliberate a murder but actually did so.

By comparison, in the usual case, competent opinion testimony on the issue of insanity is much more possible, and plausible. It is easier to demonstrate that, as a result of mental disease or defect, a defendant lacked the capacity to appreciate the wrongfulness of the conduct, for instance, than it is to demonstrate that a mental disorder, however severe it is, prevented the defendant from doing precisely what he or she seems to have done. Insanity does not invite the expert witness or anyone

else to determine what offense was committed; insanity is an affirmative defense in which an exculpatory excuse is offered for an offense that is not itself disputed. Yet this is what a positive opinion on diminished capacity implies: If the defendant could not form the requisite intent elements, those intent elements were not formed, and whatever offense was committed could not have been the specific intent crime charged.

The problem can be illustrated by a case similar to that of *Conley* and involving a 40-year-old woman charged with two counts of first-degree murder. She was an alcoholic who had been divorced but never entirely separated from her alcoholic husband who owned a bar. For nearly 24 hours before the slayings, she had been drinking and talking with her former husband and his girlfriend in his tavern. Talk turned to the daughter of the defendant and her ex-husband, the defendant's fitness as a mother, and whether it was right that she should retain custody of the child. Drunk and angry, the defendant drove home and later returned to the bar with a loaded shotgun in the trunk of her car. Once back inside the bar, she resumed her hot exchange with her husband's girlfriend, during which she threatened the woman's life. The defendant again left the bar, this time telling a barmaid on her way out to "hit the floor" when she came back in and telling her that she would not like what she was going to do. The defendant returned to the locked door of the bar with the shotgun, and, when her husband and his girlfriend opened the door for her, she fatally shot them in rapid succession. She then drove back to her apartment, hid the shotgun, told her boyfriend to deny that she had gone out, and passed out on her bed.

It might fairly be asked whether any crime would have been committed but for the defendant's drunkenness, and even whether she would have been drunk had she not been an alcoholic. But these are not the questions that must be answered with respect to the diminished capacity question that was raised at trial. It may be fair for the jury to consider whether the defendant was guilty of first-degree murder, specifically whether anyone as thoroughly intoxicated as she was should be held to have engaged in true premeditation and deliberation, or whether her thinking, even if infused with alcohol, may be taken to be actual premeditation and deliberation. The jury in this instance agreed with the prosecutor that they were, and the defendant was convicted of first-degree murder.

So what is the responsible expert witness left to say when diminished capacity is raised? There are several options, depending on the circumstances of the offenses and the findings on examination.

TRUE DIMINISHED CAPACITY?

Although genuine cases of diminished capacity should not be expected to occur often, some individuals lack the capacity to form requisite intent and therefore must not have formed it. Because larceny, a specific intent offense, necessitates an understanding of the concept of property, some individuals are so developmentally

disabled or otherwise limited cognitively that they are demonstrably incapable of larceny. Such a person walking off with some item from a home or a store, for instance, may not have formed any intent to commit larceny, simply because of an inability to comprehend the basic notion of property. Similarly, it may be possible to provide competent expert testimony that an individual was so affected by delirium or dementia as to be incapable of planning his or her next step, much less premeditating and deliberating a murder.

Although they are theoretically possible, there probably are reasons why it is rare to encounter these types of diminished capacity cases in practice. Substantially developmentally delayed individuals who take things may be unlikely to be charged with larceny. If they are, someone in a position to gate them into either the mental health or the criminal justice systems, including the police, prosecutor, or arraigning magistrate, is likely to have formed the impression that they were not so intellectually impaired in the first place. And those who most clearly have diminished capacity because of gross cognitive impairment are as unlikely to commit specific intent offenses as they are to engage in any other sophisticated, purposeful, and goal-directed behavior.

It is possible that in other more subtle ways a defendant may evidence diminished capacity in the sense of having been incapable of formulating specific intent. A man with a history of epilepsy was charged with two counts of involuntary manslaughter with a motor vehicle and with failure to stop at a personal injury accident. Historically, his seizure disorder had never been fully controlled by medication, and he was restricted from driving. Despite this, he acknowledged that he sometimes drove his mother's car. On one prior occasion he had a seizure while driving and ran the car into a parked car at low speed; that collision did not cause any injuries and he was not prosecuted. On the second occasion, however, the defendant, who reported no memory of the accident, was driving over the speed limit in a residential area when he drove over the curb and onto the sidewalk and struck a sign before running into a group of people. The collision killed two young children and injured their mother and grandfather. After the collision, the car driven by the defendant continued on the sidewalk until stopped by a tree. Witnesses reported that, once the car was stopped, the rear wheels of the vehicle continued to spin; apparently the driver was continuing to depress the accelerator. A witness related that the defendant continued to sit behind the wheel with a dazed appearance and that he was not responsive to what the angry and excited witness shouted at him, asking him if he knew what he had just done. That witness and others ran to the aid of the victims who were lying injured some distance back down the street. While the victims were being attended to, the defendant got out of his car, looked at the damage to the front end, got back in, and drove away.

Diminished capacity was raised in this man's defense. There was, aside from anything else, a question of whether he had been conscious of what he had done at the time he left the scene of the accident. Leaving the scene of a personal injury accident was viewed by the prosecution as a specific intent offense, as it required

knowledge by the individual that he had caused personal injury to others. It appeared likely in that case that the defendant had one of his characteristic seizures at the time of the accident. From the vivid reports of witnesses, particularly one man who observed him just after the fatal collision, the defendant may have been experiencing postictal clouding of consciousness prior to driving off. It is plausible that he did not comprehend the witness who demanded to know if he understood what he had just done; he was observed to be dazed and unresponsive. Postictal confusion may have prevented him from understanding what had just happened. He would not have remembered it if he had been having a seizure at the moment of the collision, and, before he drove away, he did not come back or even look toward the crowd gathered around the people he had struck down some distance behind him. Of course, this says nothing about his guilt for the more serious charges of involuntary manslaughter that he also faced; those are general intent offenses for which his epilepsy was, if anything, an aggravating factor, as he drove a vehicle when he knew he was restricted from doing so and had even had a prior accident also brought on by a seizure.

ACTUALITY OF INTENT

In some instances, it is possible to offer testimony that pertains to the actuality of intent rather than the capacity for intent, which may be assumed or not in dispute. These are cases in which consideration of a defendant's psychological functioning and the circumstances of the alleged offense suggest a plausible factual alternative to the crime charged.

A case in point involved a man without any criminal history who had been charged with breaking and entering a department store with intent to commit larceny. He claimed that he had entered the store late at night solely to find some anticonvulsant medication that he was afraid he had dropped there the day before. On examination he proved to be rigidly obsessive and hypochondriacally preoccupied; he also had a history of mild neurological impairment. Over the years, he had persistently worried about the health of his daughter and of children in general. The day before the break-in, he claimed, he had been in the department store cafeteria and had dropped a vial containing his pills. He thought he had recovered all of the pills, but he later worried that some were missing. He especially worried, he said, that some child would come into the cafeteria, pick up the pills, and swallow them. He resolved to get back into the store and search for the pills. He called the now-closed store, trying without success to talk with security officers. He assumed that, if he called the police, they would simply tell him to wait until morning, but he worried that he might be too late if he waited until morning. Finally, with what he described as mounting anxiety, he forced a door open. A silent alarm brought the police, who found him in the store and arrested him. He had no store merchandise with him. Police reported only that an ice machine in the cafeteria had been moved aside.

In this case, it is not likely, despite the defendant's history and odd presentation, that he could be said to be incapable of larceny, unless that term were reinterpreted to mean something more than it ordinarily does. The real question in this case was not whether the defendant could have formed the intent to commit larceny but whether he actually did so. Because the *mens rea* question cast in this light necessarily involves a judgment of what occurred in fact, in the final analysis it must be answered by the trier of fact. However, this is a case in which the results of a psychological evaluation lent plausibility to an alternative explanation of the behavior at issue, one that, if true, would mean that, even if the defendant had entered the store illegally and could be subject to prosecution for that, he was not guilty of breaking and entering with intent to commit larceny.

It is uncertain to what extent expert testimony bearing on the actuality of intent rather than capacity to form intent would be admissible. In Michigan, for example, when that state was still permitting diminished capacity to be raised at trial (the state supreme court, in *People v. Carpenter*, effectively abolished it in 2001), appellate decisions indicated that the only question to be considered was whether a defendant was incapable of forming specific intent, not whether he or she did so or not (*People v. Savoie*, 1984). That issue is unclear in federal courts. For example, the Third Circuit in *Pohlot* (1987), in noting that "evidence of mental abnormality may help indicate lack of *mens rea* even when a defendant is legally sane" (827 F.2d 900), appeared to suggest that such testimony would be acceptable. However, FRE Rule 704(b) appears to preclude such testimony in declaring that "[n]o expert witness . . . may state an opinion or inference as to whether the defendant did or did not have the mental state or condition constituting an element of the crime charged or of a defense thereto." The rule identifies these as "ultimate issues" that are matters for the judge or jury alone.

The forensic examiner faced with this issue in an actual case must be aware of the admissibility in that jurisdiction of opinion testimony concerning the actuality of intent. It is not recommended in any case that the expert offer an opinion on the ultimate issue and testify in effect that the defendant is or is not guilty. Opinions of this type exceed the scope of other expert opinions, such as whether a defendant is competent to stand trial or meets the insanity criteria, and should be objectionable to expert witnesses if not to courts and lawyers. However, there are occasions when the contributions an expert can make to the deliberations by judge or jury may be valuable.

NO OPINION OR NEGATIVE OPINION: THE COMMON CASE?

A young man was arrested for armed robbery in the theft of narcotics from a drugstore. There was evidence that he was a narcotics addict who had run out of both drugs and money and who was in great need. Defense counsel also cited, in support of a claim that because of his addiction he was not forming the intent for armed robbery, that the man had been under stress from conflict with his wife, including an argument the very morning of the robbery.

In a great many cases in which diminished capacity is at issue, it is not reasonable for the expert to offer testimony other than that there is no good or plausible basis for concluding that the defendant was not capable of forming intent; that the defendant was capable of engaging in conscious, goal-directed behavior; and even that the behavior alleged is not consistent with loss of consciousness or inability to make and execute plans. In this instance, if the man was not capable of forming the intent to commit armed robbery—a combination of felonious assault and larceny—it is entirely unclear what he was doing in the drugstore with a gun. If that was not an armed robber, who was that masked man? His addiction and his need for drugs serve only to explain why he might have formed the intent to commit a robbery; they do not negate the intent or make it something less than the requisite intent for this specific intent crime. The stress he cited does not do even that much. The defendant in this case provided the examiner with little to work with in terms of identifying a basis for a positive opinion on diminished capacity.

This case is an example of the sort of situation in which a positive opinion on diminished capacity does not seem possible. Most cases in which a defendant appears to meet the criteria to be considered legally insane, in those jurisdictions that retain an insanity defense, also fall into this category. Because of the different approaches the two defenses take to the issue of *mens rea*, it typically is easier to demonstrate, for example, that a defendant met the insanity test by lacking substantial capacity to appreciate the wrongfulness of a criminal act than that the defendant could not have formed specific intent. A mentally disordered defendant who committed a homicide for delusional reasons—for instance, believing that the victim intended to kill him or that God ordered him to kill the victim, or that not to do so would bring about some delusional calamity—ordinarily is certainly capable of premeditating and deliberating the killing and will not dispute that he formed the intent to kill in just this way.

WHAT DO LAWYERS REALLY WANT?

In some cases in which diminished capacity is raised, defense counsel may not be interested in eliciting an opinion on capacity for intent so much as a description or explanation of the defendant's conduct to present to the judge or jury. If that is the point, the forensic examiner may have something to contribute. The danger here lies in the possibility that the expert will conflate explanation and exculpation and testify that, because the defendant's intentions and conduct may be understood in terms of his or her psychology, they do not meet the criteria for requisite criminal intent. This sort of testimony would recapitulate the diminished-capacity-as-diminished-responsibility approach taken in California prior to retrenchment there, and it should be understood as such.

As long as the law defines intent as it does, as a skeletal cognitive and pragmatic affair, testimony of this sort may be misleading to a jury even if courts permit it. Psychologists and psychiatrists are not in a position to define what requisite legal

intent must involve or even what sorts of mental or emotional disorders preclude it. Inevitably, expert testimony that suggests otherwise involves insupportable claims by the expert, for example, a claim that, although it appears that a defendant premeditated and deliberated a murder—having actually thought about it, planned it, and even announced his intentions—he actually could not have formed such intent because of mental abnormality. Psychologists and psychiatrists have a great deal to offer in terms of delineating data within their purview concerning defendants and their behavior. However, they are not capable of transforming what the law considers requisite intent into something less than that because of unsupported assumptions about the mental factors that must be involved in a genuine criminal act.

CONCLUSION

The history of diminished capacity and *mens rea* conceptualizations indicate a potential for contributions by forensic psychologists and psychiatrists to the resolution of questions of intent. However, conceptualizations of intent embodied in the law's approach to criminal liability sharply limit the extent to which mental health expert testimony can ever bear on the ultimate issue to be determined by the judge or jury.

REFERENCES

- American Law Institute. (1962). Model Penal Code [Proposed official draft].
- Arenella, P. (1977). The diminished capacity and diminished responsibility defenses: Two children of a doomed marriage. *Columbia Law Review*, 77, 827–865.
- Barratt, E. S., & Felthous, A. R. (2003). Impulsive versus premeditated aggression: Implications for *mens rea* decisions. *Behavioral Sciences and the Law*, 21, 619–630.
- Black, H. C. (1979). *Black's law dictionary* (5th ed.). St. Paul, MN: West.
- Blinder, M. (1981–1982). My psychiatric examination of Dan White. *American Journal of Forensic Psychiatry*, 2, 12–27.
- Bonnie, R., & Slobogin, C. (1980). The role of mental health professionals in the criminal justice process: The case for informed speculation. *Virginia Law Review*, 43, 427–522.
- California Senate Bill No. 54. (1981a).
- California Senate Bill No. 590. (1981b).
- Clark, C. R. (1982). Clinical limits of expert testimony on diminished capacity. *International Journal of Law and Psychiatry*, 5, 155–170.
- Clark, C. R. (1988). *Diminished capacity in Michigan: Factors associated with forensic evaluation referrals*. Paper presented at American Psychology-Law Society conference, Miami Beach, FL.
- Clark v. Arizona, 548 U.S. 735 (2006).
- Denno, D. W. (2003). A mind to blame: New views on involuntary acts. *Behavioral Sciences and the Law*, 21, 601–618.
- Delling v. Idaho, 133 S.Ct. 504 (2012), *cert. denied*.
- Diamond, B. (1961). Criminal responsibility of the mentally ill. *Stanford Law Review*, 14, 59–86.

- Dix, G. E. (1971). Psychological abnormality as a factor in grading criminal liability: Diminished capacity, diminished responsibility, and the like. *Journal of Criminal Law, Criminology, and Police Science*, 62, 313–334.
- Erlinder, C. P. (1983). Post-traumatic stress disorder, Vietnam veterans and the law: A challenge to effective representation. *Behavioral Sciences and the Law*, 1, 25–50.
- Federal Rules of Evidence, FRE 704(b), amended by Comprehensive Crime Control Act of 1984, Public Law No 98-473.
- Finger v. State, 117 Nev. 548, 27 P.3d 66 (2001).
- Gardner, M. (1993) The *mens rea* enigma: Observations on the role of motive in the criminal law past and present. *1993 Utah Law Review*, 635–750.
- Idaho Statutes §18-207 (1982).
- In re Winship*, 397 U.S. 358 (1970).
- Insanity Defense Reform Act of 1984, Public Law 98-473, 18 U.S.C. § 17.
- Kansas Statutes §22-321 (1995).
- Kraus, F. R. (1983). The relevance of innocence: Proposition 8 and the diminished capacity defense. *California Law Review*, 71, 1197–1215.
- LaFave, W., & Scott, A., Jr. (1972). *Handbook of criminal law*. St. Paul, MN: West.
- LeBlanc, S. M. (2007). Cruelty to the mentally ill: An Eighth Amendment challenge to the abolition of the insanity defense. *American University Law Review*, 56, 1281–1328.
- Malle, B. F., & Nelson, S. E. (2003). Judging mens rea: The tension between folk concepts and legal concepts of intentionality. *Behavioral Sciences and the Law*, 21, 563–580.
- Marx, M. L. (1977). Prison conditions and diminished capacity—A proposed defense. *Santa Clara Law Reporter*, 17, 855–883.
- Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C., with Lyons, P. M. & Otto, R. K. (2007). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (3rd ed.). New York, NY: Guilford Press.
- M’Naghten’s Case, 10 Cl. & F. 200, 8 Eng. Rep. 718 (H.L. 1843).
- Montana Code §46-14-101 (1979).
- Morse, S. J. (1979). Diminished capacity: A moral and legal conundrum. *International Journal of Law and Psychiatry*, 2, 271–298.
- Morse, S. J. (1984). Undiminished confusion in diminished capacity. *Journal of Criminal Law and Criminology*, 75, 1–55.
- Morse, S. J. (1999). Crazy and criminal responsibility. *Behavioral Sciences and the Law*, 17, 147–164.
- Morse, S. J., & Cohen, E. (1982, June). Diminishing diminished capacity in California. *California Lawyer*, 24–26.
- Nevada Revised Statutes § 174.035 and 193.220 (1995).
- Nevins-Saunders, E. (2012). Not guilty as charged: The myth of mens rea for defendants with mental retardation. *University of California, Davis, Law Review*, 45, 1419–1486.
- People v. Cantrell, 8 Cal. 3d 672, 504 P. 2d 1256, 105 Cal. Rptr. 792 (1973).
- People v. Carpenter, 464 Mich. 223 (2001).
- People v. Conley, 64 Cal. 2d 310, 411 P. 2d 911, 49 Cal. Rptr. 815 (1966).
- People v. Drew, 22 Cal. 3d 333, 583 P. 2d 1318, 149 Cal. Rptr. 910 (1974).
- People v. Gorshen, 51 Cal. 2d 716, 336 P. 2d 492 (1959).
- People v. Henderson, 60 Cal. 2d 482, 386 P. 2d 677 (1963).
- People v. Poddar, 10 Cal. 3d 750, 518 P. 2d 342, 111 Cal. Rptr. 910 (1974).

- People v. Savoie, 419 Mich 118 (1984).
- People v. Wells, 33 Cal. 2d 330, 202 P. 2d 53 (1949).
- People v. Wetmore, 22 Cal. 3d 318, 583 P. 2d 1308, 149 Cal. Rptr. 264 (1978).
- People v. Wolff, 61 Cal. 2d 795, 394 P. 2d 959, 40 Cal. Rptr. 271 (1964).
- Phillips, J., & Woodman, R. (2008). The insanity of the mens rea model: Due process and the abolition of the insanity defense. *Pace Law Review*, 28, 455–494.
- Powell v. Texas, 392 U.S. 514 (1968).
- Senate Report No. 225, 98th Cong., 2d Sess. 230, reprinted in U.S. Code Cong. & Ad. News 3182, 3412 (1984).
- Steadman, H. J., McGreevy, M. A., Morrissey, J. P., Callahan, L. A., Robbins, P. C., & Cirincione, C. (1993). *Before and after Hinckley: Evaluating insanity defense reform*. New York, NY: Guilford Press.
- Stimpson, S. (1993). *State v. Cowan*: The consequences of Montana's abolition of the insanity defense. *Montana Law Review*, 55, 503–524.
- State v. Albright, 273 Kan. 811, 46 P.3d 1167 (2002).
- State v. Bethel, 275 Kan. 456, 66 P.3d 840 (2003).
- State v. Delling, 267 P.3d 709 (Idaho, 2011).
- State v. Herrera, 895 P.2d 359 (Utah, 1995).
- State v. Jorrick, 269 Kan. 72, 4 P.3d 610 (2000).
- State v. Korell, 213 Mont. 316, 690 P.2d 992 (1984).
- State v. Searcy, 118 Idaho 632, 798 P.2d 914 (1990).
- Szasz, T. (1981–1982). The political use of psychiatry—The case of Dan White. *American Journal of Forensic Psychiatry*, 2, 1–11.
- Tarasoff v. Regents of the University of California, 551 P. 2d 334 (1976).
- U.S. v. Bartlett, 856 F. 2d 1071 (8th Cir. 1988).
- U.S. v. Brown, 880 F. 2d 1012 (9th Cir. 1989).
- U.S. v. Cameron, 907 F. 2d 1051 (11th Cir. 1990).
- U.S. v. Childress, 58 F. 3d 693 (D.C. Cir. 1995).
- U.S. v. Frisbee, 623 F. Supp. 1217 (D.C. Cal. 1985).
- U.S. v. Gold, 661 F. Supp. 1127 (D.D.C. 1987).
- U.S. v. Hood, 857 F. 2d 1469 (4th Cir. 1988).
- U.S. v. Newman, 889 F. 2d 88 (6th Cir. 1989).
- U.S. v. Pohlott, 827 F. 2d 889 (3rd Cir. 1987).
- U.S. v. Schneider, 111 F. 3d 197 (1st Cir. 1997).
- U.S. v. Twine, 853 F. 2d 676 (9th Cir. 1988).
- U.S. v. White, 766 F. 2d 22 (1st Cir. 1985).
- Utah Code, §76-2-305 (1990).
- Waddell, C. W. (1979). Diminished capacity and California's new insanity test. *Pacific Law Journal*, 10, 751–771.

PART FOUR

SPECIAL APPLICATIONS

Assessing Violence Risk

KEVIN S. DOUGLAS, STEPHEN D. HART, JENNIFER L. GROSCUP, AND THOMAS R. LITWACK

IN a variety of contexts, our legal system allows for or requires assessments of risk for violence of certain individuals. That is, the law requires assessments of the risk that those individuals will cause certain types of harm under particular conditions within particular periods of time (*Schall v. Martin*, 1984; Shah, 1978). Such assessments, currently most commonly labeled violence risk assessments, can significantly affect the lives of those individuals (*Barefoot v. Estelle*, 1983; *Kansas v. Hendricks*, 1997; *United States v. Salerno*, 1987) and, if a serious proclivity toward violence goes undetected, perhaps others—that is, potential victims—as well (Monahan, 1993; Schlesinger, 1996). Mental health professionals are often called on, and may even be obliged, to participate in these decisions (*Addington v. Texas*, 1979; Buckner & Firestone, 2000; Faigman & Monahan, 2009; Felthous & Kachigian, 2001; *Tarasoff v. Regents of the University of California*, 1976; VandeCreek & Knapp, 2001; Walcott, Cerundolo, & Beck, 2001; Wilson & Douglas, 2009). Although the flurry of foundational risk-relevant legal activity waned to some extent in North America throughout the 1980s and 1990s, if one looks globally, there is no shortage of the enshrinement of risk within law and policy (Wilson & Douglas, 2009).

This chapter addresses violence risk assessments, historically referred to as assessments of dangerousness, concerning mentally or personality disordered individuals possibly at risk for violence in the community. First, we review some of the landmark commentary, case law, and research that spurred the proliferation of risk assessment research and professional activity. We then review the major approaches to risk assessment, their advantages and disadvantages, and corresponding research. Then we provide a step-by-step overview of the clinical risk assessment process. Next, we survey recent developments in the law that concern the admissibility in court of such assessments and the legal standards for depriving individuals of their liberty based on such assessments. We conclude by recommending several fruitful areas for research and practice concerning violence risk assessment to tackle.

LANDMARK EARLY RESEARCH AND COMMENTARY

The first comprehensive review of the research literature regarding assessments of dangerousness by mental health professionals was John Monahan's influential monograph *Predicting Violent Behavior: An Assessment of Clinical Techniques*, which appeared in 1981. (An earlier, less extensive, review by Dix, 1980, was also quite useful.) Monahan updated his 1981 review in another prominent article published in 1984 (see also Wettstein, 1984). Litwack and Schlesinger (1987) reviewed the research literature through 1985 for the first edition of the present volume and arrived at conclusions somewhat different from those of Monahan. Litwack, Kirschner, and Wack (1993) reviewed the relevant studies from 1985 to 1990 and concluded, echoing both Monahan (1981, 1984) and Litwack and Schlesinger, that "research had *not* negated the possibility that clinical evaluations of dangerousness *can* have a unique and useful role to play in making determinations of dangerousness that our society has decided should be made" (p. 269; emphasis in original). Litwack and Schlesinger (1999) updated their review of the literature (through 1997) for the second edition of this *Handbook* and echoed this sentiment.

However, in 1999, Quinsey, Harris, Rice, and Cormier proposed the "complete replacement" of clinical assessments of dangerousness with actuarial methods (p. 171). In response, Litwack (2001, p. 409, emphasis added) reviewed the studies directly comparing clinical with actuarial risk assessments and concluded that "research to date has *not* demonstrated that actuarial methods of risk assessment are superior to clinical methods." Rather, he proposed, "it seems that much more research is needed to determine the relative merits of clinical versus actuarial assessments of dangerousness and that such research should be conducted in as meaningful a manner as possible" (p. 424). (The actuarial versus clinical distinction and debate is discussed further in this chapter.)

Here we review some of the most prominent early studies of violence risk assessment both because of their continuing notoriety and because they illustrate important points that should be considered in evaluating most research studies regarding assessments of dangerousness and/or the assessments themselves. Then we consider major research studies published in recent years to evaluate the current state of research findings concerning violence risk assessments.

KOZOL, BOUCHER, AND GAROFALO (1972)

The most widely cited study for the often-stated proposition that predictions of violence by mental health professionals are wrong at least two times out of three, even when based on a known history of violence and extensive clinical examinations (*Barefoot v. Estelle*, 1983, dissenting opinion; Monahan, 1984), is a study by Kozol, Boucher, and Garofalo (1972). A close examination of this study illustrates, however, how cautious one should be before drawing firm conclusions from many early studies of assessments of dangerousness (as it was then commonly called).

Using clinical examinations, extensive life histories, and psychological tests, a team of mental health professionals evaluated 592 males convicted of assaultive offenses (usually sexual in nature) and sentenced to a special facility for continued evaluation and treatment. Of these men, 386 were eventually classified as *not* dangerous by the evaluating team and released. In addition, 49 men classified as still dangerous were also released by legal authorities *against the advice* of the professional staff. During a 5-year community follow-up, 8% of the patients considered by the evaluating teams to be *not* dangerous were found to have committed a serious assaultive crime. By contrast, 35% of the patients viewed as dangerous by clinicians, but nevertheless released, were discovered to have committed a serious crime.

On the surface, the clinicians studied here did appear to do much better than chance in their evaluations because the recidivism rate of offender patients released against the advice of the evaluating teams was much higher than the recidivism rate of patients evaluated to be no longer dangerous. However, it appears from the report that the examinees deemed dangerous but nevertheless released were, on average, at risk for recidivism in the community for a significantly longer period than those released after clinical judgments of nondangerousness. Thus, it cannot be definitively concluded from this study that the clinicians at issue demonstrated at least *some* ability to assess dangerousness.

However, for the reasons that follow, neither should it be concluded from this study that “predictions of violence” by mental health professionals are wrong at least two-thirds of the time. To begin with, Kozol, Boucher, and Garofalo (1973) reported subsequently that at least 14 of their 49 patients classified as dangerous and yet released were patients who had been committed and studied during the early years of their program and who would *not* have been considered to be dangerous in the later years of their study (when, presumably, their assessment techniques and judgments were more refined). Thus, Kozol et al. eventually may have developed an assessment system that was able to predict which of their sample of patients would be dangerous if released with at least 50% accuracy.

Second, the recidivism rate reported by Kozol et al. (1972) for the patients deemed dangerous—whether 35% or 50%—was presumably lower than the *actual* recidivism rate. Hall (1982) has pointed out that there is good reason to believe that only 20% of serious crimes lead to an arrest. Thus, many seeming false positives (especially among individuals with a history of serious violence) may, in fact, be undiscovered true positives.

Third, the 49 patients released despite clinical judgments of dangerousness were *not a representative sample* of patients judged to be dangerous by the clinicians in this study. Because these 49 individuals were released by judicial or parole authorities against professional advice, they were almost certainly patients whose dangerousness was in fact uncertain (i.e., there was only possible or partial evidence supporting their dangerousness). Why else, in the main, would they have been released? However, if the far larger number of patients classified as dangerous by the teams and legal authorities and not released had been released, their rate

of recidivism—and the apparent accuracy of the clinical assessments—might have been far higher than it appeared to be for the patients whose dangerousness was questionable. Thus, this study simply does not demonstrate that predictions of violence by mental health professionals are likely to be wrong two-thirds of the time, regardless of the sample of individuals being evaluated (Litwack, 1996), the circumstances involved, or the confidence of the clinicians in their judgments (McNiel, Sandberg, & Binder, 1998; see also Douglas & Ogloff, 2003a, for a discussion of confidence and accuracy).

Finally, and most important, the judgments of dangerousness at issue in this study, like most judgments of dangerousness, were never predictions of violence to begin with. The patients supposedly predicted to be violent were actually, and more conservatively, “*not recommended for release*” (Kozol et al., 1972, p. 390). Clinical concerns regarding a patient’s potential dangerousness that lead to a conclusion that the patient cannot be recommended for release do not equate with a prediction that a patient will be violent if released (Mulvey & Lidz, 1995). A clinical judgment that a patient is dangerous, even a judgment that an individual is sufficiently dangerous to warrant confinement, is rarely, if ever, a prediction that a patient unquestionably will be violent if at liberty or, at least, if unsupervised. Rather, almost always, it is a judgment that the subject poses a *significant risk* of acting violently in certain circumstances. Indeed, in forensic settings, patients with a history of serious violence may well be deemed a continued danger by clinicians and judges even if it clearly could not be concluded with confidence that those patients would recidivate if released, as long as it is determined that the patients are still significantly *at risk* for serious recidivism (Litwack, 1996, pp. 108–115; Monahan & Silver, 2003). That is, a clinical conclusion that a patient cannot be recommended for release—or even a conclusion that a patient remains dangerous—may simply be a determination that the patient remains prone to violence *under certain circumstances* (e.g., if the patient stops taking certain prescribed medications, or reengages in substance abuse, or enters into a certain type of relationship). Violence may or may not actually occur, depending on the totality of the clinical picture. Thus, if a patient deemed by a clinician to be too dangerous to be recommended for release is nevertheless released and does not recidivate, it is incorrect to conclude that the clinician made an inaccurate prediction. To the contrary, although the clinician may have concluded that the risk of causing harm still posed by the patient was too great to recommend the patient for release, the clinician may also have concluded that the patient might well not recidivate. This fact poses serious difficulties for attempts to evaluate the validity of clinical assessments of dangerousness (Litwack, 2002) or the relative merits of clinical versus actuarial assessments (Litwack, 2001, pp. 425–426).

BAXSTROM AND DIXON STUDIES

Other historically important studies of violence risk assessments by mental health professionals are the well-known *Baxstrom* studies (*Baxstrom v. Herold*, 1966;

Steadman & Coccozza, 1974) and the similar study of the *Dixon* patients by Thornberry and Jacoby (1979; *Dixon v. Attorney General*, 1971, reviewed in detail by Litwack, 1996). Briefly stated, these studies concerned hundreds of individuals (usually convicted offenders) confined for many years in forensic hospitals because they were considered to be too dangerous to be released to civil hospitals, much less to the community. Yet, as a result of judicial decisions, these patients were nevertheless transferred to civil hospitals. Follow-up studies indicated that only a small percentage had to be returned to secure facilities and only a small number of patients ultimately released to the community were rearrested for violent offenses. (The great majority of the *Baxstrom* patients, many of whom were quite elderly, did require continued confinement in civil facilities, and of the 65% who were ultimately discharged, 11% were rearrested for violent offenses.)

These findings indicate that, in the past, many mentally ill individuals were wrongfully confined in unduly restrictive facilities because of erroneous assumptions that they were too dangerous to live in less restrictive conditions; therefore, determinations of dangerousness for the purpose of preventive detention warrant careful judicial scrutiny. But it is equally clear that the determinations of dangerousness on which the unnecessarily severe confinements were grounded were not based on careful, individualized assessments but on what have been described as administrative decisions (*Baxstrom v. Herold*, 1966, n.3), global assessments (Steadman & Coccozza, 1980, p. 212), and “political predictions” (Thornberry & Jacoby, 1979, p. 26). Nor is there any evidence that the psychiatrists who made those determinations were anything like a representative sample of psychiatrists. Indeed, given the fact that many of these patients had grown old, it is hard to believe that a representative sample of mental health professionals would have determined so many of them to be seriously dangerous.

Indeed, Litwack (1996) described a representative sample of patients confined on the grounds of dangerousness (and mental disorder) in a secure forensic facility and compared his sample to Thornberry and Jacoby’s (1979) sample. Litwack concluded that the samples and assessment techniques employed were so different that “the validity, or invalidity, of the ‘predictions’ of dangerousness at issue in the study of the *Dixon* patients [were] utterly irrelevant to evaluating the validity and/or legitimacy of the assessments of dangerousness” he surveyed (p. 118). Moreover, Litwack pointed out, “once even a semblance of an individualized examination was performed . . . on the *Dixon* patients, only a distinct minority were still deemed to be dangerous” (p. 118). In short, it is wrong to draw conclusions about assessments of dangerousness in other—much less all—circumstances from findings regarding such assessments in particular (and perhaps highly unrepresentative) circumstances.

These few studies, spurred in large part by legal decisions, are foundational within the violence risk assessment field. They raised alarm bells among the legal and mental health professions in terms of the supposed inability of mental health professionals to predict violence. Despite drawing harsh conclusions about clinical

ineptitude based on less-than-optimal data or actual clinical risk assessments, these few studies, perhaps ironically, are likely the reason that the contemporary risk assessment field has received as much attention as it has. We now shift our attention to more contemporary themes.

CONTEMPORARY MODELS OF VIOLENCE RISK ASSESSMENT

Although unstructured clinical prediction still may be the most common approach to risk assessment, generally, as described here, it is not a *model* of risk assessment per se. In terms of specific, defined models, two primary contemporary models of violence risk assessment are commonly used in applied settings: actuarial and structured professional judgment (SPJ). There has been a vast amount of research on these models, particularly SPJ, since the previous edition of this volume. The actuarial method is rooted in the classic human decision making (prediction) debate between clinical prediction and actuarial prediction (Meehl, 1954). As we reviewed earlier, correctly or otherwise, much of the early research on violence risk assessment or “dangerousness” led investigators to conclude that clinical predictions by mental health professionals were poor. Research on actuarial risk assessment was spurred by these dismal conclusions. And, in part, SPJ arose out of perceived weaknesses in the unstructured clinical and actuarial approaches. Next we review these approaches, including their primary features and their pros and cons. Following that, we review research addressing their validity.

CLINICAL JUDGMENT

Unstructured clinical judgment is not a *model* of risk assessment per se. In fact, it was described by Meehl (1954) as an approach that lacks rules. It is based on professional opinion, experience, and intuition, and clinicians have absolute discretion in terms of the risk factors they rely on and how to integrate them. As Grove and Meehl (1996) wrote, clinical prediction is an “informal, ‘in the head,’ impressionistic, subjective conclusion, reached (somehow) by a human clinical judge” (p. 294). Although we believe that clinical experience is essential for conducting risk assessments, we also believe that exercising such experience within a structured decision-making context is necessary.

Although the traditional definition of clinical prediction stressed the absence of rules and the unfettered application of unchecked intuition, clinicians may and do take into account whatever available (or obtainable) data they deem relevant to their assessments, including data that can be obtained only through clinical methods (e.g., a patient’s fantasies or level of insight). Thus, although clinicians undoubtedly vary in their knowledge of which factors to consider when evaluating dangerousness and in their ability to properly assess and weigh those factors, we in fact do not accept the notion, suggested by others (e.g., Grove & Meehl, 1996), that clinical judgments are merely “subjective” or “impressionistic.” Rather, as Holt (1970, p. 348) aptly

noted decades ago, “Disciplined analytical judgment is generally better than global, diffuse judgment; but it is not any the less clinical.” (For a further discussion of the distinctions between actuarial and clinical assessments and between actuarial and clinical variables, see Litwack, 2001, pp. 412–414.) Despite these caveats, we address the unstructured clinical judgment “model” next.

Strengths and Weaknesses of Unstructured Clinical Discretion. Using clinical judgment does have benefits, including being able to respond to rare or unusual circumstances and to important elements of the individual case. This fosters case conceptualization and individualized risk management plans. However, a *purely* unstructured approach has substantial weaknesses, given the absence of guidance to decision makers. For instance, clinicians may pay undue attention to factors that are not associated with violence. Conversely, they may fail to attend to important factors that are indeed associated with violence.

Given a complete absence of structure, there may be inconsistency across raters as well as inconsistency within raters but across cases. Therefore, over time, it is probable that use of this decision approach will produce decisions of lower reliability that are less strongly related to violence than a structured approach. Guy’s (2008) meta-analysis, for instance, showed that unstructured approaches were significantly less strongly related to violence than were structured approaches (either actuarial or SPJ).

Furthermore, from an ethical perspective, entirely unstructured approaches are problematic, because the final decision may rest on unspecified factors and processes. That is, there may be little transparency in unstructured clinical decision making, which poses problems in legal settings because such decisions cannot adequately be reviewed. This places examinees’ rights (i.e., to not be subject to arbitrary detention), as well as continuity of care, at jeopardy.

ACTUARIAL PREDICTION

By contrast, the actuarial approach to prediction is “a formal method” that “uses an equation, a formula, a graph, or an actuarial table to arrive at a probability, or expected value, of some outcome” (Grove & Meehl, 1996, p. 294). The defining feature of the actuarial prediction method is the derivation and use of replicable, routinized rules for combining or integrating predictive factors. Risk factors are selected because they are predictive of violence and are combined based on their joint predictive qualities in the derivation (and, preferably, cross-validation) sample(s). Major contemporary actuarial instruments are listed (in alphabetical order) in Table 14.1.

Strengths and Weaknesses of Actuarial Decision Making. Actuarial prediction methods may possess good reliability, because the methods for using (i.e., selecting, weighing) their risk factors are explicit. This can also produce good predictive validity in

Table 14.1
Select Actuarial Violence Risk Assessment Instruments

Instrument	Intended Application	Number of Items
BVC (Brøset Violence Checklist; Almvik, Woods, & Rasmussen, 2000)	Imminent violence among adult psychiatric inpatients	6
COVR (Classification of Violence Risk; Monahan et al., 2005)	Violence among acute psychiatric patients being discharged to the community	Up to 40
J-SORRAT-II (Juvenile Sexual Offense Recidivism Risk Assessment Tool-II; Epperson, Ralston, Fowers, DeWitt, & Gore, 2006)	Sexual recidivism among juvenile sex offenders	12
LS/CMI (Level of Service/Case Management Inventory; Andrews, Bonta, & Wormith, 2004)	General and violent recidivism among adult offenders	124
LSI-R (Level of Service Inventory-Revised; Andrews & Bonta, 1995)	General and violent recidivism among adult offenders	54
MnSOST-R (Minnesota Sex Offender Screening Tool-Revised; Epperson, Kaul, Huot, Goldman, & Alexander, 2003)	Sexual recidivism among male adult sex offenders	16
ODARA (Ontario Domestic Assault Risk Assessment; Hilton et al., 2004)	Violent recidivism against female partners by adult males with a police record for domestic assault	13
SORAG (Sex Offender Risk Appraisal Guide; Quinsey, Harris, Rice, & Cormier, 2006)	Violent recidivism among sex offenders	14
Static-99 (Hanson & Thornton, 1999)	Sexual and violent recidivism among adult male sex offenders	10
VRAG (Violence Risk Appraisal Guide; Harris, Rice, & Quinsey, 1993; Quinsey et al., 2006)	Violent recidivism among adult male offenders, forensic patients	12
VRS (Violence Risk Scale; Wong & Gordon, 1999–2003)	Violent recidivism among adult male offenders	26
YLS/CMI (Youth Level of Service/Case Management Inventory; Hoge & Andrews, 2002)	General and violent recidivism and institutional behavior among adolescent offenders	42

Note. This listing of risk assessment instruments is not intended to be exhaustive but to provide a sampling of commonly used and researched instruments.

Although the Level of Service instruments rely on numeric cut-offs for decision making, they also permit clinical overrides. Hence, the initial estimate of risk is actuarial but is subject to clinical modification. The authors of the YLS/CMI state that it can be used in either an actuarial or SPJ fashion.

comparison with unstructured approaches. In addition, the transparency of these methods is a benefit in legal contexts.

The predictive properties of most actuarial risk assessment instruments tend to be optimized within a derivation sample. Often only one sample is used, and the results may not be cross-validated prior to use (Violence Risk Appraisal Guide [VRAG; Quinsey et al., 2006]; *cf.* Classification of Violence Risk [COVR; Monahan et al., 2005]). This practice is highly problematic because derivation predictive

estimates—which are linked to the specific characteristics of unique samples—tend to change upon cross-validation in new samples. As such, actuarial risk assessment instruments produce *estimated* probabilities of violence over some future time for persons who fall into certain score ranges on the test.

The actuarial approach has a number of vulnerabilities that have yet to be overcome in the risk assessment field. First, most actuarial instruments contain risk factors that were demonstrated to be predictive of violence in one sample (or, more rarely, multiple samples). Although this procedure does indeed demonstrate what variables were predictive in that given sample, there is no guarantee that (a) the same variables will be predictive in other samples or that (b) *only* those risk factors will be associated with violence in other samples. That is, given myriad sample characteristics, it is possible that actuarial instruments contain sample-specific risk factors and exclude important risk factors with broad support in the literature.

Strict actuarial approaches disallow consideration of risk factors not included on the instrument. These approaches presume that (a) the original research considered all potentially relevant risk factors and (b) all potentially relevant risk factors are contained on the instrument. Moreover, if weighting of risk factors is involved, as it often is, the actuarial approach presumes that all risk factors would be weighted equivalently in different samples. Further, attaching weights to risk factors presumes not only that they will apply equally across samples but that they apply equally to all persons within samples. That is, if, say, substance abuse receives a weight of twice that of psychopathy on some actuarial measure, the implied presumption is that substance abuse is twice as important as psychopathy for all persons to whom the instrument might be applied. This is a very high bar to set, and one that we do not think has been met in the risk assessment field.

A major problem with the actuarial approach stems from its *sample dependence*. Risk factors are selected through certain statistical procedures and given weights through one of any number of methods. As a result, the estimated probabilities of violence are all dependent on a host of sample-specific characteristics. These include, *inter alia*, the initial selection of candidate risk factors, how they were measured, how reliable their measurement was, sample size, length of follow-up, nature of the sample mix, definition of violence, and method of detecting violence. Sample dependence means that estimated probabilities of violence may change if the instrument is used in new samples. Therefore, actuarial estimates might not be stable across new samples and, unless demonstrated to be so, ought not be assumed to be so. Indeed, the onus of demonstrating the stability lies on those who use actuarial methods.

To further muddy the waters, it is unclear whether sample-based estimates should be applied to the individuals within those samples at all. Hart, Michie, and Cooke (2007; see also Hart & Cooke, 2013) calculated confidence intervals (CIs) for the VRAG and Static-99 (Hanson & Thornton, 1999) at both the group level and the individual level. Their argument was that, if precise actuarial estimates (i.e., 44% chance of recidivism) are used at the case level, they ought to be shown to possess

meaningful precision at that level (i.e., small CIs). Their findings indicated that CIs at the individual level were so broad (i.e., the CI around the .44 VRAG probability bin was .04–.93) as to render individual level prediction meaningless. Although some commentators have taken issue with the calculation of CIs at the individual level to start with (Hanson & Howard, 2010; Harris & Rice, 2007; Mossman & Selke, 2007), at the very least their work highlights the great difficulty of applying group-based probability estimates to individual persons within those groups.

We illustrate some of these problems. Earlier we noted that one problem with actuarial methods is the exclusion of potentially important risk factors either because they were not included among the set of candidate variables in the first place or because they did not “make it” onto the final instrument in the given derivation sample. Consider the Static-99, a widely used actuarial instrument for sexual violence risk assessment. It omits sexual deviation, even though this risk factor is related to violence across numerous samples (Laws & O’Donohue, 2008). As another example, the VRAG fails to contain various risk factors that have a good deal of empirical support, such as treatment noncompliance and anger. Strictly speaking, from an actuarial perspective, this means that evaluators should not consider these risk factors, because they are not contained on the given instrument. In our view, this highly limits the scope and comprehensiveness of a risk assessment, to the extent that risk factors of potentially vital importance would not be considered by the decision maker.

Next consider the weighting issue. The actuarial approach presumes that weighting improves predictive strength and that the weights apply equivalently to all people (to use our previous example, that substance abuse is twice as predictive as psychopathy, for all people in all samples). To test this assumption, Grann and Långström (2007) compared several types of weighting procedures ranging from simple to complex in a sample of 404 Swedish forensic psychiatric patients who followed for 2 years. They used the H scale of the Historical-Clinical-Risk Management–20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997) as the predictive measure. They found that the more complex weighting procedures resulted in greater degradation of predictive accuracy on cross-validation compared to unit (equal) weighting. In other words, the weights did not hold up on cross-validation. As Dawes (1979) reminded us many years ago, unit weighting produces predictive estimates that are just as accurate as cross-validated weighting procedures, a process he eloquently described as the “robust beauty of improper [unweighted] linear models” (p. 571).

Blair, Marcus, & Boccaccini (2008) also observed the degradation of weights in their meta-analysis. They conducted a meta-analysis to test the extent to which calibration/derivation actuarial predictive estimates remained stable across cross-validation samples. Using the VRAG, SORAG (Sex Offender Risk Assessment Guide, Quinsey et al., 2006), and Static-99, they reported correlational effect sizes across the initial development samples for these instruments as well as for cross-validations conducted by the authors and by independent researchers. The VRAG

correlation was reduced from .44 in the development sample to .36 in author-conducted cross-validations and to .30 in independent samples. Similar reductions were observed for the SORAG and Static-99.

Although the Blair et al. (2008) meta-analysis focused on correlational effect sizes, the findings suggest that the frequency estimates produced by some actuarial instruments (say, e.g., that 55% of people in a given category can be expected to be violent in the future) may not generalize either. Mills, Jones, & Kroner (2005) tested the generalizability of such frequency estimates among 209 offenders using the VRAG and the Level of Service Inventory–Revised (LSI-R; Andrews & Bonta, 1995). They concluded that the “results of this study do not support the generalizability of the original probabilities associated with the prediction bins, although the LSI-R bins performed much better than the VRAG bins” and that “this study does not support the use of the initial validation probability bins of either instrument with our sample” (p. 579). They further found that for the VRAG, there were “probability reversals” across its various categories, in which categories with higher estimates that should have had correspondingly higher observed recidivism rates actually produced lower recidivism rates.

Based on the foregoing, it does not appear that the promise of actuarial instruments has been realized. It is highly likely that persons who score higher compared to lower on these instruments are indeed at higher risk for future violence. Research tells us as much. However, it is not clear that we are able to ascribe precise probability estimates at the level of the individual person as opposed to the level of the sample or population.

Some of the other criticisms of actuarial approaches concern their tenuous relevance to treatment and risk management, which are highly dependent on dynamic risk, or the ability of risk factors to change over time (Douglas & Skeem, 2005). Although actuarial instruments are not inherently static, they tend to include mainly historical factors and do not emphasize repeated measurements to capture change as part of their prescribed use. We focus more on this issue next as we discuss the SPJ approach.

STRUCTURED PROFESSIONAL JUDGMENT

The SPJ approach to violence risk assessment has been in development since the early to mid-1990s (Douglas & Kropp, 2002; Douglas & Ogloff, 2003b; Douglas, Cox, & Webster, 1999; Hanson, 1998; Hart, 1998, 2001; Webster, Harris, Rice, Cormier, & Quinsey, 1994; Webster et al., 1997). It developed, in part, to compensate for the weaknesses of both the unstructured clinical and the actuarial approaches. However, it also aims to retain some of the strengths of both approaches. Having said that, it is not a combination of approaches, and it is not an “adjusted-actuarial” approach. In general, SPJ attempts to retain some of the strengths of the clinical approach, such as its relevance to treatment and risk management and its utility in individual case formulation. In addition, like the actuarial approach, it aims to attain

Table 14.2
Select SPJ Violence Risk Assessment Instruments

Instrument	Intended Application	Number of Items
EARL-20B, Version 2 (Early Assessment Risk List for Boys, EARL-20B, Version 2; Augimeri, Koegl, Webster, & Levene, 2001)	Antisocial and violent behavior in boys under 12	22
ERASOR (Estimate of Risk of Adolescent Sexual Offense Recidivism, Version 2.0; Worling & Curwen, 2001)	Sexual violence among adolescents with histories of sexual violence	25
HCR-20 Version 2 (Historical-Clinical-Risk Management-20; Webster et al., 1997) and Version 3 (Douglas, Hart, Webster, & Belfrage, 2013)	Violence among adult males or females	20
RSVP (Risk for Sexual Violence Protocol; Hart et al., 2003)	Sexual violence among male adults with histories of sexual violence	22
SAPROF (de Vogel, de Ruiters, Bouman, & de Vries Robbé, 2012)	Violence among adults; to be used in conjunction with HCR-20 or SVR-20	17
SARA (Spousal Assault Risk Assessment Guide; Kropp, Hart, Webster, & Eaves, 1999)	Violence against a current or former intimate partner by a man or a woman	20
SAVRY (Structured Assessment of Violence Risk Among Youth; Borum et al., 2006)	Violence among adolescents	30
START (Short-term Assessment of Risk and Treatability; Webster, Martin, Brink, Nicholls, & Desmarais, 2009)	Short-term violence by adult psychiatric inpatients	20
SVR-20 (Sexual Violence Risk-20; Boer, Hart, Kropp, & Webster, 1997)	Sexual violence among male adults with histories of sexual violence	20

Note. This listing of risk assessment instruments is not intended to be exhaustive but to provide a sampling of commonly used and researched instruments.

solid empirical evidence and to couch professional judgments within a structured context. Examples of major SPJ instruments are provided (in alphabetical order) in Table 14.2. Next we briefly describe the main features of the SPJ approach. The later section on the clinical assessment of risk expands on the details of how to use it.

All SPJ risk instruments use *logical* or *rational* item selection to select risk factors of relevance to the form of violence addressed by the particular SPJ measure. This approach fosters generalizability of risk assessments across applications as well as comprehensiveness of the set of risk factors on SPJ instruments. Logical item selection is a process involving a thorough review of scientific, theoretical, and professional literatures on the topic of relevance and the selection of risk factors with broad support across a number of numerous samples and contexts. Its purpose is to derive a set of risk factors that minimizes the likelihood that important risk factors are excluded from an assessment or that irrelevant factors are

included. This approach is not sample dependent (in that items are not statistically derived from single samples), and hence risk factors generalize across settings. This approach contrasts with the empirical item selection approach used by most actuarial methods, the weaknesses of which were outlined earlier. Unlike an unstructured clinical approach, it also ensures that an *a priori*, standard set of risk factors is considered by all clinicians for any case.

For reasons reviewed earlier when discussing the weaknesses of actuarial approaches, SPJ instruments do not use numeric score cut-offs to categorize people, nor do they use numeric probability or frequency estimates of future risk for violence. SPJ approaches require decision makers to classify individuals as low, moderate, or high risk depending on their perceived level of risk and the corresponding required degree or amount of intervention to dampen this risk. Clinicians consider the number and relevance of risk factors that are present in a given case. High risk means that an examinee is considered high priority by the evaluator for receiving risk management or reduction interventions, without which the evaluator is confident that the examinee would commit a violent act. Although some commentators have criticized the lack of a numerical assignment system (i.e., Quinsey et al., 2006), research (reviewed in the “SPJ Research” section) shows that this decision-making system is as accurate as or more accurate than actuarial classification systems.

SPJ approaches do not provide *a priori weighting* to risk factors, as most actuarial instruments do. That is, in most SPJ measures, all factors are judged to be present, absent, or partially/possibly present. Then, based on the given case, clinicians decide which risk factors are more versus less relevant *in that case*. There is no presumption, as there is in actuarial models, that all risk factors are related to violence in the same manner for all persons across all samples. Decision makers consider the *presence* of risk factors that have empirical support at the *nomothetic* level as well as the risk factors’ individual *relevance* at the *idiographic* level.

All SPJ violence risk instruments include *dynamic risk factors*. This fosters treatment and risk management planning. As described, the absence of dynamic risk is not an inherent property of actuarial instruments. However, most such instruments tend not to include an emphasis on dynamic risk (Douglas & Skeem, 2005), which limits their applicability to monitoring risk over time. The SPJ model helps clinicians decide how often to reevaluate risk factors and how to link risk assessment to risk management. All SPJ violence risk instruments contain numerous (10–30) potentially changeable risk factors that inform clinicians’ judgments about what sort of risk reduction strategies are necessary in a given case.

As its name implies, SPJ risk instruments include various mechanisms to structure evaluations. These elements—to be expanded on in our “Conducting Comprehensive Clinical Risk Assessments” section—include:

1. a standard set of risk factors for a given type of concern about violence (e.g., spousal violence, sexual violence, youth violence);
2. operational definitions of risk factors;

3. coding instructions for risk factors;
4. guidance for making final judgments of low, moderate, or high risk based on the presence and relevance of risk factors and degree of intervention required; and
5. facilitation of risk management.

COMPARATIVE EVALUATIVE RESEARCH

How do these approaches fare, both singly and comparatively, in terms of evaluating risk for violence? In this section, we draw some conclusions from the literature on this topic. There are now hundreds of risk assessment studies—especially on actuarial and SPJ approaches. After a review of landmark studies of unstructured clinical judgment, we focus on meta-analytic studies devoted to contemporary actuarial and SPJ instruments as well as illustrative comparative studies. For detailed commentary on early developmental research on key actuarial instruments such as the VRAG and COVR, please see the third edition of this *Handbook* (Weiner & Hess, 2006).

LANDMARK STUDIES ON CLINICAL PREDICTION

In 1993, Lidz, Mulvey, and Gardner published a study on clinical violence risk assessments that was deemed by Monahan in 1996 to be, at that time, “surely the most sophisticated study published on the clinical prediction of violence” (p. 111). The Lidz et al. study remains apposite, and we review it in some detail. In general, this research concluded that mental health professionals’ clinical judgments were significantly predictive of psychiatric patients’ violence.

Pairs of clinicians were asked to independently rate hundreds of psychiatric emergency department patients on a scale from 1 to 5 regarding the patients’ “potential . . . violence toward others during the next 6 months.” Patients who received a summed rating of at least 3 out of a possible score of 10 were included in the “predicted violent” group. Each of these patients was then matched for sex, race, and age with another emergency room patient who had elicited less staff concern about future violence to others. Ultimately, 357 matched pairs were followed for 6 months after their discharge from the hospital. Violent incidents were detected in 36% of the comparison cases and 53% of the predicted cases, a statistically significant difference. Even when the patients’ preadmission history of violence was controlled for, the clinicians still did statistically better than chance, leading the authors to conclude that “this study . . . show[s] that clinical judgment has been undervalued in previous research” (Lidz et al., 1993, p. 1010).

However, because a significant percentage of the patients who did act violently in the community were not identified as dangerous by the clinicians (a measure of the sensitivity of the clinical judgments) and because a considerable percentage of patients who did not act violently in the community were in the predicted violent

group (a measure of the specificity of the judgments), the authors also concluded that “the low sensitivity and specificity of these judgments show that clinicians are relatively inaccurate predictors of violence” (Lidz et al., 1993, p. 1010).

Furthermore, although this study showed that clinicians, generally, did better than chance at predicting violence, this was true only for male patients. Prediction for female patients was no better than chance. Clinicians underestimated the risk posed by women. Although it was estimated to be about half that of men, in fact the men and women had similar base rates of violence in the follow-up period. And in a follow-up analysis of these data, Coontz, Lidz, and Mulvey (1994) showed that clinicians spent much less time asking questions about violence to female patients compared to male patients who had recently been violent. Moreover, in further follow-up analyses, Gardner, Lidz, Mulvey, and Shaw (1996b) showed that a simple, three-item actuarial screen outperformed clinicians’ decisions. Thus, this study, considered one of the best to test clinicians’ predictions, showed lukewarm support for clinicians’ validity. The clinical predictions showed some association with violence but also showed a good deal of error, inapplicability for women, and inferior performance to a simple actuarial screen.

Despite the value of these lessons learned, there are other questions about what exactly can be drawn from this study. To begin with, the clinicians in this study were not predicting violence but instead rating their patients’ *potential* for violence. It is simply incorrect to conclude that a patient who elicited some clinical concern regarding future dangerousness (e.g., summed ratings of 3 or 4 out of a possible 10) has been predicted to be violent (cf. Mulvey & Lidz, 1995). Lidz et al. (1993) did find that patients about whom clinicians expressed serious concern—those who had a summed score of 6 or above—were no more likely to commit violence than patients regarding whom the clinicians had expressed some but less concern. However, a high clinical rating of potential violence does not necessarily mean a judgment that the patient is very *likely* to commit violence. Such a concern may instead reflect a judgment that the patient is at risk for committing *serious* violence, even if the risk of occurrence is not high. That is, the concept of risk is not equivalent to the concept of probability or likelihood. According to the law as well as social science, the concept of risk includes consideration of the nature, severity, imminence, and frequency or duration of harm—as well as its likelihood. Thus, a clinical opinion of high risk could reflect a belief that the patient poses (a) some significant possibility of serious violence, (b) a high probability of minor violence, or (c) a moderate probability of imminent violence.

Moreover, this study certainly does not demonstrate (or refute the idea) that psychiatric emergency room clinicians have sufficient ability to assess future dangerousness to justify emergency commitments based on their assessments of dangerousness. This is because the follow-up measures and analysis, however much a step forward from past efforts, did not assess a crucial variable: *whether the patients’ violence in the community, when it occurred, was sufficiently serious and occurred sufficiently soon after their return to the community that would have justified continued confinement*

had it been foreseen. Even if clinicians can do better than chance when they assess dangerousness, this is a far cry from concluding that they can assess dangerousness sufficiently well to justify depriving a person of liberty based on such an assessment.

As is discussed in more detail later (in our “Violence Risk Assessments and the Law” section), “clear and convincing evidence” of dangerousness is required to justify an extended civil commitment (*Addington v. Texas*, 1979). That a clinical determination of dangerousness can be shown to be likely to be somewhat better than a random judgment—that is, better than chance—does not render that judgment, in and of itself, “clear and convincing evidence” of dangerousness. To put it another way, statistical significance may not amount to legal significance. Further, to have maximum legal utility, risk assessment methods should assess the risk of legally relevant possible violence—that is, violence of legally relevant seriousness, likelihood, and/or imminence (Litwack, 2001).

Furthermore, the comparison to the actuarial screen may not be entirely appropriate. These analyses compared actuarial and clinical predictions of violence for their accuracy in predicting *any* community violence versus predicting *serious* community violence. Actuarial predictions had lower rates of false-positive and false-negative errors than the clinical predictions for *any* violence. However, the actuarial instruments were not superior to clinical judgments in predicting *serious* violence, which is the issue of practical concern. Even more important, data regarding three of the most critical variables in the actuarial prediction equation were collected from patients not in the emergency room but in the community after their discharge from the hospital. These variables were the patient’s score on the Hostility subscale of the Brief Symptom Inventory (BSI; Derogatis, 1993), the patient’s recent history of drug abuse, and the patient’s recent history of violence. Gardner, Lidz, Mulvey, and Shaw (1996a) reported that a simple decision tree relying on these three variables (a BSI Hostility score greater than 2, more than three prior violent acts, and heavy drug use) and age less than 18 predicted future violence as well as a regression-based method using these and other variables.

However, it is questionable, at best, whether the data required for the decision tree could be validly and reliably collected in the emergency room, given patients’ clinical conditions (and other practical considerations) at that time. At the least, therefore, this study did not convincingly demonstrate that actuarial methods are superior to clinical methods in determining which patients evaluated in psychiatric emergency rooms should or should not be hospitalized involuntarily. To the contrary, because the clinicians in Gardner et al.’s studies did as well as the actuarial scheme in predicting future *serious* violence, and because there is every reason to believe that the necessary actuarial data would not have been nearly as valid if collected in the emergency room—if it could have been collected there at all—there is every reason to suppose that clinicians are *superior* to actuarial methods in determining short-term serious dangerousness (the only decision that is actually called for) regarding individuals brought for evaluation to psychiatric emergency rooms.

META-ANALYSES

In an early meta-analysis of 64 prospective studies of predictors of violence among mentally disordered offenders, Bonta, Law, and Hanson (1998) compared "objective risk" assessments to clinical judgment (based on a smaller subset of studies). Objective assessments produced a relatively strong predictive effect ($Zr = .39$) and were considerably more strongly related to general recidivism than clinical judgment ($Zr = .11$). A similar pattern was reported for violent recidivism ($Zrs = .27$ and $.09$). It should be noted that the "objective risk" assessments included very few of the major contemporary measures used today and also included study-specific actuarial procedures that were not cross-validated (i.e., regression equations built from multiple risk factors). Nonetheless, this meta-analysis showed that, at least among mentally disordered offenders, (unstructured) clinical judgment was systematically less strongly related to future violence than actuarial approaches.

Campbell, French, and Gendreau (2009) meta-analyzed 88 studies from between 1980 and 2006. Their primary outcome variable was general recidivism among adults, although subanalyses focused on violent recidivism. Instruments that had the largest number of evaluations were evaluated separately to facilitate cross-instrument comparison. The authors were able to do this for these instruments: HCR-20; Psychopathy Checklist-Revised (PCL-R; Hare, 1991, 2003); Statistical Information on Recidivism (SIR; Nuffield, 1982); the VRAG; and the Level of Service family of instruments, which includes the original Level of Supervision Inventory (LSI; Andrews, 1982) and its revisions, the LSI-R and the Level of Service/Case Management Inventory (LS/CMI; Andrews et al., 2004). In general, the CIs overlapped for most instruments, meaning that no clear "winner" emerged. For institutional violence, effect sizes (Zr) ranged from a low of $.08$ to a high of $.28$. For violent recidivism, they ranged from $.22$ to $.32$. For violent recidivism, instruments that contained dynamic risk factors intending to focus on risk management and treatment produced larger effect sizes than those without such a focus.

In the previously described meta-analysis by Blair et al. (2008), the authors reported effect sizes for the VRAG, SORAG, and Static-99 that were generally in the moderate range ($\sim .30$). However, there was a clear decrease in sizes of effects as one progressed from development samples, to instrument author-conducted studies employing cross-validation samples, through independent cross-validation samples. Across all three instruments, correlational effect sizes decreased from $.39$ (development) to $.36$ (author cross-validation) to $.28$ (nonauthor cross-validation). The primary point here is that one can expect the predictive accuracy of actuarial risk assessment instruments to decrease on cross-validation.

Hanson and Morton-Bourgon (2009) evaluated 118 studies examining sex offender risk assessment. They classified assessments as actuarial, SPJ, and unstructured clinical prediction, and they classified outcomes as "any," "violent," and "sexual." Their main conclusion was that actuarial assessments were more accurate than unstructured clinical judgments for each type of outcome. SPJ measures were intermediate. Actuarial estimates also tended to produce larger effect sizes than SPJ

measures for sexual recidivism, although there were only six SPJ studies, and only three of these used summary risk ratings as opposed to the summation of scores. (There was little difference, though, between the three that did and the three that did not.) The single instrument with the largest effect sizes with sexual recidivism was the SVR-20, an SPJ measure, although this was based on only three studies. The authors advised that given the small number of SPJ studies, caution is warranted in evaluating SPJ instruments based on this particular meta-analysis.

Olver, Stockdale, and Wormith (2009) focused on three specific instruments for their meta-analysis of risk assessment of young offenders. One instrument—the Hare Psychopathy Checklist–Revised: Youth Version (PCL:YV; Forth, Kosson, & Hare, 2003)—is not a risk assessment instrument per se, but they included it for comparative purposes, because psychopathy instruments are commonly used in risk assessment. The authors identified 44 usable studies and compared outcomes across general, violent, nonviolent, and sexual recidivism. None of the measures was specifically developed to assess the risk for sexual recidivism, and none performed well for this outcome. As with other meta-analyses, there were not large differences between measures, each performing comparably and moderately. More specifically, for the Structured Assessment of Violence Risk in Youth (SAVRY; Borum, Bartel, & Forth, 2006), effect sizes (weighted correlations, excluding sexual recidivism) ranged from .30 to .38. For the Youth Level of Service/Case Management Inventory (YLS/CMI; Hoge & Andrews, 2002), the range was .26 to .32, and for the PCL:YV, it was .16 to .28. In another meta-analysis focusing on young offenders, although specifically on sexual violence, Viljoen, Mordell, and Beneteau (2012) again reported no meaningful differences between the four instruments they evaluated (both actuarial and SPJ). Across 33 studies, effect sizes for sexual recidivism were in the small (weighted $r_s = .12-.20$) to moderate range (AUCs $\sim .65$).

Yang, Wong, and Coid (2010) were specifically interested in comparing individual measures in their meta-analysis of nine measures across 28 studies. They were also interested in focusing on contemporary instruments (studies from 1999–2008) and using the PCL-R as a benchmark. That is, to what extent do measures improve on the PCL-R? Further, they investigated the extent to which study features versus specific instruments accounted for differences in predictive effects of instruments across studies. Most instruments had moderate predictive validity, and there were few differences between instruments. Only about 25% of the variance in effect sizes was attributable to instruments, meaning that much of the difference between effect sizes was attributable to study design features. One standout finding was that the interpersonal/affective aspects of the PCL-R (which is not a risk assessment instrument) were substantially less predictive compared to the risk assessment instruments and to the other features of the PCL-R. Only two instruments, across various statistical models, added incremental validity to the PCL-R: the HCR-20 and the Offender Group Reconviction Scale (OGRS; Copas & Marshall, 1998). However, only two studies of the OGRS permitted this comparison, compared to 16 for the HCR-20.

Another, very focused, meta-analysis compared the HCR-20 and PCL-R in 34 samples in which both instruments were included (Guy, Douglas, & Hendry, 2010). Although focused, this meta-analysis contains more HCR-20/PCL-R comparisons than do other meta-analyses. The authors reported that, in general, the instruments performed the same (AUCs = .69 for both). When the psychopathy item was removed from the HCR-20, the instrument's AUC essentially remained unchanged (AUC = .71). Most interestingly, the authors secured seven raw data sets and were able to conduct head-to-head multivariate analyses of the HCR-20 (with the psychopathy item removed) and the PCL-R. The HCR-20 added unique, incremental validity beyond the PCL-R, whereas the converse was not true. Specifically, using meta-analytic logistic regression, the authors reported that, with both instruments included in the analysis, for every 1-point increase on the HCR-20, the probability of detecting violence increased 23%, whereas for every 1-point increase on the PCL-R, the probability of detecting violence decreased by 1%.

In yet another meta-analysis (Singh, Grann, & Fazel, 2011), the authors were interested in determining whether certain instruments were more accurate than others and whether features of study design impacted accuracy. Across 88 independent studies published between 1995 and 2008, the authors evaluated nine instruments. Although 88 studies were included in the meta-analysis, instrument-specific analyses typically were based on no more than 12 studies and typically fewer than 10. Also unlike most previous meta-analyses, they coded instruments into predictive "bins" of low versus high risk. For SPJ instruments, the analysis (for 22 of 27 studies) was based on the non-numeric summary risk ratings of low, moderate, and high risk (collapsed into either low+moderate versus high, or low versus moderate+high, for various analyses). For actuarial instruments, the authors similarly reduced the numeric score categories into two bins. Unlike most previous meta-analyses, they argued that there were indeed differences between measures in terms of predictive accuracy. They reported that the SAVRY (an SPJ instrument) produced the largest effect and the LS and PCL measures produced the smallest. The authors interpreted this finding to mean that instruments designed to assess risk in specific populations (i.e., adolescence) were more accurate than those designed for more general use (LS instruments) or those not designed specifically for risk assessment (PCL instruments). Consistent with this interpretation, the authors also reported that instruments designed to predict violence fared better than those designed to predict general recidivism. This latter finding was also observed in a recent meta-analysis on an overlapping group of 73 studies (Fazel, Singh, Doll, & Grann, 2012). In this latter meta-analysis, instruments designed to assess risk for violence (HCR-20; VRAG; SAVRY; Spousal Assault Risk Assessment Guide [SARA; Kropp, Hart, Webster, & Eaves, 1999]), collapsed across actuarial-SPJ assessment format, had superior predictive validity than did those instruments (LS or PCL family of instruments) designed or used for general criminality, odds ratios of 6.1 versus 2.84, respectively.

It should be pointed out, however, that differences between the SAVRY and the LS family of instruments are only somewhat meaningful, in that the SAVRY was

designed specifically for youth violence whereas the LS family spans youth and adulthood, and most studies were based on the adult LS measures. That is, generally, these differences in effect sizes do not compare instruments that would be used with the same populations. In terms of comparing SPJ and actuarial instruments, SPJ instruments produced diagnostic odds ratios that were larger (4.01–4.15) than those for actuarial instruments (2.77–2.88), although these differences were not significant. Most potential moderators were not significant, including gender, country, setting, and ethnicity, although there was a small trend for assessments to be more accurate among samples with a greater percentage of older, Caucasian participants.

Regrettably, most risk assessment meta-analyses suffer the same flaw: They fail to include the summary risk ratings commonly used with the SPJ approach. Singh et al. (2011) is an exception, where they were able to use summary risk ratings for 22 of 27 SPJ studies. The meta-analysis conducted by Hanson and Morton-Bourgon (2009) also is an exception, although there were only three sex offender studies available for analysis that used summary risk ratings. Although it is important to know how well SPJ instruments fare in terms of the sum of their risk factors, as a general indication of whether, in general, more risk factors equates to higher risk, this numeric index is not the one primarily intended to guide clinical practice. As discussed, raters are expected to come to decisions of low, moderate, or high risk based on their consideration of the number and relevance of risk factors and the expected degree of intervention required to reduce risk. In addition to Singh et al. (2011), we are aware of only one broad meta-analysis that includes summary risk ratings (Guy, 2008). Although it is unpublished at this time, we review it here, because it exhaustively meta-analyzed all SPJ literature.

Guy's (2008) meta-analysis evaluated 113 SPJ disseminations and, where possible, compared these to actuarial and unstructured clinical prediction. Focusing on SPJ studies (albeit the numeric use of instruments), effect sizes did not differ as a function of gender (although they trended in the direction of larger effect sizes for female-only samples, as they did in Singh et al.'s [2011] meta-analysis), Europe versus North America, adult versus adolescent, setting (civil, forensic, correctional, mixed/other), institution versus community, file versus file+interview, or whether authors or translators were involved in the research.

Guy (2008) also compared the numeric use of SPJ instruments with the use of summary risk ratings. Quite consistently, summary risk ratings were more strongly associated with violence relative to numeric use. Similar to the findings of Singh et al. (2011) and Fazel et al. (2012), Guy reported that the HCR-20 summary risk ratings were more strongly related to outcomes when they focused on violence as opposed to nonviolent or general criminality. This was particularly the case for physical violence.

Across all available instruments and effect sizes, SPJ summary risk ratings produced larger effect sizes ($AUC = .68$) than actuarial instruments ($AUC = .62$). For unstructured predictions, the effect sizes were on average smaller ($AUC = .59$).

Based on one composite effect size per study, SPJ summary risk ratings (AUC = .69) and actuarial instruments (.67) were both stronger than unstructured predictions (AUC = .58). In direct comparisons, summary risk ratings and actuarial predictions produced very similar effect sizes that did not differ significantly.

Three general observations about risk assessment meta-analytic research can be made.

1. It is remarkable how many studies of contemporary risk assessment instruments have been conducted, even in the past 10 (or 5) years.
2. Meta-analyses should focus on instruments in the ways that they were intended to be used, as was done by Guy (2008), Singh et al. (2011), and to a lesser extent Hanson and Morton-Bourgon (2009).
3. Most meta-analyses use selection criteria that remove a majority of studies that have been conducted on various instruments. For instance, there have been approximately 100 studies on the predictive validity of the HCR-20, yet in most meta-analyses only a small sampling of these studies (10 to 20) is included.

From a substantive perspective, four conclusions can be drawn.

1. There is very little if any evidence that actuarial methods are more accurate than SPJ instruments, as is argued by some (Quinsey et al., 2006; Rice, Harris, & Hilton, 2010). Indeed, one can fairly state—based on meta-analytic evidence—that professional, nonactuarial judgments, derived in an SPJ context, are as strongly or more strongly associated with violence than actuarial methods.
2. There is some evidence that, when instruments are used with the specific outcomes that they were designed to predict, they perform better than when used with nonspecific outcomes. This was evident with the HCR-20 performing better with violence and physical violence (the outcomes it was designed to be used with) than with general antisocial outcomes.
3. Instruments designed specifically to evaluate risk for violence seem to perform better than those designed or used to predict general criminal behavior.
4. There is fairly clear evidence for validity shrinkage when using actuarial instruments.

SPJ RESEARCH

In this section, we review the studies that have evaluated the summary risk ratings of low, moderate, and high risk used in the SPJ approach, especially those that have compared them to the numeric (actuarial) use of such instruments or to actuarial instruments. We do so because this topic has only recently started to work its way into meta-analyses, and there remains less research on this use of the SPJ model relative to simply summing its risk factors.

By our count, a total of 34 published studies have investigated whether summary risk judgments are predictive of violence (see the appendix to this chapter for a list of these studies). Of these 34 studies, 30 (88%) support the use of final SPJ judgments in violence risk assessment, in that they were significantly predictive of violence. Four studies did not find evidence of predictive validity of final SPJ judgments in predicting violence (Braithewaite, Charette, Crocker, & Reyes, 2010; Schaap, Lammers, & de Vogel, 2009; Sjöstedt & Långström, 2002; Viljoen et al., 2008). In one of these studies that used one SPJ and three actuarial instruments (Sjöstedt & Långström, 2002), *none* of the instruments predicted violence. Similarly, in Schaap et al. (2009), *none* of the HCR-20 or PCL-R indices (total scores, subscale scores, summary risk ratings) were predictive of either general or violent recidivism in this sample of 45 female forensic psychiatric patients. The HCR-20 summary risk ratings produced an AUC of .65 (a moderate effect), suggesting that low power might have contributed to null findings. In Braithewaite et al. (2010), the sample size was only 34. In a sample of 169 male adolescent sex offenders in residential treatment, Viljoen et al. (2008) found that total SAVRY scores were predictive of nonsexual violence in youth whereas final structured professional ratings were not.

Of note, the 34 studies spanned multiple countries—Canada, Denmark, Finland, the Netherlands, Norway, Portugal, Serbia, Spain, Sweden, the United Kingdom, and the United States—and included numerous different measures: Early Assessment Risk List for Boys (EARL-20B; Augimeri et al., 2001); Estimate of Risk of Adolescent Sexual Offense Recidivism (ERASOR; Worling & Curwen, 2001); HCR-20; Structured Assessment of Protective Factors for Violence Risk (SAPROF; de Vogel et al., 2012); SARA; SAVRY; Short-Term Assessment of Risk and Treatability (START; Webster et al., 2009); START: Adolescent Version (START:AV; Nicholls, Viljoen, Cruise, Desmarais, & Webster, 2010); Sexual Violence Risk–20 (SVR-20; Boer, Hart, Kropp, & Webster, 1997; de Vogel, de Ruiter, van Beek, & Mead, 2004). As such, the finding that summary risk ratings are predictive of violence is robust and is so across country and instrument.

Of the published studies supporting the use of summary risk ratings, half (17) have tested whether final SPJ judgments added incrementally to the prediction of violence over and above the numerical (or actuarial) use of the instrument, or of a PCL measure, an actuarial measure, or unstructured clinical prediction. To test incremental validity, investigators typically use a multivariate regression approach in which the comparison index is entered in the first step of the model and the summary risk ratings are entered as the second step. The key outcome is whether the addition of the summary risk rating adds in a statistically significant manner to the predictive power of the multivariate model. In all but two of the 17 studies, incremental validity was observed.

For instance, HCR-20 summary risk ratings add incremental validity beyond the numeric use of the instrument among mentally disordered offenders (de Vogel & de Ruiter, 2006), psychiatric inpatients (Arbach-Lucioni, Andres-Pueyo, Pomarol-Clotet, & Gomar-Sones, 2011), forensic patients (Douglas, Hart, & Ogloff, 2003;

Pedersen, Rasmussen, & Elsass, 2010), and criminal offenders released from prison (Douglas, Yeomans, & Boer, 2005) or on probation or parole (Neves, Goncalves, & Palma-Oliveira, 2011). A Dutch measure modeled after and highly similar to the HCR-20 also showed incremental validity among forensic patients (van den Brink, Hooijschuur, van Os, Savenije, & Wiersma, 2010).

The START summary risk rating showed incremental validity beyond its own numeric total among forensic psychiatric inpatients (Desmarais, Nicholls, Wilson, & Brink, 2012), although its adolescent version, the START:AV, did not (for violence; it did for substance abuse, victimization, and suicidal ideation). Incremental validity has been shown with the SAVRY in three samples (Dolan & Rennie, 2008; Lodewijks, de Ruiter, & Doreleijers, 2008; Lodewijks, Doreleijers, & de Ruiter, 2008) but not in two others (Schmidt, Campbell, & Houlding, 2011; Vincent, Chapman, & Cook, 2011). Incremental validity of summary risk ratings has also been found with the EARL-20B in 6- to 12-year-old boys (Enebrink, Långström, & Gumpert, 2006) and the SARA in a sample of adult domestic violence offenders (Kropp & Hart, 2000).

When compared to actuarial tools such as the VRAG, the SORAG, and the Static-99, summary risk ratings of SPJ instruments tend to have better predictive utility for the outcome of interest (Guy, 2008; Heilbrun, Douglas, & Yasuhara, 2009). More important, researchers have investigated whether summary risk judgments add predictive validity above and beyond other instruments typically used for violence risk assessment. Most such studies have found that summary risk ratings add incrementally to actuarial risk assessment tools. Some studies using multivariate analyses have demonstrated that the HCR-20 summary risk judgment added incrementally to the PCL-R when used to predict violence in samples of criminal offenders and forensic patients (de Vogel & de Ruiter, 2005; Douglas et al., 2003, 2005). The START summary risk ratings have shown similar incremental validity beyond the screening version of the PCL:SV (Desmarais, Nicholls, Wilson, & Brink, 2012), as have SAVRY summary risk ratings beyond the PCL:YV (Dolan & Rennie, 2008). Douglas and colleagues (2005) found that the HCR-20 and the VRAG both accounted for unique variance in predicting violence; that is, they both added predictive power over the other. Lodewijks, Doreleijers et al. (2008) showed that SAVRY summary risk ratings possess incremental validity beyond unstructured clinical predictions, a finding consistent with Guy's (2008) meta-analytic finding that unstructured predictions were substantially less accurate than SPJ or actuarial methods.

Considering all of the empirical evidence, it appears safe to conclude that SPJ instruments, used in practice as intended, are as accurate as or more accurate than actuarial indices derived from their numeric use, than actuarial instruments, than unstructured clinical prediction, and than PCL instruments. Furthermore, they do not suffer the same drawbacks as actuarial methods, discussed earlier. We now turn to a discussion of conducting clinical assessments of risk, with a focus on the steps laid out in the SPJ model.

CONDUCTING COMPREHENSIVE CLINICAL RISK ASSESSMENTS

Clinical evaluations of violence risk have been described and discussed for decades (Kozol et al., 1972; Litwack & Schlesinger, 1999; Litwack, Zapf, Groscup, & Hart, 2006; Monahan, 1981; Scott, 1977; Tardiff, 1996). There is good consensus (if not unanimity) that these evaluations comprise two distinct phases. The first step is to understand an examinee's potential for violence. It involves systematic analysis of the examinee's violence history (violent acts, attempts, threats, ideation, and intent), psychosocial adjustment more generally (e.g., problems, vulnerabilities, and strengths), and living situation (e.g., social and physical environment). The goal is to understand what kinds of violence examinees might perpetrate, against which people, for which reasons, and under which circumstances. This is sometimes characterized as a process of *prediction*, although this is true only if the latter term is used loosely to mean understanding what might happen in the future rather than making precise, quantitative, probabilistic estimates of violence. (Understanding potentials and making probabilistic predictions are quite different things.)

The second step is to determine what events and occurrences might increase or decrease examinees' potential for violence. It involves systematic analysis of possible future living situations, both with and without the influence of external agents, to identify putative controlling factors. The goal is to prevent (i.e., mitigate the risk of) future violence by developing a plan for intervention. This is sometimes characterized as a process of *management*.

Although these phases are conceptually distinct, there is some debate concerning whether they are independent or separable. Some (Heilbrun, 1997; Quinsey et al., 2006) have argued in favor of this view. They see a simple, linear temporal relation between the two. Professionals start by completing the assessment/prediction phase; then, if needed, they complete the management phase. Commentators who hold this view believe that, in certain legal contexts, violence risk assessment requires only prediction—that is, the only legally relevant issue is a subject's potential for future violence—whereas in other legal contexts, as well as all clinical contexts, it requires both prediction and management.

Others (Hart, Douglas, & Webster, 2001) have argued that the relation between prediction and management is bidirectional or recursive. First professionals start the prediction phase; but the prediction phase segues into the management phase, which may stimulate reiteration of the prediction and, in turn, management phases. According to this view, the prediction and management phases are flip sides of the same coin. Each requires the other: Good prediction is impossible without systematic consideration of the subject's possible future living situations and interventions; and, conversely, good management is impossible without systematic consideration of the potential for future violence. Commentators with this perspective believe that the prediction phase is conditional or contingent on the management phase and that prediction is thus not legally relevant without consideration of management.

Regardless of how one conceptualizes or defines the process of evaluating violence risk, then, there seems to be unanimous agreement that *comprehensive* clinical evaluations incorporate both prediction and management. As unstructured clinical judgment by definition does not provide a well-defined framework for either prediction or management, actuarial decision making (i.e., reliance on actuarial risk assessment instruments [ARAI]) provides a framework only for prediction. In the remainder of this section we present a framework for comprehensive clinical evaluations of violence risk based on the SPJ approach.

FRAMEWORK BASED ON THE SPJ APPROACH

As discussed previously, the hallmark of the SPJ approach is that professionals use evidence-based guidelines to structure evaluations of violence risk, but their decisions rely on the exercise of discretion rather than strict adherence to a fixed and explicit algorithm developed a priori. In some respects, it is misleading to speak of an SPJ approach, as this gives the impression that there is a single, monolithic set of rules to which all those who develop or use SPJ guidelines adhere. The reality is that various guidelines have been developed by multiple (overlapping) groups working around the world over the past 20 years or so, each geared for different populations of subjects, different groups of users, and different forms of violence. Yet these SPJ guidelines are consistent at the most broad or general level in positing that the task of comprehensive clinical evaluations of violence risk can be broken down into about six major steps (e.g., Hart & Logan, 2011):

1. Gather critical information.
2. Identify the presence of risk factors.
3. Evaluate the relevance of risk factors.
4. Develop scenarios of violence in light of risk factors that are present and relevant.
5. Develop management plans in light of scenarios of violence.
6. Communicate the evaluation findings.

Some SPJ guidelines collapse these steps. For example, the Workplace Assessment of Violence Risk (WAVR-21; White & Meloy, 2010) discusses only three steps: gather information, consider presence of risk factors, and communicate findings. But in the WAVR-21, the final step incorporates Steps 3, 4, and 5 above, even though it does not explicitly discuss and structure them. Other SPJ guidelines add additional steps. For example, one of the newest SPJ guidelines, Version 3 of the Historical-Clinical-Risk Management-20 (HCR-20 V3; Douglas, Hart, Webster, & Belfrage, 2013; see Table 14.3) separates Step 3 into two steps: Evaluate the relevance of risk factors individually and develop a multifactorial formulation of violence risk.

We turn next to a discussion of each of the six steps of the general SPJ process of comprehensive clinical evaluations of violence risk. We break down Step 3 into two substeps, following the HCR-20 V3.

Table 14.3

Violence Risk Factors Contained Within the HCR-20 V3

Historical Scale (History of Problems With . . .)

- H1. Violence
 - a. As a Child (12 and Under)
 - b. As an Adolescent (13–17)
 - c. As an Adult (18 and Over)
 - H2. Other Antisocial Behavior
 - a. As a Child (12 and Under)
 - b. As an Adolescent (13–17)
 - c. As an Adult (18 and Over)
 - H3. Relationships
 - a. Intimate
 - b. Non-Intimate
 - H4. Employment
 - H5. Substance Use
 - H6. Major Mental Disorder
 - a. Psychotic Disorder
 - b. Major Mood Disorder
 - c. Other Major Mental Disorders
 - H7. Personality Disorder
 - a. Antisocial, Psychopathic, and Dissocial
 - b. Other Personality Disorders
 - H8. Traumatic Experiences
 - a. Victimization/Trauma
 - b. Adverse Childrearing Experiences
 - H9. Violent Attitudes
 - H10. Treatment or Supervision Response
-

Clinical Scale (Recent Problems With . . .)

- C1. Insight
 - a. Mental Disorder
 - b. Violence Risk
 - c. Need for Treatment
 - C2. Violent Ideation or Intent
 - C3. Symptoms of Major Mental Disorder
 - a. Psychotic Disorder
 - b. Major Mood Disorder
 - c. Other Major Mental Disorders
 - C4. Instability
 - a. Affective
 - b. Behavioral
 - c. Cognitive
 - C5. Treatment or Supervision Response
 - a. Compliance
 - b. Responsiveness
-

Table 14.3 (Continued)**Risk Management Scale (Future Problems With . . .)**

-
- R1. Professional Services and Plans
 - R2. Living Situation
 - R3. Personal Support
 - R4. Treatment or Supervision Response
 - a. Compliance
 - b. Responsiveness
 - R5. Stress or Coping
-

Note. Reprinted with permission from Douglas et al. (2013).

Step 1: Gather Critical Information. As Scott (1977) pointed out, “Before factors can be considered, they must be gathered” (p. 129). The quantity and quality of information reviewed by evaluators sets fundamental limits on the reliability and validity of their subsequent judgments.

With respect to the quantity of information, evaluators should try to gather all the information that is necessary to reach opinions regarding risk in the case, given the context of the evaluation. Gathering information requires “patience, thoroughness and persistence . . . rather than diagnostic or interviewing brilliance” (Scott, 1977, p. 129) and “the painstaking assembling of facts and the checking of information from a variety of sources” (Prins, 1988, p. 600). As we noted, however, the context of the evaluation sets limits on the quantity and quality of information available to clinicians (Borum, 1996). For instance, emergency room clinicians do not have the information or the time available to gather information that clinicians typically have in long-term forensic facilities.

Information should be gathered about multiple issues or topics and from multiple sources. Most SPJ guidelines make explicit recommendations for gathering information (e.g., Douglas et al., 2013; Hart et al., 2003; Kropp, Hart, Webster, & Eaves, 1995), although the specific topics and sources relevant in a given case vary according to the nature of the risks being assessed and the specifics of the case. For example, information about an examinee’s sexual functioning is likely to be directly relevant to considering his or her risk for sexual violence but is likely to be less relevant to an evaluation of risk for terroristic violence. One important topic about which information should be gathered in every case is the examinee’s history of violence. We agree with the sage advice from years ago: “Of paramount importance is a meticulous description of the actual assault. . . . The description of the aggressor in action is often the most valuable single source of information” (Kozol et al., 1972, p. 384). If there is such a history, the evaluator should construct a timeline of violent incidents to look for evidence of patterns, such as a clear trajectory or triggering events. For each major incident, the evaluator should attempt to determine who was involved (e.g., perpetrators, victims, witnesses), what happened and why (i.e., intentions, actions, motivations), where and when it happened (e.g., physical and

social context), and the examinee's reactions (e.g., thoughts and feelings before, during, and after the incident).

Typically, sources of information will include:

1. Interviews with and direct observation of the examinee.
2. Interviews with victims of or witnesses to past violence or review of their documented accounts.
3. Interviews with collateral informants, such as family members, friends, and coworkers.
4. Review of collateral records including but not limited to
 - criminal histories from various agencies, police reports, statements by the examinee, and statements by victims and witnesses concerning past violence;
 - prosecution and court files;
 - daily logs from institutional and community corrections;
 - presentence or predisposition reports;
 - mental health and medical treatment records;
 - assessment and treatment reports by civil and forensic mental health consultants;
 - daily logs from outpatient and inpatient treatment facilities; and
 - school, employment, and military records.

However, in many cases it is impossible to gather and consider all the information available in a given case. Evaluators must balance comprehensiveness with efficiency by focusing on information that is useful (i.e., directly relevant to risk assessment), unique (i.e., nonredundant), and trustworthy (i.e., obtained from reliable sources).

With respect to the quality of information, evaluators must make judgments about the credibility of various sources of information on which they relied and attempt to reconcile any contradictory information. For example, there may be contradictory information concerning the extent of the examinee's history of violence or the examinee's behavior during the most recent act of violence; the examinee's history of employment or intimate relationship problems may be unclear; or there may be no information concerning the examinee's plans for release from hospital or prison into the community.

Step 2: Identify the Presence of Risk Factors. Evaluators cannot keep in their heads all the information they gathered in its raw or original form; they simplify it by sorting it into useful units: risk factors. The primary problem here is that evaluators may not use consistent or sensible frameworks to sort information. They may give insufficient weight to risk factors with established validity or too much weight to risk factors of little or questionable validity. Worse still, they may use a framework

that is implicit, so they are not entirely conscious or self-aware of what they did or did not consider risk factors.

SPJ guidelines structure this part of the evaluation process by giving evaluators a checklist or *aide mémoire* of risk factors whose validity is supported by systematic review of the scientific and professional literature. (It is important to emphasize here that the term *risk factor* is used here broadly to include any characteristics of examinees, their living situations, and their plans for the future, and in some cases even characteristics of potential victims, that might increase or decrease risk.) The goal is to focus the attention of evaluators in the first instance on those issues or topics that are considered important in general for evaluations of that sort—things that should be considered at a minimum. The list of risk factors is intended to be reasonably thorough or comprehensive but by no means exhaustive; evaluators are encouraged to go beyond the standard risk factors to identify rare or case-specific risk factors, should they be able to provide a compelling clinical or logical rationale for doing so. We stress that clinicians should be able to articulate a clear rationale for considering case-specific factors and should do so only if it is not possible to capture risk factors with the pre-specified list available on the particular instrument. This will help to avoid the inclusion of risk factors that are unlikely to be associated with violence.

In addition to identifying risk factors, SPJ guidelines discuss the nature and definition of the risk factors, summarize the evidence supporting their validity, and provide guidance for determining if they are or have been present over time in the case at hand. Most SPJ guidelines include a minimum of about 10 and a maximum of about 30 risk factors, although the number, conceptualization, and definition of risk factors vary across guidelines, according to the nature of the risks being evaluated and the nature of the evaluators who will be using the guidelines or the settings in which they will be used. We have opted in this chapter not to provide a detailed review of important violence risk factors, as they are—for the most part—very well cataloged on contemporary violence risk assessment instruments (see, e.g., Table 14.3 for a list of the risk factors included on the HCR-20 V3).

Step 3: Evaluate the Relevance of Risk Factors. Risk factors may be relevant in general or on average according to the scientific and professional literature but apparently irrelevant in the case at hand on the basis of idiographic evaluation. (More rarely, the converse may be true.) The next step, then, is to analyze the risk factors identified as present to make sure they are relevant—that is, germane to the examinee’s risk for violence and the management of that risk. As noted previously, a recent development in the SPJ approach, reflected most clearly in the HCR-20 V3 (Douglas et al., 2013), is to encourage evaluators to consider the relevance of each risk factor both individually (i.e., one at a time) and collectively (i.e., taken together). The crux of the issue is aptly delineated by Kluckhohn and Murray (1953): “[E]very man is in certain respects (a) like all other men, (b) like some other men, (c) like no other

man” (p. 53). In essence, violence risk assessments must investigate all three of these issues.

Determining relevance is an exercise in abductive inference, that is, inference to the best explanation—or, as it is more commonly referred to in clinical practice, formulation (Hart, Sturmeay, Logan, & McMurrin, 2011). Formulation of violence risk should be guided by theory, and it is common in the SPJ approach to use decision theory (Hart & Logan, 2011). Decision theory may be considered a version of a well-established theory of criminal behavior known as the psychology of criminal conduct (PCC) or the general personality and cognitive social learning (GPCSL) perspective (Andrews & Bonta, 2006) but tailored specifically to violence. Decision theory views violence as a choice (i.e., purposive behavior intended to achieve one or more goals). The decision may be made quickly, based on poor information, and with little care and attention—that is, it may be a bad decision or a decision made badly—but it is a decision nonetheless. The bottom line is that, with very rare exceptions, all people think before they commit violence: They choose whom they commit violence against, when they will commit it, and what kinds of violence they will commit. Within the framework of decision theory, it is assumed that before people engage in violence, they have gone through a four-step thought process (Hart & Logan, 2011):

1. The possibility of acting violently in a given situation entered their conscious awareness, and they entertained this notion rather than dismissing it or pushing it out of their minds.
2. They evaluated the possible positive consequences of violence and determined that it might result in reward or benefit for them. Put simply, they perceived that violence might pay off.
3. They evaluated the possible negative consequences of violence and determined that the costs were acceptable.
4. They evaluated their options for committing violence and determined it was feasible.

According to the decision theory framework, comprehensive clinical evaluations of violence risk require evaluators to understand how and *why* examinees decided to engage in violence—as well as why they decided *not* to engage in violence—in the past, including the various factors that impinged on or influenced their decision making, and also to understand what interventions, events, and occurrences might encourage examinees’ decisions to act pro-socially and discourage their decisions to act nonviolently.

Risk factors, then, are things that influence decision making. They can motivate, disinhibit, or destabilize decisions. *Motivators* increase the perceived rewards or benefits of violence. *Disinhibitors* decrease the perceived costs or negative consequences of violence. *Destabilizers* generally disturb or disorganize people’s ability to monitor and control their decision making.

Step 4: Develop Scenarios of Violence. Next, evaluators must make judgments about the examinee's potential for violence in light of the presence and relevance of risk factors. Simply concluding that an examinee is a "high risk for violence" or has a "54% chance of violence within 10 years" is not sufficient for making any kind of reasonable clinical—or even legal—decision. Instead, evaluators need to consider what kinds of violence the examinee might perpetrate, for which motivations, against which victims, with what kinds of consequences, and at which times. In short, they need to consider *scenarios* of violence.

Each scenario is a story about violence the examinee *might* commit. It is not a prediction about what will happen; rather, it is a general forecast or speculation about what reasonably *could* happen, in light of the evaluator's general knowledge and experience and the specifics of the case at hand. Although the number of possible scenarios that could be constructed is virtually infinite, in any given case only a few distinct scenarios should seem reasonable, credible, or plausible in light of what is known about fact and theory (e.g., Chermack & van der Merwe, 2003). Other scenarios will be perceived as implausible and subsequently dismissed, or "pruned" (e.g., Pomerol, 2001).

The SPJ approach to developing scenarios of violence was derived from the more general management strategy known as scenario planning, which has been used for more than 50 years in such fields as business, health care, and the military (Ringland, 1998; van der Heijden, 1997). According to Chermack and Lynham (2002, p. 366), "Scenario planning is a process of positing several informed, plausible and imagined alternative future environments in which decisions about the future may be played out, for the purpose of changing current thinking, improving decision making, enhancing human and organization learning and improving performance." It is most appropriate for situations of complexity and unbounded uncertainty (e.g., van der Heijden, 1994)—in our view, an accurate characterization of the state of affairs in which many violence risk assessments are conducted.

The SPJ approach encourages evaluators to consider four broad scenarios of violence.

1. The evaluator considers a scenario in which the examinee commits violence similar to his or her most recent act—what might be called a repeat, "flat trajectory," linear projection, or point projection scenario.
2. The evaluator considers a scenario in which the trajectory of violence decreases, and the examinee commits a less serious act or even chooses to desist altogether—a best-case or optimistic scenario.
3. The evaluator considers a worst-case scenario, also known as a pessimistic or "doom" scenario, one in which the trajectory increases, and the examinee commits a more serious, and perhaps even life-threatening, act of violence.
4. The evaluator considers a "twist" or "sideways trajectory" scenario in which the nature of violence changes or evolves, such as with respect to the manner of victim selection or the type of coercion used.

Multiple scenarios could be developed within each of these four broad categories. Then, for each scenario, the evaluator develops a detailed description in terms of the nature, severity, imminence, frequency or duration, and likelihood of violence. Finally, the evaluator trims the scenarios that seem implausible based on theory, research, experience, and case facts. In our experience, three to five general scenarios usually are sufficient to capture the range of plausible outcomes in a given case.

Step 5: Consider Management Plans. Next, consistent with general principles of scenario planning, evaluators develop case management plans based on plausible scenarios of violence, which in turn were based on the presence and relevance of risk factors. It is common in the SPJ approach to structure development of plans by considering four general categories of strategies: monitoring, supervision, treatment, and victim safety planning (e.g., Hart, Douglas, & Webster, 2001; Hart et al., 2003; Kropp, Hart, & Lyon, 2008). Within each general category of strategies, evaluators identify specific strategies, then translate these into more detailed plans by considering tactics and even logistics.

Step 6: Communicate Findings. In Step 6, evaluators should document and communicate their judgments regarding the overall risk in the case. Evaluators are encouraged to make judgments concerning such things as case prioritization or overall risk for violence, risk for serious physical harm, any indication of other risks the examinee may pose, any immediate actions taken or required, and critical dates or triggers for case review. It is during this step that evaluators can assign ratings of low, moderate, or high risk (i.e., for any violence; serious violence; imminent violence; etc.)—the summary risk ratings discussed in detail earlier. These are the key, encapsulating judgments that professionals offer in order to inform risk management. As our research review demonstrated, they are as or more accurate than actuarial estimates. Critically, as discussed earlier, the SPJ approach does not make these judgments using a fixed and explicit algorithm based on some combination of risk factors; instead, evaluators use their discretion to consider, decide, and explain the relevance or meaningfulness of any factors that are present with respect to the risks posed and management of those risks. Also, evaluators are encouraged to limit or qualify their opinions based on the quantity and quality of case information they reviewed.

ADDITIONAL CONSIDERATIONS

Four additional themes consistently emerge from both the literature and court decisions regarding what a reasonably competent assessment of risk entails, once clinicians are aware (or ought to be aware) that there exist reasonable grounds to believe patients may pose a risk for violence. When *Tarasoff* liability has been imposed on clinicians (at least in reported cases), it has usually been because of the *failure* of the clinician to abide by one or more of these themes (see our “Violence

Risk Assessments and the Law” section; see also, e.g., Monahan, 1993; *Peck v. The Counseling Service of Addison County*, 1985).

1. *Throughout the process, clinicians should make all reasonable efforts to obtain details of the patient’s history of violence and response to treatment for violence.* For any reasonably comprehensive assessment, “the painstaking assembling of facts and the checking of information from a variety of sources are essential” (Prins, 1988, p. 600). The examinee’s versions of events should be checked against collateral information, including police and victim versions of events.
2. *Clinicians must be alert to their own possible tendencies to avoid, deny, or wishfully minimize violent (or violence-related) themes and affects* (Kutzer & Lion, 1984, p. 71). That is, if the possibility of violence risk arises in a clinical context, it must be investigated in a reasonably full and complete manner. In our view, this entails the use of comprehensive violence risk assessment procedures, such as those provided under the SPJ approach. The use of such an approach minimizes the chances that important risk factors will be overlooked.

When patients appear to be at risk for violence, they should be asked if they are thinking of harming anyone and, if so, how they have dealt with such thoughts and feelings in the past (Beck, 1990). Appelbaum and Gutheil (1991) suggested that it is often useful to ask examinees whether they have ever caused the death or serious injury of another person, even if accidentally. Borum, Swartz, and Swanson (1996) suggested: “Are you the sort of person who has trouble controlling your temper?” (p. 209) and “Have you found yourself hitting people or damaging things when you are angry?” (p. 209). (See also Mulvey, Shaw, & Lidz, 1994.) Monahan (1993) observed:

Directly asking patients about violent behavior and possible indices of violent behavior (e.g., arrest or hospitalization as “dangerous to others”) is surely the easiest and quickest way to obtain this essential information. Open-ended questions such as “What is the most violent thing you have ever done?” or “What is the closest you have ever come to being violent?” may be useful probes, as might “Do you ever worry that you might physically hurt somebody?” The obvious problem, of course, is that patients may lie or distort their history or their current thoughts. . . . Quite often, however, patients are remarkably forthcoming about violence. (p. 244)

3. *Consider the circumstances the examinee may be facing in the future (including whether the examinee will be in secure confinement).* Are these circumstances similar to those that have led to violence in the past (e.g., a discordant family situation)? Or are these circumstances that have reduced the risk of violence in the past (e.g., a supportive family environment or social network or protective conditions of confinement)? If the examinee will be returning to a less restrictive environment, does he or she demonstrate a meaningful understanding of how to avoid violence in the future? Has the examinee

demonstrated a commitment to avoiding violence in the future? In general, it is worth noting Lion's (1987, p. 5) observation that an "appearance of tranquility in a [recently] violent person can be deceptive. Discharging the patient can be an error when the problem has not been really resolved. The clinician must consider whether anything has really changed."

4. *When in doubt, consult.* Indeed, when and where feasible, even when not in doubt, consult. Simply put, it is both ethically and legally advisable to obtain a second, knowledgeable opinion about what to do when one is uncertain whether a patient poses a serious risk of causing serious harm to another person or persons (or self).

We make the next additional recommendations that are useful to consider in fulfilling the demands of comprehensive risk assessments that are ethically and legally defensible.

As the violent history of the examinee becomes more distant in time, more effort may be required to accurately reconstruct the details, but such efforts should be made. At the same time, *all* potential sources of information regarding the examinee's former violence, current behavior, and mental status should be considered.

In institutional contexts, it is generally sensible to have patients assessed by clinicians who are not attached to the patient's treatment and care before final decisions or recommendations are made. This recommendation—offered to facilitate objectivity in assessment—is consistent with ethical guidelines for mental health professionals (see, e.g., American Psychological Association, 2010, 2013; the Specialty Guidelines are reprinted as the appendix to this volume with permission of the APA). Treating clinicians, and ward staff in general, may become so invested in believing that particular patients under their care have made adequate progress, or may so want to avoid disrupting the relative equilibrium achieved by a formerly more disorganized patient, or may so want to support the aspirations of a well-liked patient, that they avoid seeing negative signs and confronting the patient with difficult but necessary questions (Prins, 1988, 1996). Conversely, a patient who is uncooperative or challenging toward staff on the ward may be viewed as more dangerous than he or she really is. Therefore, except in obvious cases, a more detached evaluator may be called for.

When inpatients pose a risk of serious violence, recommendations for release should not be made without subjecting the patient to stressful questions regarding the sources of his or her previous violence and what the patient needs to do to avoid violence in the future. As Borum et al. (1996) observed:

Many potentially violent patients can and will appear calm and nonthreatening when not challenged, frustrated, or irritated. The clinician needs to be able to gingerly increase the frustration or challenge in the interview to test the frustration tolerance and impulse control of the patient without precipitating a dangerous outburst. A highly structured, unchallenging interview can dull the examiner into underestimating the violent potential of the patient. (p. 211)

In certain circumstances, with certain patients, we would make the point even more strongly: Sometimes it is precisely the question or challenge that *will* precipitate a “dangerous outburst” that must be posed (although, as Borum et al. [1996] emphasized, always in conditions of safety for the clinician). Indeed, sometimes the clinician’s inner sense that certain areas should be avoided lest the patient become overly disturbed is the best guide to determining what areas *require* further exploration (Glasser, 1996). Of course, there is no legitimate reason to provoke a patient needlessly. Thus, there may be no reason to subject a patient to a stressful interview if it is clear from other information that less secure supervision is not in order in any event. But when a patient with a history of serious violence is being considered for transfer to a less secure setting, it is hard to imagine concluding that transfer is in order without at least determining how the patient reacts to stressful questions regarding the circumstances he or she is likely to face in a new setting.

Even when stressful questions are indicated, clinicians should be careful not to be excessively, or needlessly, aggressive in their manner or confrontations. And clinicians should take reasonable steps to calm a disturbed patient once a stressful interview has ended. In any event, however, in order to pose suitably stressful questions to the patient, the assessor must be knowledgeable beforehand about the particular vulnerabilities of the patient. That is, in order to conduct an adequate stress interview, the clinician must first thoroughly inform him- or herself about what has led to and triggered violence or regressions in the patient in the past. Indeed, it cannot be overly stressed that comprehensive risk assessments can be accomplished only if the assessor is first well versed about the examinee’s history. It has often been said of legal trial practice that preparation is the key to cross-examination. Equally so, preparation is the key to a fully comprehensive risk assessment.

A number of commentators have suggested that assessments of violence risk can be improved if assessors avoid making errors that have frequently been observed to pertain to dangerousness evaluations (Ennis & Litwack, 1974; Hall, 2002; Monahan, 1981; Pfohl, 1978; Webster & Menzies, 1989; Webster & Polvi, 1995). Among the relevant recommendations are that the clinicians should:

1. *Recognize their own legitimate doubts and legitimate disagreements between evaluators and among staff regarding an examinee’s risk level to avoid unjustifiably confident determinations of dangerousness in either direction.* Clinicians should avoid a judgmental perspective or aligning themselves too much with the patient’s wishes and desires. They should be open to information that contradicts their initial, or even stated, opinion. Indeed, as Appelbaum and Gutheil (1991) stressed, it may be useful in many forensic evaluations for clinicians to imagine that they were retained as an expert by the “other side” of the case, and to imagine how their evaluation or conclusions might be different if that were so.
2. *Be careful not to underestimate the potential for violence in female patients who have a history of violence.* It is true that, on the whole in our society, women

have a much lower rate of violence than men, yet research indicates that women *who evidence risk factors for violence* are about as likely as men to commit violence (Newhill, Mulvey, & Lidz, 1995) and that, relative to their estimations regarding men, clinicians tend to underestimate the risk of future violence with women (Coontz et al., 1994; McNiel & Binder, 1995).

With these guidelines and recommendations in mind, we now turn to a discussion of how violence risk assessment has fared in the courts.

VIOLENCE RISK ASSESSMENTS AND THE LAW

In this section, we review certain important legal developments concerning assessments of dangerousness by mental health professionals.

ASSESSMENTS OF DANGEROUSNESS AND THE SUPREME COURT

Despite qualms expressed by various professional organizations and legal commentators, and some of their own brethren (all well summarized by Faigman, Kaye, Saks, & Sanders, 1997), the Supreme Court of the United States has been receptive to assessments of dangerousness and risk (and even predictions of violence) by mental health professionals in a variety of circumstances. Most notable perhaps (and perhaps most notorious), in *Barefoot v. Estelle* (1983), by a 6–3 vote, the Supreme Court upheld the constitutionality of a sentence of death that was based in part on the prediction of a testifying psychiatrist (who had not interviewed the defendant) that the “probability” that the defendant would commit additional crimes of violence in prison if not executed was “one hundred percent and absolute” (*Texas v. Barefoot*, Record at 2131; quoted in Appelbaum, 1984, p. 169).

Although not unmindful of the questions that existed regarding the validity of predictions of violence by mental health professionals, in *Barefoot*, the majority opined that such questions could adequately be dealt with by the trier of fact, “We are not persuaded,” stated the court, “that such testimony is almost entirely unreliable [i.e., invalid] and that the factfinder and the adversary system will not be competent to uncover, recognize and take due account of its shortcomings” (463 U.S. at 899). In *Schall v. Martin* (1984), the Supreme Court observed that “from a legal point of view there is nothing inherently unattainable about a prediction of future criminal conduct” and that the lower court had “specifically rejected the contention, based on . . . sociological data . . . , ‘that it is impossible to predict future violent behavior’” (467 U.S. at 278–279). More recently, U.S. Supreme Court Justice Stevens reasoned in his dissent in *United States v. Scheffer* (1998) that “[t]here is no legal requirement that expert testimony must satisfy a particular degree of reliability to be admissible. Expert testimony about a defendant’s ‘future dangerousness’ to determine his eligibility for the death penalty, even if wrong ‘most of the time,’ is routinely admitted” (523 U.S. at 334).

Similarly, in *Kansas v. Hendricks* (1997) and *Kansas v. Crane* (2002), the Court voiced approval of risk assessment in the sexual predator context. In *Hendricks*, the Supreme Court determined the constitutionality of a sexually violent predator statute, permitting the post-incarceration civil commitment of sex offenders. The intent of the statute was to civilly detain people who presented a risk to public safety due to tendencies to perpetrate sexual violence. As with other types of civil commitment, a determination of future dangerousness was required by the statute. Future dangerousness could be determined by past sexually violent behavior, such as the crime for which the individual was incarcerated, and a connection between that behavior and a mental “abnormality” reducing the individual’s control over the dangerous behavior. The Supreme Court determined that post-conviction civil commitment in this manner did not violate substantive due process rights under the Constitution.

In *Kansas v. Crane* (2002), the Court further interpreted this statute and held that the lack of control over the sexually violent behavior did not have to be absolute. In its reasoning, the Court continued to demonstrate amenability toward risk assessment and a reliance on mental health professionals in stating that the lack of behavioral control, “when viewed in light of such features of the case as *the nature of the psychiatric diagnosis* and the severity of the mental abnormality itself, must be sufficient to distinguish the dangerous sexual offender whose serious mental illness, abnormality, or disorder subjects him to civil commitment from the dangerous but typical recidivist convicted in an ordinary criminal case” (p. 413, emphasis added).

RECENT DEVELOPMENTS CONCERNING THE ADMISSIBILITY OF TESTIMONY ABOUT VIOLENCE RISK

Our own survey of state and federal cases considering the admissibility of expert testimony on risk assessment reveals that it is rarely excluded. Testimony on risk assessment can be based on clinical interviewing techniques, reliance on results from actuarial instruments, or a combination of these techniques. Clinical assessments of dangerousness have long been admissible under the *Frye* test (*Frye v. United States*, 1923), which still prevails in many states and which allows for expert testimony regarding the results of professional procedures when those procedures have gained “general acceptance” in their field. Moreover, little has changed in the federal courts since *Frye* was supplanted by *Daubert* (1993). When testimony is based on the results of a clinical interview alone, courts typically admit the testimony, stating that clinical interviewing techniques are generally accepted (*People v. Ward*, 1999). However, some courts have determined that reliance only on a clinical interview and failure to use actuarial instruments renders expert testimony unreliable (*Coble v. State*, 2010).

When the testimony is based in part on the results of actuarial instruments, typically it is also admitted. Many courts have indicated that testimony based in part on actuarial instruments is admissible because it is combined with clinical

opinion (*Lee v. State*, 2003). Courts have not specifically stated that testimony based solely on the results of actuarial instruments would be excluded. However, this has been implied by court arguments that the testimony is admissible only because combined with clinical judgment. For example, one court stated:

By this ruling, we are not concluding that actuarial risk assessment instruments are reliable per se or have our approval when used alone and not in conjunction with a full clinical evaluation. We note this was not the situation or issue presented in the instant case. The instruments *were* used in conjunction with a full clinical evaluation and their limitations were clearly made known to the jury. (*In re Detention of Holtz*, 2002, p. 619, emphasis in original; see also *People v. Stevens*, 2004)

Regardless of the basis for the testimony, courts have rarely given substantive consideration to the reliability of risk assessment in their admission decisions, even after the *Daubert* and *Kumho Tire* (1999) opinions. In their assessment of reliability, courts have taken several approaches. Typically, courts argue that a *Frye* or *Daubert* evaluation of the evidence is inappropriate. Many courts draw a distinction between clinical observations, which have a medical basis, and the use of actuarial instruments, which have a scientific basis (*People v. Ward*, 1999). When this distinction is made, it is reasoned that a reliability analysis does not apply to the clinical observations because they are not scientific or novel, they are based on standard psychological analysis of behavioral observations, and they are based on experience (*In re Commitment of R. S.*, 2001; *In re Detention of Berry*, 2011; *Westerheide v. State*, 2000; see generally *Logerquist v. McVey*, 2000).

Courts are more likely to subject actuarial tools to *Frye* or *Daubert* analyses, but the application of these standards to actuarial instruments varies across jurisdictions. As noted, some courts have argued that *Frye* and *Daubert* do not apply to the scientifically based actuarial instruments when they are combined with clinical interviewing, but that they would be if presented alone. They reason that the instruments do not act as a source of “infallible truth” or “scientific infallibility” when combined with clinical judgment (*People v. Stevens*, 2004; *People v. Therrien*, 2003; *State ex rel. Romley v. Fields*, 2001). Other courts have argued that a *Frye* or *Daubert* analysis must be conducted when any part of the testimony is based on the actuarial instruments because they are scientific (*In re Commitment of R. S.*, 2001; *In re Detention of Hargett*, 2003). These courts have largely determined that the use of actuarial instruments in risk assessment is generally accepted and reliable, and the testimony is admitted (*In re Commitment of R. S.*, 2001; *In re Commitment of Simmons*, 2004; *In re Detention of Campbell*, 1999; *In re Detention of Thorell*, 2003; *People v. Stevens*, 2004).

Some courts have pronounced the instruments generally accepted and their results admissible even without a *Frye* hearing (*In re Detention of Strauss*, 2001). Regarding the actuarial instruments, one court concluded that they were “at least as good, if not in most cases better, in terms of reliability and predictability than clinical interviews” (*In the Matter of Registrant C. A.*, 1996, p. 106). Some courts

that previously determined that testimony based on actuarial instruments was inadmissible because the tools were not generally accepted (reasoning that they were still in an experimental phase; their reliability and validity was not yet established, as evidenced by validation on a limited sample, lack of replication, lack of peer review, and scoring inconsistencies) were later overruled by courts finding that they are generally accepted (*People v. Taylor*, 2002, overruled by *In re Commitment of Simmons*, 2004).

However, some courts continue to reason that actuarial instruments are irrelevant and speculative in demonstrating lack of propensity for future risk of violence (*Leon v. McDonald*, 2011). Other courts concluding that testimony based on actuarial instruments satisfies *Frye* when used to predict recidivism require *Frye* hearings when these tools are used for other purposes, such as demonstrating the existence of a mental illness (*In the Matter of State v. Rosado*, 2009). Most of these decisions have arisen from *Frye* jurisdictions, but reasoning about these cases in *Daubert* jurisdictions does not differ substantially. The *Daubert* factors are rarely applied to this type of testimony—or are applied loosely (see, e.g., *United States v. Barnette*, 2000, admitting testimony based in part on the PCL-R and actuarial data under *Daubert*). Finally, many courts decline to assess reliability for the purpose of admission decisions, reasoning that questions of reliability go to the weight, not to the admissibility, of the evidence (*In re Detention of Holtz*, 2002), or have determined that *Frye* and *Daubert* do not apply to the legal context of the case, such as sentencing (see *United States v. Fields*, 2007, following *Barefoot*). In short, whether in *Frye* or in *Daubert* jurisdictions, courts have been receptive to violence risk assessments, whether clinical or actuarial, by mental health professionals. (For an argument in favor of the admissibility of actuarial risk assessments and for a review of other relevant cases, see Janus & Prentky, 2003.)

CIVIL COMMITMENT DECISIONS

In *Addington v. Texas* (1979), the U.S. Supreme Court ruled that individuals could be involuntarily committed to a mental hospital for an *extended* period of time only if there was “clear and convincing evidence” that they met a legitimate legal standard for confinement. Therefore, to the extent that dangerousness as well as mental illness is required by law to justify an extended commitment, *Addington* requires that such dangerousness be proven by “clear and convincing evidence” (more than a “preponderance of the evidence” but less evidence than is required for proof “beyond a reasonable doubt”). The Court in *Addington* also observed, “Whether the individual is mentally ill and dangerous to either himself or others and is in need of confined therapy turns on the meaning of the facts which *must be interpreted by expert psychiatrists and psychologists*” (441 U.S. at 429, emphasis added).

There are two points to note about the *Addington* decision. First, it applied only to extended confinements. Therefore, presumably (and as is current practice), less than “clear and convincing evidence” of mental illness and dangerousness could

justify a relatively brief commitment for the purpose of further evaluation. Second, it is important to recognize that the requirement of “clear and convincing evidence” of dangerousness, when it exists, is not a requirement of proof that the individual is more likely than not to be violent if not hospitalized. Rather, it is a requirement for “clear and convincing evidence” *of enough risk of enough harm to justify the confinement at issue* (see, e.g., Monahan & Silver, 2003; *Rogers v. Okin*, 1980). Indeed, in a 1990 decision, the Ninth Circuit Court of Appeals ruled that “a finding of [a] ‘substantial risk’ [of violence sufficient to justify an extended civil commitment] may be based on any activity that evinces a *genuine possibility* of future harm to persons or property” (*United States v. Sahhar*, 1990, 917 F.2d at 1207, emphasis added). And the court also rejected the notion that, to be constitutional, a civil commitment must be based on a *recent* overt act or threat of violence. Rather, the court stated, “Whether [worrisome] activity occurred recently is but one factor . . . to consider in weighing the evidence” (p. 1207; see also *United States v. Evanoff*, 1993; *United States v. Williams*, 2002).

The decision in the *Sahhar* case is emblematic of a recent trend away from the strict standards for civil commitments that were established in some jurisdictions in the 1970s to more flexible (and perhaps more realistic) criteria. This trend can best be appreciated by comparing more recent decisions with the rulings in *Lessard v. Schmidt* (1972), an often-cited federal court decision of a generation ago that strictly limited the government’s power to civilly commit mentally ill individuals. In *Lessard*, the court ruled that a commitment could be justified only by proof of “an *extreme likelihood* that if the person is not confined he will do *immediate harm* to himself or others” (349 F. Supp. 1078, emphasis added).

Moreover, the *Lessard* court held that the necessary determination of dangerousness had to be “based upon a finding of a *recent* overt act, attempt, or threat to do substantial harm to oneself or another” (p. 1078, emphasis added). In addition to rejecting the notion that violence must be likely or based on a recent overt act or threat in order to justify a commitment, other court decisions have rejected the notion that future dangerousness must be “imminent” to justify confinement. For example, in 1991, the Supreme Court of Massachusetts held that “to the degree that the anticipated harm is serious . . . some lessening of a requirement of imminence seems justified” (*Commonwealth v. Rosenberg*, 1991, 573 N.E. 2d at 958); and in *Seltzer v. Hogue* (1993), a New York State appellate court upheld the continued confinement of a mentally ill person who in the past had “invariably” become violent following his release from hospitalization, even though he had not been “imminently” violent upon release. Thus, as one of us (Litwack, 1993) has written elsewhere:

It appears that as the earlier abuses of the civil commitment system (see, e.g., *O’Connor v. Donaldson*, 1975) are supplanted in judicial and public concern by concern about potential violence by mentally ill persons who perhaps could not be committed under a strict reading of earlier and more libertarian oriented decisions, the judicial pendulum is swinging toward a greater willingness to allow civil commitments to protect the public from potential danger, and to allow that danger to be assessed broadly, rather than by rigid, impractical rules. (p. 363)

Moreover, it should be pointed out, the U.S. Supreme Court has ruled that insanity acquittees may be required to prove that they are no longer dangerous before being released from confinement (*Jones v. United States*, 1983). And in the case of *In the Matter of George L.* (1995), the New York Court of Appeals approvingly quoted from an earlier decision regarding insanity acquittees that “compliance or lack of dangerousness in a facility does not necessarily mean that an individual does not suffer from a dangerous mental disorder” (624 N.Y.S. 2d at 103). Even more recently, regarding the retention of insanity acquittees, the New York Court of Appeals opined:

In addition to recent acts of violence . . . a court may consider the nature of the conduct that resulted in the initial commitment, the likelihood of relapse or cure, history of substance or alcohol abuse, the effects of medication, the likelihood that the patient will discontinue medication without supervision, the length of confinement and treatment, the lapse of time since the underlying criminal acts *and any other relevant factors that form a part of an insanity acquittee’s psychological profile.* (*In the Matter of David B.*, 2002, p. 279, emphasis added)

Of course, it is ultimately for the courts, rather than clinicians, to decide when mentally disordered individuals pose a sufficient risk of causing significant harm to justify depriving them of their liberty. But emergency room clinicians must make such decisions regarding emergency admissions, and judges frequently look to clinicians for their input and insights before making their decisions. Therefore, before clinicians deem a patient to be dangerous for commitment purposes, they should be mindful of the fact that they are, indeed, making a *risk* assessment—and that, just as a sufficient risk (rather than a certainty) of future violence may justify a patient’s confinement, so, too, a patient’s right to liberty should be weighed in the balance.

THE TARASOFF CASE AND THE DUTY TO PROTECT

In the well-known case of *Tarasoff v. Regents of the University of California* (1976), the Supreme Court of California ruled:

Once a therapist does in fact determine, or under applicable professional standards reasonably should have determined, that a patient poses a serious danger of violence to others, he bears a duty to exercise reasonable care to protect the foreseeable victim of that danger. (17 Cal. 3d at 439)

The *Tarasoff* ruling has been followed in most states (though with significant variations from state to state) and, as Monahan (1993) observed in a seminal article:

In jurisdictions in which appellate courts [or legislatures] have not yet ruled on the question, the prudent clinician is well advised to proceed under the assumption that some version of *Tarasoff* liability will be imposed. . . . The duty to protect, in short, is

now a fact of professional life for nearly all American clinicians and, potentially, for clinical researchers as well. (p. 242)

The case law, research, and voluminous commentary that have followed from the *Tarasoff* decision cannot be reviewed here; nor can the various laws and professional regulations defining *Tarasoff*-like duties that have been adopted in many states, or the differing responsibilities held by inpatient and outpatient clinicians. For examples, see Binder and McNeil (1996); *Emmerich v. Philadelphia Center for Human Development* (1998); *Fraser v. U.S.* (1996); and N.J. Stat. §21:62A-16 (2001). For recent commentary, see Bersoff (2008); Borum and Reddy (2001); Buckner and Firestone (2000); Felthous and Kachigian (2001); Gutheil and Brodsky (2010) (for *Tarasoff's* application to medical examinations); Pinta (2010) (for *Tarasoff's* application to a prison population); Simone and Fulero (2005); Tolman (2001); VandeCreek and Knapp (2001); Walcott et al. (2001); and Werth, Welfel, and Benjamin (2009). Suffice it to say for the purposes of this chapter that clearly one component of the duty to protect is the duty to conduct a professionally adequate risk assessment when such an assessment is called for.

CONCLUSION AND FUTURE DIRECTIONS

The field of violence risk assessment has made considerable strides in recent years. Indeed, since the publication of the previous version of this chapter, several hundred empirical studies have been published on violence risk assessment instruments. We have indeed made progress in terms of understanding how to develop and validate risk assessment instruments and how instruments from different families fare relative to one another. Perhaps one of the biggest developments is that, contrary to previous opinion (Meehl, 1954; Quinsey et al., 2006), clinical judgments of risk—so long as they are derived in a structured context, such as that provided by the SPJ model—are as or more accurate compared to actuarial predictions of violence. This is a major and liberating finding that can facilitate focus on other important topics within risk assessment. For instance, there has been great conceptual progress on the *process* of risk assessment at the individual case level (where it takes place in practice)—this topic would benefit greatly from a comparable amount of empirical attention (Litwack, 2002; Mulvey & Lidz, 1985). For instance, how do clinicians decide which risk factors are most relevant in a given case? How do they integrate this information into formulations? Does this improve the quality of risk management?

The risk assessment field could benefit from turning its attention to several other areas of focus in the future. First, although commentators have been stressing the importance of dynamic risk for some time, the empirical evaluation of changes in risk factors over time lags (Douglas & Skeem, 2005). This is understandable—it is hard research to conduct. Yet many questions remain: Which factors change? What is the nature of change? Does change relate to changes in violence? Can we impact

change deliberately? For which risk factors? Sustained attention to dynamic risk will allow the field to continue to move toward a risk reduction or risk management focus. That is, studies of whether risk assessment instruments can be used to assign persons to treatment intensities, with specified relevant risk factors, are necessary. There is evidence from the correctional field that such a focus could indeed reduce violence (Andrews, 2012).

There has been a movement over the past decade or so to increase the focus on protective factors (sometimes called strengths, resiliency, or buffers). As yet, the field of forensic mental health has not come to terms with the conceptual nature of protective factors (how should they be defined, exactly?) and how they interact with risk factors. Are they distinct from risk factors? Or are they the opposite pole of risk factors? Are they relevant in the absence of risk factors, or only when risk factors are present? Regardless of the answer to these important conceptual questions, it is clear that important strides are under way to understand the role that focusing on positive aspects of our clients' lives might have. For example, several assessment instruments focus on positive aspects of clients. The START requires ratings of patient strengths along with vulnerabilities. The SAVRY contains a subscale devoted to protective factors. The SAPROF is devoted entirely to protective factors and is intended to be used in conjunction with an SPJ risk assessment instrument such as the HCR-20. Empirical support for the utility of each of these measures is accruing (as reviewed earlier), and we encourage the continued attention to this topic.

Although hardly a new concern, we advocate continued efforts to test and ensure that risk (and protective) assessment instruments can be used ethically with persons of all ethnic backgrounds, whether male or female, and regardless of sexual orientation. Evidence drawn from meta-analyses suggests that it is unlikely that there are major problems in this regard (although to our knowledge there has not yet been any data on point concerning sexual orientation). Some concerns were evident in the findings reported by Singh et al. (2011). Further, it is important to understand whether, even if a risk assessment instrument predicts violence comparably well across, say, men and women, its risk factors have comparable meaning or relevance across gender (Garcia-Mansilla, Rosenfeld, & Nicholls, 2009). To some extent, this question can be integrated into the formulation stage of clinical assessment, where clinicians are grappling with making sense of the meaning of certain risk factors for individuals. Continued work on this topic is necessary.

Finally, we think it is important that researchers expand their topics beyond those concerning the predictive accuracy of instruments. Assessing and improving predictive accuracy is crucial, to be sure. However, as described earlier concerning the process of risk assessment, there are other important topics to understand. One such topic is related to both assessing and improving predictive validity: To what extent might postrelease (or within-institution) intervention or management moderate, and possibly improve, predictive accuracy? We want our assessments to influence practice and reduce risk (e.g., persons evaluated to be high risk

should receive higher doses of treatment). However, if we study accuracy issues within real-world settings, where risk assessments may influence interventions, those interventions interfere with the outcomes that we are interested in (see, e.g., the discussion about threats of violence by Litwack et al., 2006). Hence, research should attempt to build expected treatment or intervention dosage into predictive models. One such example was recently provided by Belfrage et al. (2012), who demonstrated that persons judged to be higher risk on the SARA by police, who also received higher doses of risk management, were less likely to recidivate in the future relative to high-risk persons who received lower doses of management.

Although great strides have been made in the risk assessment field, it is clear that much important work remains.

APPENDIX: STUDIES INVESTIGATING WHETHER SUMMARY RISK JUDGMENTS ARE PREDICTIVE OF VIOLENCE

- Arbach-Lucioni, K., Andres-Pueyo, A., Pomarol-Clotet, E., & Gomar-Sones, J. (2011). Predicting violence in psychiatric inpatients: A prospective study with the HCR-20 violence risk assessment scheme. *Journal of Forensic Psychiatry & Psychology, 22*, 203–222.
- Belfrage, H., Strand, S., Storey, J. E., Gibas, A. L., Kropp, P. R., & Hart, S. D. (2012). Assessment and management of risk for intimate partner violence by police officers using the Spousal Assault Risk Assessment Guide. *Law and Human Behavior, 36*, 60–67.
- Braithwaite, E., Charette, Y., Crocker, A. G., & Reyes, A. (2010). The predictive validity of clinical ratings of the Short-Term Assessment of Risk and Treatability (START). *International Journal of Forensic Mental Health, 9*, 271–281.
- Catchpole, R. E. H., & Gretton, H. M. (2003). The predictive validity of risk assessment with violent young offenders: A 1-year examination of criminal outcome. *Criminal Justice and Behavior, 30*, 688–708.
- de Vogel, V., & de Ruiter, C. (2005). The HCR-20 in personality disordered female offenders: A comparison with a matched sample of males. *Clinical Psychology & Psychotherapy, 12*, 226–240.
- de Vogel, V., & de Ruiter, C. (2006). Structured professional judgment of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20. *Psychology, Crime & Law, 12*, 321–336.
- de Vogel, V., de Ruiter, C., Hildebrand, M., Bos, B., & van de Ven, P. (2004). Type of discharge and risk of recidivism measured by the HCR-20: A retrospective study in a Dutch sample of treated forensic psychiatric patients. *International Journal of Forensic Mental Health, 3*, 149–165.
- de Vogel, V., de Ruiter, C., van Beek, D., & Mead, G. (2004). Predictive validity of the SVR-20 and Static 99 in a Dutch sample of treated sex offenders. *Law and Human Behavior, 28*, 235–251.
- de Vries Robbé, M., de Vogel, V., & de Spa, E. (2011). Protective factors for violence risk in forensic psychiatric patients: A retrospective validation study of the SAPROF. *International Journal of Forensic Mental Health, 10*, 178–186.
- Desmarais, S. L., Nicholls, T., Wilson, C. M., & Brink, J. (2012). Reliability and validity of the Short-Term Assessment of Risk and Treatability (START) in assessing risk for inpatient aggression. *Psychological Assessment, 24*, 685–700.

- Dolan, M. C., & Rennie, C. E. (2008). The Structured Assessment of Violence Risk in Youth as a predictor of recidivism in a United Kingdom cohort of adolescent offenders with conduct disorder. *Psychological Assessment, 20*, 35–46.
- Douglas, K. D., Ogloff, J. R. P., & Hart, S. D. (2003). Evaluation of a model of violence risk assessment among forensic psychiatric patients. *Psychiatric Services, 54*, 1372–1379.
- Douglas, K. S., Yeomans, M., & Boer, D. P. (2005). Comparative validity analysis of multiple measures of violence risk in a sample of criminal offenders. *Criminal Justice and Behavior, 32*, 479–510.
- Enebrink, P., Långström, N., & Gumpert, C. H. (2006). Predicting aggressive and disruptive behavior in referred 6- to 12-year-old boys: Predictive validation of the EARL-20B risk/needs checklist. *Assessment, 13*, 356–367.
- Gammelgård, M., Koivisto, A., Eronen, M., & Kaltiala-Heino, R. (2008). The predictive validity of the Structured Assessment of Violence Risk in Youth (SAVRY) among institutionalised adolescents. *Journal of Forensic Psychiatry & Psychology, 19*, 352–370.
- Jovanović, A. A., Toševski, D. L., Ivković, M., Damjanović, A., & Gašić, M. J. (2009). [Predicting violence in veterans with posttraumatic stress disorder]. *Vojnosanitetski Pregled: Military Medical & Pharmaceutical Journal of Serbia & Montenegro, 66*, 13–21.
- Kropp, P. R., & Hart, S. D. (2000). The Spousal Assault Risk Assessment (SARA) guide: Reliability and validity in adult male offenders. *Law and Human Behavior, 24*, 101–118.
- Langton, C. M., Hogue, T. E., Daffern, M., Mannion, A., & Howells, K. (2009). Prediction of institutional aggression among personality disordered forensic patients using actuarial and structured clinical risk assessment tools: Prospective evaluation of the HCR-20, VRS, Static-99, and Risk Matrix 2000. *Psychology, Crime & Law, 15*, 635–659.
- Lodewijks, H. P. B., Doreleijers, Th. A. H., & Ruiter, C. de (2008). SAVRY risk assessment in relation to sentencing and subsequent recidivism in a Dutch sample of violent juvenile offenders. *Criminal Justice and Behavior, 35*, 696–709.
- Lodewijks, H. P. B., Doreleijers, Th. A. H., Ruiter, C. de, & Borum, R. (2008). Predictive validity of the Structured Assessment of Violence in Youth (SAVRY) during residential treatment. *International Journal of Law and Psychiatry, 31*, 263–271.
- Lodewijks, H. P. B., de Ruiter, C., & Doreleijers, Th. A. H. (2008). Gender differences in violent outcome and risk assessment in adolescent offenders after residential treatment. *International Journal of Forensic Mental Health, 7*, 133–146.
- Meyers, J. R., & Schmidt, F. (2008). Predictive validity of the Structured Assessment for Violence Risk in Youth (SAVRY) with juvenile offenders. *Criminal Justice and Behavior, 35*, 344–355.
- Neves, A. C., Goncalves, R. A., & Palma-Oliveira, J. M. (2011). Assessing risk for violent and general recidivism: A study of the HCR-20 and the PCL-R with a non-clinical sample of Portuguese offenders. *International Journal of Forensic Mental Health, 10*, 137–149.
- Nonstad, K., Nettet, M. B., Kroppan, E., Pedersen, T. W., Nøttestad, J. A., Almvik, R., & Palmstierna, T. (2010). Predictive validity and other psychometric properties of the Short-Term Assessment of Risk and Treatability (START) in a Norwegian high secure hospital. *International Journal of Forensic Mental Health, 9*, 294–299.
- Pedersen, L., Rasmussen, K., & Elsass, P. (2010). Risk assessment: The value of structured professional judgments. *International Journal of Forensic Mental Health, 9*, 74–81.

- Penney, S. R., Lee, Z., & Moretti, M. M. (2010). Gender differences in risk factors for violence: An examination of the predictive validity of the Structured Assessment of Violence Risk in Youth. *Aggressive Behavior, 36*(6), 390–404.
- Rajlic, G., & Gretton, H. (2010). An examination of two sexual recidivism risk measures in adolescent offenders. *Criminal Justice & Behavior, 37*, 1066–1085. doi:10.1177/0093854810376354
- Schaap, G., Lammers, S., & de Vogel, V. (2009). Risk assessment in female forensic psychiatric patients: A quasi-prospective study into the validity of the HCR-20 and PCL-R. *Journal of Forensic Psychiatry & Psychology, 20*, 354–365.
- Schmidt, F., Campbell, M. A., & Houlding, C. (2011). Comparative analyses of the YLS/CMI, SAVRY, and PCL:YV in adolescent offenders: A 10-year follow-up into adulthood. *Youth Violence and Juvenile Justice, 9*, 23–42.
- Sjöstedt, G., & Långström, N. (2002). Assessment of risk for criminal recidivism among rapists: A comparison of four different measures. *Psychology, Crime and Law, 8*, 25–40.
- van den Brink, R. H. S., Hooijschuur, A., van Os, T. W. D. P., Savenije, W., & Wiersma, D. (2010). Routine violence risk assessment in community forensic mental healthcare. *Behavioral Sciences & the Law, 28*, 396–410.
- Verbrugge, H. M., Goodman-Delahunty, J., & Frize, M. C. J. (2011). Risk assessment in intellectually disabled offenders: Validation of the suggested ID supplement to the HCR-20. *International Journal of Forensic Mental Health, 10*, 83–91.
- Viljoen, J. L., Scalora, M., Cuadra, L., Bader, S., Chávez, V., Ullman, D., & Lawrence, L. (2008). Assessing risk for violence in adolescents who have sexually offended: A comparison of the J-SOAP-II, J-SORRAT-II, and SAVRY. *Criminal Justice and Behavior, 35*, 5–23.
- Vincent, G. M., Chapman, J., & Cook, N. (2011). Predictive validity of the SAVRY, racial differences, and the contribution of needs factors. *Criminal Justice and Behavior, 38*, 42–62.

REFERENCES

- Addington v. Texas, 441 U.S. 418 (1979).
- Almvik, R., Woods, P., & Rasmussen, K. (2000). The Brøset Violence Checklist: Sensitivity, specificity, and interrater reliability. *Journal of Interpersonal Violence, 15*, 1284–1296.
- American Psychological Association. (2010). *Ethical principles of psychologists and code of conduct: With the 2010 amendments*. Washington, DC: Author.
- American Psychological Association. (2013). Specialty guidelines for forensic psychology. *American Psychologist, 68*(1), 7–19. doi:10.1037/a0029889
- Andrews, D. A. (1982). *The Level of Supervision Inventory (LSI): The first follow-up*. Toronto, Canada: Ontario Ministry of Correctional Services.
- Andrews, D. A. (2012). The risk-need-responsivity (RNR) model of correctional assessment and treatment. In J. A. Dvoskin, J. L. Skeem, R. W. Novaco, & K. S. Douglas (Eds.), *Using social science to reduce violent offending* (pp. 127–156). New York, NY: Oxford University Press.
- Andrews, D. A., & Bonta, J. (1995). *Level of Service Inventory—Revised*. Toronto, Canada: Multi-Health Systems.
- Andrews, D. A., & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Cincinnati, OH: Anderson.

- Andrews, D. A., Bonta, J., & Wormith, S. J. (2004). *The Level of Service/Case Management Inventory (LS/CMI)*. Toronto, Canada: Multi-Health Systems.
- Appelbaum, P. S. (1984). Hypothetical, psychiatric testimony, and the death sentence. *Bulletin of the American Academy of Psychiatry and the Law*, 12, 169–177.
- Appelbaum, P. S., & Gutheil, T. G. (1991). *Clinical handbook of psychiatry and the law* (2nd ed.). Baltimore, MD: Williams & Wilkins.
- Arbach-Lucioni, K., Andres-Pueyo, A., Pomarol-Clotet, E., & Gomar-Sones, J. (2011). Predicting violence in psychiatric inpatients: A prospective study with the HCR-20 violence risk assessment scheme. *Journal of Forensic Psychiatry & Psychology*, 22, 203–222.
- Augimeri, L. K., Koegl, C. J., Webster, C. D., & Levene, K. S. (2001). *Early assessment risk list for boys: Version 2*. Toronto, Canada: Earls Court Child and Family Centre.
- Barefoot v. Estelle, 463 U.S. 880 (1983).
- Baxstrom v. Herold, 383 U.S. 107 (1966).
- Beck, J. C. (1990). Clinical aspects of the duty to warn or protect. In R. Simon (Ed.), *Review of clinical psychiatry and the law* (Vol. 1, pp. 191–204). Washington, DC: American Psychiatric Press.
- Belfrage, H., Strand, S., Storey, J. E., Gibas, A. L., Kropp, P. R., & Hart, S. D. (2012). Assessment and management of risk for intimate partner violence by police officers using the Spousal Assault Risk Assessment Guide. *Law and Human Behavior*, 36, 60–67.
- Bersoff, D. (2008). *Ethical conflicts in psychology* (4th ed.). Washington, DC: American Psychological Association.
- Binder, R. L., & McNiel, D. E. (1996). Application of the Tarasoff ruling and its effect on the victims and the therapeutic relationship. *Psychiatric Services*, 47, 1212–1215.
- Blair, P. R., Marcus, D. K., & Boccaccini, M. T. (2008). Is there an allegiance effect for assessment instruments? Actuarial risk assessment as an exemplar. *Clinical Psychology: Science and Practice*, 15, 346–360.
- Boer, D. P., Hart, S. D., Kropp, P. R., & Webster, C. D. (1997). *Manual for the Sexual Violence Risk –20: Professional guidelines for assessing risk of sexual violence*. Vancouver, Canada; British Columbia Institute on Family Violence and Mental Health, Law, and Policy Institute, Simon Fraser University.
- Bonta, J., Law, M., & Hanson, R. K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin*, 123, 123–142.
- Borum, R. (1996). Improving the clinical practice of violence risk assessment: Technology, guidelines, and training. *American Psychologist*, 51, 945–956.
- Borum, R., Bartel, P., & Forth, A. (2006). *Manual for the Structured Assessment of Violence Risk in Youth (SAVRY)*. Odessa, FL: Psychological Assessment Resources.
- Borum, R., & Reddy, M. (2001). Assessing violence risk in Tarasoff situations: A fact-based model of inquiry. *Behavioral Sciences and the Law*, 19, 375–385.
- Borum, R., Swartz, M., & Swanson, J. (1996). Assessing and managing violence risk in clinical practice. *Journal of Practical Psychiatry and Behavioral Health*, 2, 205–215.
- Braithwaite, E., Charette, Y., Crocker, A. G., & Reyes, A. (2010). The predictive validity of clinical ratings of the Short-Term Assessment of Risk and Treatability (START). *International Journal of Forensic Mental Health*, 9, 271–281.
- Buckner, F., & Firestone, M. (2000). “Where the public peril begins”: 25 years after Tarasoff. *Journal of Legal Medicine*, 21, 187–222.

- Campbell, M. A., French, S., & Gendreau, P. (2009). The prediction of violence in adult offenders: A meta-analytic comparison of instruments and methods of assessment. *Criminal Justice and Behavior, 36*, 567–590.
- Chermack, T. J., & Lynham, S. A. (2002). Definitions and outcome variables of resource planning. *Human Resource Development Review, 1*, 366–383.
- Chermack, T. J., & van der Merwe, L. (2003). The role of constructivist learning in scenario planning. *Futures, 35*, 445–460.
- Coble v. State, 330 S.W.3d 253 (Tex. Crim. App., 2010).
- Commonwealth v. Rosenberg, 410 Mass. 347, 573 N.E. 2d 949 (1991).
- Coontz, P. D., Lidz, C. W., & Mulvey, E. P. (1994). Gender and the assessment of dangerousness in the psychiatric emergency room. *International Journal of Law and Psychiatry, 17*, 369–376.
- Copas, J., & Marshall, P. (1998). The Offender Group Reconviction Scale: The statistical reconviction score for use by probation officers. *Journal of the Royal Statistical Society, 47C*, 159–171.
- Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786 (1993).
- Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34*, 571–582.
- Derogatis, L. R. (1993). *BSI, Brief Symptom Inventory: Administration, scoring, and procedures manual* (4th ed.). Minneapolis, MN: National Computer Systems.
- de Vogel, V., & de Ruiter, C. (2005). The HCR-20 in personality disordered female offenders: A comparison with a matched sample of males. *Clinical Psychology & Psychotherapy, 12*, 226–240.
- de Vogel, V., & de Ruiter, C. (2006). Structured professional judgment of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20. *Psychology, Crime & Law, 12*, 321–336.
- de Vogel, V., de Ruiter, C., Bouman, Y., & de Vries Robbé, M. (2012). *SAPROF. Guidelines for the assessment of protective factors for violence risk* (2nd ed.). Utrecht, the Netherlands: Forum Educatief.
- de Vogel, V., de Ruiter, C., van Beek, D., & Mead, G. (2004). Predictive validity of the SVR-20 and Static 99 in a Dutch sample of treated sex offenders. *Law and Human Behavior, 28*, 235–251.
- Desmarais, S. L., Nicholls, T., Wilson, C. M., & Brink, J. (2012). Reliability and validity of the Short-Term Assessment of Risk and Treatability (START) in assessing risk for inpatient aggression. *Psychological Assessment, 24*, 685–700.
- Dix, G. E. (1980). Clinical evaluation of the “dangerousness” of “normal” criminal defendants. *Virginia Law Review, 66*, 523–581.
- Dixon v. Attorney General of the Commonwealth of Pennsylvania, 325 F.Supp. 966 (1971).
- Dolan, M. C., & Rennie, C. E. (2008). The Structured Assessment of Violence Risk in Youth as a predictor of recidivism in a United Kingdom cohort of adolescent offenders with conduct disorder. *Psychological Assessment, 20*, 35–46.
- Douglas, K. S., Cox, D. N., & Webster, C. D. (1999). Violence risk assessment: Science and practice. *Legal and Criminological Psychology, 4*, 149–184.
- Douglas, K. S., Hart, S. D., & Ogloff, J. R. P. (2003). Evaluation of a model of violence risk assessment among forensic psychiatric patients. *Psychiatric Services, 54*, 1372–1379.

- Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *Historical-Clinical-Risk Management-20 (Version 3): Violence risk assessment guidelines (HCR-20 V3)*. Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Douglas, K. S., & Kropp, P. R. (2002). A prevention-based paradigm for violence risk assessment: Clinical and research applications. *Criminal Justice and Behavior*, 29, 617–658.
- Douglas, K. S., & Ogloff, J. R. P. (2003a). The impact of confidence on the accuracy of structured professional and actuarial violence risk judgments in a sample of forensic psychiatric patients. *Law and Human Behavior*, 27, 573–587.
- Douglas, K. S., & Ogloff, J. R. P. (2003b). Multiple facets of risk for violence: The impact of judgmental specificity on structured decisions about violence risk. *International Journal of Forensic Mental Health*, 2, 19–34.
- Douglas, K. S., & Skeem, J. L. (2005). Violence risk assessment: Getting specific about being dynamic. *Psychology, Public Policy, and Law*, 11, 347–383.
- Douglas, K. S., Yeomans, M., & Boer, D. P. (2005). Comparative validity analysis of multiple measures of violence risk in a sample of criminal offenders. *Criminal Justice and Behavior*, 32, 479–510.
- Emmerich v. Philadelphia Center for Human Development, 720 A. 2d 1032 (1998).
- Enebrink, P., Långström, N., & Gumpert, C. H. (2006). Predicting aggressive and disruptive behavior in referred 6- to 12-year-old boys: Predictive validation of the EARL-20B risk/needs checklist. *Assessment*, 13, 356–367.
- Ennis, B. J., & Litwack, T. R. (1974). Psychiatry and the presumption of expertise: Flipping coins in the courtroom. *California Law Review*, 62, 693–752.
- Epperson, D. L., Kaul, J. D., Huot, S. J., Goldman, R., & Alexander, W. (2003). *Minnesota Sex Offender Screening Tool—Revised (MnSOST-R): Development, validation, and recommended risk-level cut scores* [Technical paper]. St. Paul, MN: Minnesota Department of Corrections.
- Epperson, D. L., Ralston, C. A., Fowers, D., DeWitt, J., & Gore, K. S. (2006). Actuarial risk assessment with juveniles who sexually offend: Development of the Juvenile Sexual Offense Recidivism Risk Assessment Tool-II (JSORRAT-II). In D. S. Prescott (Ed.), *Risk assessment of youth who have sexually abused* (pp. 118–169). Oklahoma City, OK: Wood ‘N’ Barnes.
- Faigman, D., Kaye, D., Saks, M., & Sanders, J. (1997). *Modern scientific evidence: The law and science of expert testimony*. St. Paul, MN: West.
- Faigman, D., & Monahan, J. (2009). Standards of legal admissibility and their implications for psychological science. In K. S. Douglas, J. L. Skeem, & S. O. Lilienfeld (Eds.), *Psychological science in the courtroom: Consensus and controversy* (pp. 3–25). New York, NY: Guilford Press.
- Fazel, S., Singh J. P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24,827 people: Systematic review and meta-analysis. *British Medical Journal*, 345:e4692 doi: 10.1136/bmj.e4692
- Felthous, A. R., & Kachigian, C. (2001). To warn and to control: Two distinct legal obligations or variations of a single duty to protect? *Behavioral Sciences and the Law*, 19, 355–373.
- Forth, A. E., Kosson, D. S., & Hare, R. D. (2003). *The Hare Psychopathy Checklist—Revised: Youth Version*. Toronto, Canada: Multi-Health Systems.
- Fraser v. U.S., 674 A. 2d 811 (Conn., 1996).
- Frye v. United States, 293 F. 1013 (DC Cir., 1923).
- Garcia-Mansilla, A., Rosenfeld, B., & Nicholls, T. (2009). Risk assessment: Are current methods applicable to women? *International Journal of Forensic Mental Health*, 8, 50–61.

- Gardner, W., Lidz, C. W., Mulvey, E. P., & Shaw, E. C. (1996a). Clinical versus actuarial predictions of violence in patients with mental illness. *Journal of Consulting and Clinical Psychology, 64*, 602–609.
- Gardner, W., Lidz, C. W., Mulvey, E. P., & Shaw, E. C. (1996b). A comparison of actuarial methods for identifying repetitively violent patients with mental illness. *Law and Human Behavior, 20*, 35–48.
- Glasser, M. (1996). The management of dangerousness: The psychoanalytic contribution. *Journal of Forensic Psychiatry, 7*, 271–283.
- Grann, M., & Långström, N. (2007). Actuarial assessment of violence risk: To weigh or not to weigh? *Criminal Justice and Behavior, 34*, 22–36.
- Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, and Law, 2*, 293–323.
- Gutheil, T., & Brodsky, A. (2010). Commentary: Tarasoff duties arising from a forensic independent medical examination. *Journal of the American Academy of Psychiatry and the Law, 38*, 57–60.
- Guy, L. S. (2008). *Performance indicators of the structured professional judgement approach for assessing risk for violence to others: A meta-analytic survey* (Unpublished doctoral dissertation). Simon Fraser University, Burnaby, British Columbia, Canada.
- Guy, L. S., Douglas, K. S., & Hendry, M. (2010). The role of psychopathic personality disorder in violence risk assessments using the HCR-20. *Journal of Personality Disorders, 24*, 551–580.
- Hall, H. V. (1982). Dangerous predictions and the maligned forensic professional: Suggestions for detecting distortions of true basal violence. *Criminal Justice and Behavior, 9*, 3–12.
- Hall, H. V. (2002). *Violence prediction: Guidelines for the forensic practitioner*. Springfield, IL: Charles C Thomas.
- Hanson, R. K. (1998). What do we know about sex offender risk assessment? *Psychology, Public Policy, and Law, 4*, 50–72.
- Hanson, R. K., & Howard, P. D. (2010). Individual confidence intervals do not inform decision-makers about the accuracy of risk assessment evaluations. *Law and Human Behavior, 34*, 275–281.
- Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment, 21*, 1–21.
- Hanson, R. K., & Thornton, D. (1999). *Static-99: Improving actuarial risk assessments for sex offenders*. User Report 99-02. Ottawa, Canada: Department of the Solicitor General.
- Hare, R. (1991). *The Hare Psychopathy Checklist–Revised*. Toronto, Canada: Multi-Health Systems.
- Hare, R. (2003). *The Hare Psychopathy Checklist–Revised* (2nd ed.). Toronto, Canada: Multi-Health Systems.
- Harris, G. T., & Rice, M. E. (2007). Characterizing the value of actuarial violence risk assessments. *Criminal Justice and Behavior, 34*, 1638–1658.
- Harris, G. T., Rice, M. E., & Quinsey, V. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior, 20*, 315–335.

- Hart, S. D. (1998). The role of psychopathy in assessing risk for violence: Conceptual and methodological issues. *Legal and Criminological Psychology, 3*, 121–137.
- Hart, S. D. (2001). Forensic issues. In J. W. Livesley (Ed.), *Handbook of personality disorders: Theory, research, and treatment* (pp. 555–569). New York, NY: Guilford Press.
- Hart, S. D., & Cooke, D. J. (2013). Another look at the (im-)precision of individual risk estimates made using actuarial risk assessment instruments. *Behavioral Sciences and the Law, 31*, 81–102.
- Hart, S. D., Douglas, K. S., & Webster, C. D. (2001). Risk management using the HCR-20: A general overview focusing on historical factors. In K. S. Douglas, C. D. Webster, S. D. Hart, D. Eaves, & J. R. P. Ogloff (Eds.), *HCR-20 violence risk management companion guide* (pp. 27–40). Burnaby, Canada: Mental Health, Law, & Policy Institute, Simon Fraser University.
- Hart, S. D., Kropp, P. R., Laws, D. R., Klaver, J., Logan, C., & Watt, K. A. (2003). The Risk for Sexual Violence Protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence. Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Hart, S. D., & Logan, C. (2011). Formulation of violence risk using evidence-based assessments: The structured professional judgment approach. In P. Sturmey & M. McMurrin (Eds.), *Forensic case formulation* (pp. 83–106). Chichester, England: Wiley-Blackwell.
- Hart, S. D., Michie, C., & Cooke, D. J. (2007). Precision of actuarial risk assessment instruments: Evaluating the “margins of error” of group *v.* individual predictions of violence. *British Journal of Psychiatry, 190*, s60–s65. doi:10.1192/bjp.190.5.s60
- Hart, S. D., Sturmey, P., Logan, C., & McMurrin, M. (2011). Forensic case formulation. *International Journal of Forensic Mental Health, 10*, 118–126. doi:10.1080/14999013.2011.577137
- Heilbrun, K. S. (1997). Prediction versus management models relevant to risk assessment: The importance of legal decision-making context. *Law and Human Behavior, 21*, 347–359.
- Heilbrun, K. S., Douglas, K. S., & Yasuhara, K. (2009). Violence risk assessment: Core controversies. In J. L. Skeem, K. S. Douglas, & S. O. Lilienfeld (Eds.), *Psychological science in the courtroom: Consensus and controversy* (pp. 333–357). New York, NY: Guilford Press.
- Hilton, N. Z., Harris, G. T., Rice, M. E., Lang, C., Cormier, C. A., & Lines, K. J. (2004). A brief actuarial assessment for the prediction of wife assault recidivism: The Ontario Domestic Assault Risk Assessment. *Psychological Assessment, 16*, 267–275.
- Hoge, R. D. & Andrews, D. A. (2002). *Youth Level of Service/Case Management Inventory: User's manual*. Toronto, Canada: Multi-Health Systems.
- Holt, R. R. (1970). Yet another look at clinical and statistical prediction: Or, is clinical psychology worthwhile? *American Psychologist, 25*, 337–349.
- In re* Commitment of R. S., 773 A. 2d 72 (N.J. Super. A.D., 2001).
- In re* Commitment of Simmons, 821 N.E.2d 1184 (2004).
- In re* Detention of Berry, 248 P.2d 592 (Wash., 2011).
- In re* Detention of Campbell, 986 P. 2d 771 (Wash., 1999).
- In re* Detention of Hargett, 786 N.E. 2d 557 (Ill. App. 3 Dist., 2003).
- In re* Detention of Holtz, 653 N.W. 2d 613 (Iowa, App., 2002).
- In re* Detention of Strauss, 20 P. 3d 1022 (Wash., App. Div. 1, 2001).
- In re* Detention of Thorell, 72 P. 3d 708 (Wash., 2003).
- In the Matter of David B., 97 N.Y. 2d 267 (2002).
- In the Matter of George L., 85 N.Y. 2d 295, 624 N.Y.S. 2d 99, 648 N.E. 2d 475 (1995).

- In the Matter of Registrant C. A., 679 A. 2d 1153 (1996).
- In the Matter of State v. Rosado, 889 N.Y.S.2d (2009).
- Janus, E. S., & Prentky, R. A. (2003). Forensic use of actuarial risk assessment with sex offenders: Accuracy, admissibility and accountability. *American Criminal Law Review*, 40, 1443–1499.
- Jones v. United States, 463 U.S. 354 (1983).
- Kansas v. Crane, 534 U.S. 407 (2002).
- Kansas v. Hendricks, 117 S. Ct. 2072 (1997).
- Kluckhohn, C., & Murray, H. A. (1953). Personality formulation: The determinants. In C. Kluckhohn, H. Murray, & D. Schneider (Eds.), *Personality in nature, society and culture* (pp. 53–67). New York, NY: Knopf.
- Kozol, H. L., Boucher, R. J., & Garofalo, R. F. (1972). The diagnosis and treatment of dangerousness. *Crime and Delinquency*, 19, 371–392.
- Kozol, H. L., Boucher, R. J., & Garofalo, R. F. (1973). Dangerousness: A reply to Monahan. *Crime and Delinquency*, 19, 554–555.
- Kropp, P. R., & Hart, S. D. (2000). The Spousal Assault Risk Assessment (SARA) Guide: Reliability and validity in adult male offenders. *Law and Human Behavior*, 24, 101–118.
- Kropp, P. R., Hart, S. D., & Lyon, D. (2008). Risk assessment of public figure stalkers. In J. R. Meloy, J. Hoffmann, & L. Sheridan (Eds.), *Stalking, threatening and attacking public figures* (pp. 343–361). Oxford, England: Oxford University Press.
- Kropp, P. R., Hart, S. D., Webster, C. D., & Eaves, D. (1995). *Manual for the Spousal Assault Risk Assessment Guide* (2nd ed.). Vancouver, Canada: British Columbia Institute on Family Violence.
- Kropp, P. R., Hart, S. D., Webster, C. D., & Eaves, D. (1999). *Spousal Assault Risk Assessment Guide user's manual*. Toronto, Canada: Multi-Health Systems/British Columbia Institute on Family Violence.
- Kumho Tire Co. v. Carmichael, 119 S. Ct. 1167 (1999).
- Kutzer, D., & Lion, J. R. (1984). The violent patient: Assessment and intervention. In S. Saunders, A. M. Anderson, C. A. Hart, & G. M. Rubenstein (Eds.), *Violent individuals and families: A handbook for practitioners* (pp. 69–86). Springfield, IL: Charles C Thomas.
- Laws, D. R., & O'Donohue, W. T. (2008). *Sexual deviance: Theory, assessment, and treatment*. New York, NY: Guilford Press.
- Lee v. State, 854 So. 2d 709 (Fla. App. 2 Dist., 2003).
- Leon v. McDonald, 2011 U.S. Dist. LEXIS 132211 (Cent. Dist. CA, 2011).
- Lessard v. Schmidt (1972) 349 F. Supp. 1978, vacated on other grounds, 414 U.S. 473 (1974).
- Lidz, C. W., Mulvey, E. P., & Gardner, W. (1993). The accuracy of predictions of violence to others. *Journal of the American Medical Association*, 269, 1007–1011.
- Lion, J. R. (1987). Clinical assessment of violent patients. In L. Roth (Ed.), *Clinical treatment of the violent person* (pp. 1–19). New York, NY: Guilford Press.
- Litwack, T. R. (1993). On the ethics of dangerousness assessments. *Law and Human Behavior*, 17, 479–482.
- Litwack, T. R. (1996). "Dangerous" patients: A survey of one forensic facility and review of the issue. *Aggression and Violent Behavior*, 1, 97–122.
- Litwack, T. R. (2001). Actuarial versus clinical assessments of dangerousness. *Psychology, Public Policy and Law*, 7, 409–443.

- Litwack, T. R. (2002). Some questions for the field of violence risk assessment and forensic mental health: Or, "back to basics" revisited. *International Journal of Forensic Mental Health, 1*, 171–178.
- Litwack, T. R., Gerber, G. L., & Fenster, C. A. (1980). The proper role of psychology in child custody disputes. *Journal of Family Law, 18*, 269–300.
- Litwack, T. R., Kirschner, S. M., & Wack, R. C. (1993). The assessment of dangerousness and predictions of violence: Recent research and future prospects. *Psychiatric Quarterly, 64*, 245–273.
- Litwack, T. R., & Schlesinger, L. B. (1987). Assessing and predicting violence: Research, law, and applications. In I. B. Weiner & A. K. Hess (Eds.), *Handbook of forensic psychology* (pp. 205–207). New York, NY: Wiley.
- Litwack, T. R., & Schlesinger, L. B. (1999). Dangerousness risk assessments: Research, legal, and clinical considerations. In A. K. Hess & I. B. Weiner (Eds.), *Handbook of forensic psychology* (2nd ed., pp. 171–217). New York, NY: Wiley.
- Litwack, T. R., Zapf, P. A., Groscup, J. L., & Hart, S. D. (2006). Violence risk assessment: Research, legal, and clinical considerations. In I. B. Weiner, & A. K. Hess (Eds.), *The handbook of forensic psychology* (3rd ed., pp. 487–533). Hoboken, NJ: Wiley.
- Lodewijks, H. P. B., de Ruiter, C., & Doreleijers, Th. A. H. (2008). Gender differences in violent outcome and risk assessment in adolescent offenders after residential treatment. *International Journal of Forensic Mental Health, 7*, 133–146.
- Lodewijks, H. P. B., Doreleijers, Th. A. H., & de Ruiter, C. (2008). SAVRY risk assessment in relation to sentencing and subsequent recidivism in a Dutch sample of violent juvenile offenders. *Criminal Justice and Behavior, 35*, 696–709.
- Logerquist v. McVey, 1 P.3d 113 (Ariz., 2000).
- McNiel, D. E., & Binder, R. L. (1995). Correlates of accuracy in the assessment of psychiatric inpatients' risk of violence. *American Journal of Psychiatry, 152*, 901–906.
- McNiel, D. E., Sandberg, D. A., & Binder, R. L. (1998). The relationship between confidence and accuracy in clinical assessment of psychiatric patients' potential for violence. *Law and Human Behavior, 22*, 655–669.
- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press.
- Mills, J. F., Jones, M. N., & Kroner, D. G. (2005). An examination of the generalizability of the LSI-R and VRAG probability bins. *Criminal Justice and Behavior, 32*, 565–585.
- Monahan, J. (1981). *Predicting violent behavior: An assessment of clinical techniques*. Beverly Hills, CA: Sage.
- Monahan, J. (1984). The prediction of violent behavior: Toward a second generation of theory and policy. *American Journal of Psychiatry, 141*, 10–15.
- Monahan, J. (1993). Limiting therapist exposure to Tarasoff liability. *American Psychologist, 48*, 242–250.
- Monahan, J. (1996). Violence prediction: The past twenty years and the next twenty years. *Criminal Justice and Behavior, 23*, 107–120.
- Monahan, J., & Silver, E. (2003). Judicial decision thresholds for violence risk management. *International Journal of Forensic Mental Health, 2*, 1–6.
- Monahan, J., Steadman, H., Appelbaum, P., Grisso, T., Mulvey, E., Roth, L., . . . Silver, E. (2005). *The classification of violence risk*. Lutz, FL: Psychological Assessment Resources.

- Mossman D., & Selke, T. M. (2007). Avoiding errors about "margins of error." *British Journal of Psychiatry*, *191*, 561.
- Mulvey, E. P., & Lidz, C. W. (1985). Back to basics: A critical analysis of dangerousness research in a new legal environment. *Law and Human Behavior*, *9*, 209–219.
- Mulvey, E. P., & Lidz, C. W. (1995). Conditional prediction: A model for research on dangerousness to others in a new era. *International Journal of Law and Psychiatry*, *18*, 129–143.
- Mulvey, E. P., Shaw, E., & Lidz, C. W. (1994). Editorial: Why use multiple sources in research on patient violence in the community? *Criminal Behaviour and Mental Health*, *4*, 253–258.
- N.J. Stat. §21:62A-16 (2001).
- Neves, A. C., Goncalves, R. A., & Palma-Oliveira, J. M. (2011). Assessing risk for violent and general recidivism: A study of the HCR-20 and the PCL-R with a non-clinical sample of Portuguese offenders. *International Journal of Forensic Mental Health*, *10*, 137–149.
- Newhill, C. E., Mulvey, E. P., & Lidz, C. W. (1995). Characteristics of violence in the community by female patients seen in a psychiatric emergency service. *Psychiatric Services*, *46*, 785–789.
- Nicholls, T. L., Viljoen, J. L., Cruise, K. R., Desmarais, S. L., & Webster, C. D. (2010). *Short-Term Assessment of Risk and Treatability: Adolescent Version (START:AV)* (abbreviated manual). Coquitlam, Canada: BC Mental Health and Addiction Services.
- Nuffield, J. (1982). *Parole decision-making in Canada*. Ottawa, Canada: Communication Division, Solicitor General of Canada.
- O'Connor v. Donaldson, 422 U.S. 563 (1975).
- Olver, M., Stockdale, K., & Wormith, J. (2009). Risk assessment with young offenders: A meta-analysis of three assessment measures. *Criminal Justice and Behavior*, *36*, 329–353.
- Peck v. the Counseling Service of Addison County. 499 A. 2d 422 (VT, 1985).
- Pedersen, L., Rasmussen, K., & Elsass, P. (2010). Risk assessment: The value of structured professional judgments. *International Journal of Forensic Mental Health*, *9*, 74–81.
- People v. Stevens (*In re Stevens*), 803 N.E. 2d 1036 (2004).
- People v. Taylor, 782 N.E. 2d 920 (Ill. App. 2 Dist., 2002).
- People v. Therrien, 113 Cal. App. 4th 609, 6 Cal. Rptr. 3d 415 (2003).
- People v. Ward, 71 Cal. App. 4th 368, 6 Cal. Rptr. 2d 828 (1999).
- Pfohl, S. J. (1978). *Predicting dangerousness*. Lexington, MA: Heath.
- Pinta, E. (2010). Tarasoff duties in prisons: Community standards with certain twists. *Psychiatric Quarterly*, *81*, 177–182. doi:10.1007/s11126-010-9127-1
- Pomerol, J.-C. (2001). Scenario development and practical decision making under uncertainty. *Decision Support Systems*, *31*, 197–204.
- Prins, H. (1988). Dangerous clients: Further observations on the limitation of mayhem. *British Journal of Social Work*, *18*, 593–609.
- Prins, H. (1996). Risk assessment and management in criminal justice and psychiatry. *Journal of Forensic Psychiatry*, *7*, 42–62.
- Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. (1999). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.
- Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. (2006). *Violent offenders: Appraising and managing risk* (2nd ed). Washington, DC: American Psychological Association.
- Rice, M. E., Harris, G. T., & Hilton, N. Z. (2010). The Violence Risk Appraisal Guide and Sex Offender Risk Appraisal Guide for violence risk assessment and the Ontario Domestic

- Assault Risk Assessment and Domestic Violence Risk Appraisal Guide for wife assault risk assessment. In R. Otto & K. Douglas (Eds.), *Handbook of violence risk assessment tools*. (pp. 99–120). Oxford, England: Routledge/Taylor & Francis.
- Ringland, G. (1998). *Scenario planning: Managing for the future*. Chichester, England: Wiley.
- Rogers v. Okin, 634 F. 2d 650 (1st Cir., 1980).
- Schaap, G, Lammers, S., & de Vogel, V. (2009). Risk assessment in female forensic psychiatric patients: A quasi-prospective study into the validity of the HCR-20 and PCL-R. *Journal of Forensic Psychiatry & Psychology*, 20, 354–365.
- Schall v. Martin, 467 U.S. 253 (1984).
- Schlesinger, L. B. (1996). The catathymic crisis (1912–present): A review and clinical study. *Aggression and Violent Behavior*, 1, 307–316.
- Schmidt, F., Campbell, M. A., & Houlding, C. (2011). Comparative analyses of the YLS/CMI, SAVRY, and PCL:YV in adolescent offenders: A 10-year follow-up into adulthood. *Youth Violence and Juvenile Justice*, 9, 23–42.
- Scott, P. D. (1977). Assessing dangerousness in criminals. *British Journal of Psychiatry*, 131, 127–142.
- Seltzer v. Hogue, 594 N.Y.S. 2d 781 (1993).
- Shah, S. A. (1978). Dangerousness: A paradigm for exploring some issues in law and psychology. *American Psychologist*, 33, 224–238.
- Simone, S., & Fulero, S. (2005). Tarasoff and the duty to protect. *Journal of Aggression, Maltreatment & Trauma*, 11, 145–168. doi:10.1300/J146v11n01_12
- Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and meta-regression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review*, 31, 499–513.
- Sjöstedt, G., & Långström, N. (2002). Assessment of risk for criminal recidivism among rapists: A comparison of four different measures. *Psychology, Crime and Law*, 8, 25–40.
- State ex rel. Romley v. Fields, 35 P. 3d 82 (Ariz. App. Div. 1, 2001).
- Steadman, H. J., & Cocozza, J. (1974). *Careers of the criminally insane*. Lexington, MA: Heath.
- Steadman, H. J., & Cocozza, J. (1980). The prediction of dangerousness—Baxstrom: A case study. In G. Cooke (Ed.), *The role of the forensic psychologist* (pp. 204–215). Springfield, IL: Charles C Thomas.
- Tarasoff v. Regents of the University of California, 17 Cal. 3d 425, 551 P. 2d 334 (1976).
- Tardiff, K. (1996). *Concise guide to assessment and management of violent patients* (2nd ed.). Washington, DC: American Psychiatric Press.
- Thornberry, T. P., & Jacoby, J. E. (1979). *The criminally insane: A community follow-up of mentally ill offenders*. Chicago, IL: University of Chicago Press.
- Tolman, A. O. (2001). Clinical training and the duty to protect. *Behavioral Sciences and the Law*, 19, 387–404.
- United States v. Barnette, 212 F. 3d 803 (2000).
- United States v. Evanoff, 10 F. 3d 559 (8th Cir., 1993).
- United States v. Fields, 483 F.3d 313 (5th Cir., 2007).
- United States v. Sahhar, 917 F. 2d 1197 (1990); cert. denied, 111 S. Ct. 1591 (1991).
- United States v. Salerno, 481 U.S. 739 (1987).
- United States v. Scheffer, 523 U.S. 303 (1998).
- United States v. Williams, 299 F.3d 673 (Ill., 2002).

- VandeCreek, L., & Knapp, S. (2001). *Tarasoff and beyond: Legal and clinical considerations in the treatment of life-endangering patients* (3rd ed.). Sarasota, FL: Professional Resource Press/Professional Resource Exchange.
- van den Brink, R. H. S., Hooijschuur, A., van Os, T. W. D. P., Savenije, W., & Wiersma, D. (2010). Routine violence risk assessment in community forensic mental healthcare. *Behavioral Sciences & the Law*, *28*, 396–410.
- van der Heijden, K. (1994). Probabilistic planning and scenario planning. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 549–572). Chichester, England: Wiley.
- van der Heijden, K. (1997). *Scenarios: The art of strategic conversation*. New York, NY: Wiley.
- Viljoen, J. L., Mordell, S., & Beneteau, J. L. (2012). Prediction of adolescent sexual reoffending: A meta-analysis of the J-SOAP-II, ERASOR, J-SORRAT-II, and Static-99. *Law and Human Behavior*, *36*, 423–438.
- Viljoen, J. L., Scalora, M., Cuadra, L., Bader, S., Chávez, V., Ullman, D., & Lawrence, L. (2008). Assessing risk for violence in adolescents who have sexually offended: A comparison of the J-SOAP-II, J-SORRAT-II, and SAVRY. *Criminal Justice and Behavior*, *35*, 5–23.
- Vincent, G. M., Chapman, J., & Cook, N. (2011). Predictive validity of the SAVRY, racial differences, and the contribution of needs factors. *Criminal Justice and Behavior*, *38*, 42–62.
- Walcott, D. M., Cerundolo, P., & Beck, J. C. (2001). Current analysis of the Tarasoff duty: An evolution towards the limitation of the duty to protect. *Behavioral Sciences and the Law*, *19*, 325–343.
- Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *The HCR-20 scheme: The assessment of dangerousness and risk (Version 2)*. Burnaby, British Columbia, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Webster, C. D., Harris, G. T., Rice, M. E., Cormier, C., & Quinsey, V. L. (1994). *The violence prediction scheme: Assessing dangerousness in high risk men*. Toronto, Canada: University of Toronto Press.
- Webster, C. D., Martin, M. L., Brink, J., Nicholls, T. L., & Desmarais, S. L. (2009). *Manual for the Short-Term Assessment of Risk and Treatability (START) (Version 1.1)*. Coquitlam, Canada: British Columbia Mental Health & Addiction Services.
- Webster, C. D., & Menzies, R. J. (1989). The clinical prediction of dangerousness. In D. N. Weisstaub (Ed.), *Law and mental health: International perspectives* (Vol. 3). Toronto, Canada: Pergamon Press.
- Webster, C. D., & Polvi, N. H. (1995). Challenging assessments of dangerousness and risk. In J. Ziskin (Ed.), *Coping with psychiatric and psychological testimony* (5th ed., pp. 22–38). Los Angeles, CA: Law and Psychology Press.
- Werth, J., Jr., Welfel, E., & Benjamin, G. (2009). *The duty to protect: Ethical, legal, and professional considerations for mental health professionals*. Washington, DC: American Psychological Association.
- Westerheide v. State, 767 So. 2d 637 (Fla. App. 5 Dist., 2000).
- Wettstein, R. M. (1984). The prediction of violent behavior and the duty to protect third parties. *Behavioral Sciences and the Law*, *2*, 291–317.
- White, S. G., & Meloy, J. R. (2010). *WAVR-21: A structured professional guide for the workplace assessment of violence risk* (2nd ed.). San Diego, CA: Specialized Training Services.
- Wilson, C., & Douglas, K. S. (2009). Assessment of dangerousness. In C. Edwards (Ed.), *Encyclopedia of forensic sciences* (pp. 667–676). Hoboken, NJ: Wiley.

- Weiner, I. B., & Hess, A. K. (Eds.). (2006). *The handbook of forensic psychology* (3rd ed.). Hoboken, NJ: Wiley.
- Wong, S., & Gordon, A. E. (1999–2003). *The Violence Risk Scale*. Saskatoon, Canada: University of Saskatchewan, Department of Psychology.
- Worling, J. R., & Curwen, T. (2001). *Estimate of Risk of Adolescent Sexual Offense Recidivism, Version 2.0*. Toronto, Canada: Ontario Ministry of Community and Social Services.
- Yang, M., Wong, S. C. P., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, 136, 740–767.

Psychology and Law Enforcement

ELLEN M. SCRIVNER, DAVID M. COREY, AND LORRAINE W. GREENE

THE application of psychology to law enforcement has created opportunities for psychologists to contribute to shaping the public safety in the nation's communities. In some respects, the development of the psychology–law enforcement relationship parallels the forensic collaboration between the larger law and psychology enterprise, as reflected in similar developmental trends and the growing numbers of psychologists involved in these activities. At one time, the disciplines of psychology and law enforcement seemed mutually exclusive, yet over the past few decades, fruitful collaboration has occurred and the outcomes reflect an expanding body of knowledge that demonstrates how psychology adds value to the operation of law enforcement systems. These outcomes culminated in a significant achievement in 2011 when the American Board of Professional Psychology (ABPP) established the American Board of Police & Public Safety Psychology, an event heralded as the single most important development in police psychology since its inception (Trompetter, 2011).

Relatively unheard of until the 1960s (Reese, 1987; Reiser, 1972), the practice of applying psychology in law enforcement evolved as forward-thinking public safety executives came to recognize that repeated exposure to a difficult environment can take a toll on human beings. Acknowledging the unique culture of law enforcement, many sought proactive approaches to optimize the psychological functioning and personal adjustment of officers and reduce occupational stress. Others viewed the services of psychologists as a way to reduce risk and protect departments from liability claims; they viewed psychological services as a type of cover. Still, their response could be considered tempered when compared to those who resisted these efforts and believed that psychological services were of little value. Consequently, psychologists who were trying to practice in this unique environment had to carefully negotiate significant hurdles.

Notwithstanding the varied reasons responsible for bringing psychology into law enforcement, over time psychologists have established a presence in law

enforcement agencies and now deliver a range of direct psychological services to individual officers, their families, and their respective agencies. There is little question today that psychologists have made a difference and have had an impact on the delivery of law enforcement services across the country.

This chapter addresses those differences and how they translate into professional activity. We summarize the evolution of law enforcement psychology and the core technologies that are generally accepted as the basic framework of law enforcement psychological services. Also, we address issues that psychologists encounter when they attempt to practice in this unique setting as well as new trends that are evolving as this partnership continues to grow. Some of these trends are responses to the range of public safety issues that emerged following September 11, 2001, but others signify practice changes that developed as the field expanded.

EVOLUTION AND GAINING ACCEPTANCE

Law enforcement is a high-risk but also highly structured occupation that has been characterized as tight-knit, paramilitary, and rigid and as a hidebound bureaucracy not given to innovation. Although organizational changes have occurred in relation to the increasing presence of more highly educated law enforcement executives and through reforms such as community policing, this clearly was not the environment psychologists encountered when first starting to work with law enforcement.

Initially, the tradition-clad agencies were uncertain about the need for psychological services, and psychologists had an uphill battle to gain credibility and develop an understanding of the law enforcement culture. In addition, there was an immediate need to respond to professional practice issues that emerged when trying to establish services in that environment. Questions surfaced about maintaining overall integrity of services, and certain activities raised distinct issues specific to the nature of the practice. An overarching question was: Who was the client—the applicants psychologists screened for police department jobs or the police organizations that wanted to hire them? In the context of providing clinical interventions, was the client the officer seeking treatment or the organization that referred the officer because of concern about his or her capacity to enforce the law in a fair and impartial manner?

The client question had clear implications with respect to the confidentiality of communications, the cornerstone of psychological services and a fundamental concern of law enforcement officers. However, confidentiality was not fully understood in non-health-care organizations, and particularly so in those that operated as closed systems. Therefore, limiting disclosure to the minimum information needed to fulfill the purpose of evaluations, safeguarding privileged information, and maintaining privacy of records in a system that expects definitive answers and is known for reacting quickly to solve problems presented challenges. All too often, psychologists were confronted with statements such as “Tell us who has a problem and we’ll fire them.”

Fortunately, many of these issues were resolved by state laws and the professional standards that govern the practice and licensure of psychologists. However, for psychologists providing services in a law enforcement setting, practice issues continue to require ongoing and careful monitoring. One misstep that raises questions about professional practices could undo a well-functioning program, and any breach of confidentiality could undermine the years of work it took to build credible services. The earlier and simplistic question, Who is the client?, has since given way to the recognition that police psychologists owe duties to multiple parties in complex organizational systems (Fisher, 2009). Ongoing training on psychological professional standards is recommended and conducted within some agencies as an integral part of the training provided to new police personnel promoted into leadership positions.

Overcoming the initial resistance and resolving professional issues, psychologists went on to make significant inroads into improving psychological functioning in the tradition-clad occupations that are responsible for community public safety. Today, virtually every police agency in America engages psychological services for purposes of preemployment psychological screening, fitness-for-duty evaluations, critical incident stress management or psychological first aid, employee assistance counseling, operational support, or organizational consulting, and often as a result of statutes or regulations that require the service.

Overall, it appears that the concept of psychological services has been institutionalized in law enforcement agencies and that psychology has played a major role in breaking down barriers, expanding the use of mental health services in this specialized environment, and creating opportunities to facilitate organizational learning. In essence, psychology has been instrumental in shaping a major shift in the culture of law enforcement.

KEY EVENTS SIGNIFYING A CULTURE SHIFT

Broad-based use and acceptance of psychologists and other mental health service providers in law enforcement have been supported by a series of key events that represent the building blocks of professional development and signify a major shift in the culture. These events are described briefly next.

- Five police psychology conferences were hosted by the Federal Bureau of Investigation (FBI) Training Academy (1984 to 2001). These meetings brought psychologists together to discuss issues relative to providing services in law enforcement agencies. In addition to general practice issues, topics included critical incident stress, organizational issues, the impact on families, and suicide in law enforcement.
- Professional organizations acknowledged the work of psychologists in law enforcement and developed sections in their organizations to shape policy and guide the development of professional procedures. Division 18 (Psychologists

in Public Service) of the American Psychological Association (APA) developed a section on Police and Public Safety Psychology, and the International Association of Chiefs of Police developed the Police Psychological Services Section (IACP-PPSS). These sections provide venues for peer exchange and ongoing training to address unique professional issues that psychologists encounter when working with law enforcement.

- Police psychologists, supported by APA, provided congressional testimony on police stress and family well-being (*On the Front Lines: Police Stress and Family Well-Being*, 1991). This was the impetus for an amendment to the 1994 Omnibus Crime Act that provided for the development of the federally funded Corrections and Law Enforcement Family Support Program managed by the National Institute of Justice (NIJ). From 1996 to 2003, NIJ funded approximately 30 innovative programs to treat stress, deliver training, and conduct survey research to help understand the needs of law enforcement and corrections officers. These programs supported partnerships with law enforcement agencies, labor unions, and professional organizations. In an APA-sponsored Police Chiefs Roundtable Series, 15 years after establishing a public safety presence in Division 18, police psychologists brought police chiefs together, and they sought input from the APA Committee on Urban Initiatives on strategies to manage a range of problems that affect the quality of American policing. From 1998 to 2000, the *APA Monitor on Psychology* published articles about the breadth of the activities of psychologists working with law enforcement and showcased their professional activities (Rabasca, 2000). For over two decades, the Police and Public Safety Section of APA Division 18 has offered a miniconference, one-day training focused on the needs of law enforcement and public safety personnel and agencies. A growing literature on police psychology (Blau, 1994; Kitaeff, 2011; Kurke & Scrivner, 1995; Toch, 2002, 2012) also includes the compilation of papers from the FBI Conference series and articles in peer-reviewed journals. Following September 11, 2001, the FBI Academy and APA convened a meeting on how psychology could assist in counterterrorism initiatives. This conference represented a significant benchmark for police psychology in that it combined operational and theoretical viewpoints and also included psychologists representing the APA Science Directorate, federal agents, and police officials (Smith, 2002).
- The IACP-PPSS has been instrumental in providing guidelines for professional practice that address a range of issues but are particularly salient for preemployment psychological evaluations as well as psychological fitness-for-duty evaluations. Updated every 5 years, the current standards were published in 2009. These guidelines were featured in the August 2011 edition of *The Police Chief*, the official publication of the IACP, in a special edition that featured psychological issues in policing.

- In 2008, APA recognized police psychology as a professional proficiency, but clearly the most significant event in the history of this field was the 2011 affiliation of Police and Public Safety Psychology as a specialty board of the ABPP (Corey, Cuttler, Cox, & Brower, 2011).

These developments have strengthened the professional dimensions of the field. They have facilitated a growing law enforcement–psychology literature, encouraged presentations at professional conferences, and fostered an interest in scholarly research, including a body of work devoted to police stress. More recently, Corey and Stewart (in press) broadened the research portfolio through their examination of research on officers involved in violent acts and proposed a model to classify offenders that will facilitate managing risk in an occupation where the legitimate use of force is intrinsic to the role of police. Their proposal is based on the Workplace Violence Model developed by the Occupational Safety and Health Administration (OSHA) and the National Institute of Occupational Safety and Health (NIOSH). This type of work is critically important both to law enforcement and to the field of psychology in that it demonstrates how the evidence-based knowledge of psychology can contribute to helping law enforcement nationwide to manage police forces, discharge their authority under the rule of law, and provide police services that can withstand community scrutiny and engender community trust.

As a result of these achievements, in total, a more comprehensive and multifaceted role has been established for an active presence of law enforcement psychology in the nation’s law enforcement agencies. Further, most of the major law enforcement agencies now provide some level of mental health services for employees and include psychologists as consultants in police operations and in developing training exercises.

WHAT DO LAW ENFORCEMENT PSYCHOLOGISTS DO?

The introduction of psychologists into law enforcement brought new sets of skills to these agencies, and these skills defined the core technologies used by law enforcement psychology. Subsequently, four primary domains have been identified that, with some exceptions, are similar to and consistent with the competencies that psychologists provide in other settings. They include evaluation/assessment activities, clinical intervention services, organizational consulting and training, and operational and investigation support. The latter brings psychologists into closer contact with law enforcement and also provides an opportunity for psychology to contribute to the broader world of criminal justice.

ASSESSMENT DOMAIN

In the early 1900s, Terman and then Thurstone used psychological tests to try to identify successful candidates for law enforcement positions (Super, 1999).

However, it was federal funding from the Law Enforcement Assistance Administration (LEAA; President's Commission on Law Enforcement and Administration of Justice, 1967), established by the Omnibus Crime Control & Safe Street Act of 1968, that encouraged law enforcement agencies to seek the expertise of psychologists to help them select emotionally stable candidates with personal characteristics suitable for law enforcement work. The awarding of these funds also supported recommendations from the 1968 National Advisory Commission on Civil Disorder and subsequently created a law enforcement psychological preemployment screening specialization (Kurke & Scrivner, 1995). The screening emphasis was a direct challenge to police departments nationwide to ensure that they were fulfilling their duty to their communities by selecting police officers in their agencies who were emotionally, behaviorally, and cognitively fit to perform the duties and functions of a law enforcement officer.

Although psychological screening of job candidates was a fairly traditional responsibility for psychologists, it was quite new to the police personnel function. It has now grown to the point that findings from a 1994 study showed that almost all of the 50 largest cities participating in the Major City Chiefs Association used psychological tests as part of their preemployment applicant screening process (Scrivner, 1994). Further, 38 states now mandate the use of psychological evaluations in the preemployment screening of candidates for police positions (Corey & Borum, 2013). This area of specialization has positioned psychologists to have a major impact on the quality of law enforcement services.

Critical issues affecting preemployment psychological suitability assessments are well known and involve using up-to-date tests that can be justified in personnel decision making, having substantial statistical evidence to justify use of cut-off scores, using appropriate interview techniques, communicating test results appropriately, engaging in ongoing validation of the evaluation process, and conforming to civil rights legislation and Equal Employment Opportunity Commission requirements. In addition, preemployment evaluation procedures are influenced by the Americans with Disabilities Act (ADA, 1990) and now the Genetic Information Nondiscrimination Act (GINA, 2008). In fact, the ADA has had a major impact on the sequencing of psychological evaluations in preemployment assessment and affects decisions on when to use some psychological tests, particularly those that screen for mental impairment. The continual updating of the IACP-PPSS Guidelines has been very helpful in framing the significant issues that the psychologist must be aware of when engaged in all components of these assessments, including the face-to-face assessment interview. Ben-Porath et al. (2011) define the assessment of psychological suitability for the law enforcement position as one of the essential functions of police psychology, and they cite the guidelines as the set of principles and recommended procedures for conducting these assessments.

Preemployment assessments raised other questions about *screening in* viable candidates versus *screening out* those with problems that reflect some level of mental impairment and about the use of clinical judgment paradigms to make job-related

decisions versus predictive statistical models. It might seem obvious that one would be concerned with both screening in and screening out. However, much of the early validation work used test instruments that were developed and normed with clinical populations. As such, they were more appropriate to screening-out decisions. Gradually, psychological tests were developed specifically for preemployment screening of police candidates, including the Hilson Personnel Profile/Success Quotient (Inwald & Brobst, 1988), which measures dimensions such as work ethic and social skills, the CPI Form 434 (Roberts, 1995), and the Matrix-Predictive Uniform Law Enforcement Selection Evaluation (M-PULSE), a self-report measure designed to measure attitudes and behaviors relevant to law enforcement (Davis & Rostow, 2008). These instruments examine candidates for suitability to perform law enforcement functions and compare candidate scale scores to norms developed from large samples of police applicants as well as a sample of incumbents (i.e., applicants who were subsequently hired and successfully completed probation). Dantzker (2011) identified the most commonly cited assessment instruments used in screening police applicants as the Minnesota Multiphasic Personality Inventory-2 (MMPI-2); the Inwald Personality Inventory (IPI); the California Psychological Inventory (CPI); the Personality Assessment Inventory (PAI); the NEO Personality Revised (NEO PI-R); and the Sixteen Personality Factor, Fifth Edition (16PF). Some of these assessment instruments may be more appropriate for pre-offer use since they assess normal traits and behaviors, whereas others are preferred for assessing psychopathology at the postoffer stage. Further, some are better than others for predicting on-the-job performance (Ben-Porath, Corey, & Stewart, 2011; Corey & Borum, 2013). Irrespective of the tests selected for a suitability battery, it is very clear that the research component of this domain has expanded considerably, and today there is a rich literature available on assessing police candidates for psychological suitability.

Although validation data initially were somewhat limited, that picture continues to change. Many psychologists providing suitability assessments for departments collect validation data, and there is a greater emphasis on publishing results of validation studies. Others have taken a somewhat different approach. Hough & Ones (2002) assessed characteristics needed for effective community policing and provided a comprehensive review of the literature that incorporates research done on other jobs in which similar performance dimensions are important for successful performance, and documented assessment scales that are not directed exclusively at eliminating mental impairment. In an earlier study, Hough (2000) contended that validity of some predictors could be expected to generalize and would be useful for predicting law enforcement performance, particularly community policing. She provided a thorough assessment of a wide range of promising selection measures as well as recommendations for criterion validation studies.

The research of Hough and Ones also is relevant to the screening-in issue and is consistent with the work of the California Peace Officer Standards and Training Commission (POST). The POST approach to preemployment law enforcement

assessment has been broadened to include a range of suitability criteria related to effective law enforcement performance. These developments make clear distinctions between mental impairment and the traits or characteristics that are necessary to perform the essential functions of the law enforcement job, such as ability to communicate with diverse groups of people, manage conflict, and tolerate stress. Consequently, these selection criteria are not driven by models based solely on psychopathology. This changing focus on preemployment psychological evaluations is expanding the utility of screening and adds value to current preemployment evaluation processes.

FITNESS-FOR-DUTY EVALUATIONS

Although preemployment screening has dominated the assessment conversation, post-hire concerns about an incumbent police officer's emotional and behavioral functioning often result in another kind of mandatory evaluation: namely, the psychological fitness-for-duty evaluation (FFDE). These evaluations are highly sensitive and present a different set of issues from those raised by preemployment screening. In this context, there is a critical need to differentiate the fitness evaluation from other forms of psychological services and to ensure that the officers undergoing examination understand these differences. These evaluations are the result of mandatory referrals, and a verbal or written report is provided to the client agency summarizing the results of the evaluation. Under these conditions, confidentiality is limited. Administrative personnel and supervisors also need to understand these differences and frequently require training in this area so that they are able to make an appropriate referral for an FFDE. In this regard, Corey's (2011) detailed discussion of the ethical, legal, and practice issues pertaining to FFDEs of police officers provides information useful to police psychologists, police administrators, and human resource personnel.

One useful supplement to FFDE referrals could be assisting with the development and implementation of an early intervention system (EIS) wherein supervisors learn to recognize certain types of behavior and help the employee get assistance before a problem develops to a level that a mandatory evaluation is required. This proactive form of risk management is being used in many departments, including those that have consent decrees or memoranda of agreement with the United States Department of Justice due to a pattern or practice of inappropriate police behavior (Batts, Smoot, & Scrivner, 2012).

The distinction crafted by Fischler et al., (2011) is important in differentiating the preemployment psychological suitability assessment from the FDE. They define the methodology of the psychological screening assessment as providing a standardized protocol for all applicants to determine suitability for employment. In contrast, an FFDE involves a more individualized approach based on the need to understand the underlying concerns that were responsible for the evaluation referral. Given the mandatory and nonconfidential nature of the process, law pertaining

to this type of evaluation has been established, and it confirms a police chief's right—even obligation—to order an evaluation, take reasonable steps to ensure the psychological suitability of employees, and to ensure that communications regarding fitness evaluations are properly protected (see Corey & Borum, 2013; Flanagan, 1995; Ostrov, 1995). A related issue involves the potential for conflict of interest. Psychologists who provide therapeutic services to officers should avoid fitness evaluations because of the inherent conflict in roles. Similar issues pertain when evaluating officers for work-related disabilities. Federal legislation needs to be considered in disability evaluations and also in other core activities, specifically the Civil Rights Act of 1964, the Rehabilitation Act of 1973, the ADA of 1990, the ADA Amendments Act of 2008, and GINA of 2008.

CLINICAL INTERVENTION DOMAIN

As preemployment screening gained some measure of success, law enforcement agencies began to request clinical services for police personnel experiencing the effects of cumulative or posttraumatic stress, substance use disorders, marital and relationship conflicts, and other problems. By 1980, both applicant screening and clinical interventions designed to help officers cope with the stressful nature of policing had been identified as primary activities of psychologists working with law enforcement (Stratton, 1980). Responding to an array of personal problems believed to be intensified by this line of work, psychologists developed services to assist officers in dealing with a range of problems, including marital conflict, family problems, substance abuse, depression, anxiety, anger, and aggressive behavior directed to themselves or others (Reese & Scrivner, 1994).

In 2001, there was a concentrated focus on preventing police suicide (Sheehan & Warren, 2001), perhaps in part due to publicity surrounding 26 police officer suicides at the New York City Police Department between 1993 and 1995 (Genet & Dowling, 2011). Finn, Talucci, and Wood (2000) reported that negative press coverage and allegations of corruption may have been responsible for these high numbers. Volanti (1996), however, advanced a theoretical framework that included “role constriction theory” based on the premise that, as the police role begins to dominate work life, officers' cognitive coping style becomes constricted and affects how they cope with psychological issues in their personal lives. He contended that with time on the job, law enforcement officers risk becoming overly constricted, and their options to change or get help for a problem become limited. The interaction between the nature of the job and the socialization into the police culture generally creates the constricted environment, Volanti argued, but specific factors drive the constriction, such as law enforcement selection procedures, training experiences, job stress, and department and public expectations.

Reports on police suicide for the most part are provided by statistics from large metropolitan agencies; only limited information is available regarding the

prevalence among small or midsize police departments, which constitute the majority of American law enforcement agencies. The IACP Police Psychology Section, the Bureau of Justice Assistance, and EEI Communications partnered to collect prevalence data, develop resources to prevent suicide from occurring, and provide effective department interventions. *Preventing a Law Enforcement Officer Suicide: A Compilation of Resources and Best Practices* (IACP, 2009) is a CD-ROM that contains a collection of materials to implement a suicide prevention and intervention program. Findings from the Montreal Police Service (Mishara & Martin, 2012) affirm the positive impact of comprehensive intervention programs in reducing the incidence of suicide among police officers. Page and Jacobs (2011) have contributed an important and valuable focus on the nature of clinical support required in rural agencies.

The focus on suicide remains today and is accentuated by the military's experience with unprecedented numbers of suicides among returning veterans. However, law enforcement also confronts a phenomenon highly unique to the law enforcement role: namely, "suicide by cop," which is a form of victim-precipitated homicide (Kennedy, Homant, & Hupp, 1998; Mohandie, Meloy, & Collins, 2009). "Suicide by cop" refers to a situation in which a person with suicidal intent uses or feigns the use of deadly force to threaten a law enforcement officer, with the intention that the officer will use fatal force against him or her. This is a difficult phenomenon for the general public to understand, much less accept, and it usually ends up in headlines about "trigger-happy cops" and costly lawsuits. This phenomenon occurs in perhaps as many as one-third of all officer-involved shootings (Mohandie et al., 2009) and plays a significant role in police shootings (Kennedy et al., 1998). Moreover, it has a strong impact on an officer's emotional and behavioral functioning.

LAW ENFORCEMENT TRAUMA

A significant component of clinical intervention services includes providing a crisis response to help officers adjust to on-duty traumatic incidents, a situation that occurs with some frequency in public safety occupations and at rates that exceed those in the general population. Consequently, psychologists developed focused interventions designed for officers involved in traumatic incidents. Initially developed as a response to potential trauma following an on-duty shooting incident and labeled "post-shooting trauma" (Reese, Horn, & Dunning, 1991), this reaction soon became known as critical incident stress in order to incorporate other traumatic incidents unrelated to police shootings (Bohl, 1995). Unfortunately, a growing number of incidents have required intervention. Psychologists were active at the Oklahoma City bombing, the TWA airliner crash on Long Island, the 1993 World Trade Center bombing, the 2001 World Trade Center Twin Towers collapse, and the related airliner crashes in Pennsylvania and Virginia.

Traumatic incident interventions are initiated to reduce immediate stress and prevent posttraumatic stress disorder (PTSD). From the prevention perspective,

they also seek to identify officers who may require further treatment. The process that is used is based on a short-term, crisis intervention response that involves critical incident stress management. Frequently it is delivered in an individual or group defusing/debriefing format, and it is considered to be a type of psychological first aid, in contrast to psychotherapy.

The effectiveness of posttraumatic incident debriefings for victims of trauma, not just law enforcement officers, has come under scrutiny, and the questioning has intensified since the horrific events of September 11, 2001. Investigators researching the issue have questioned the effectiveness of debriefings in preventing PTSD (McNally, Bryant, & Ehlers, 2003). McNally et al. concluded that, although most of the research is faulty, the preponderance of evidence suggests that debriefings do not prevent PTSD and in some instances can be harmful. This debate is ongoing, and research continues to emerge that questions the timing (Carlier, Voerman, & Gersons, 2000), effectiveness, and need for mandatory debriefing as well as the long-term effects of posttraumatic stress (Choe, 2005). Given the sensitivity of responding appropriately to those exposed to trauma, it is clear that further research is needed to resolve some of these issues.

Others promote the use of peer support teams to provide emotional support to fellow police personnel. Kamena, Gentz, Hays, Bohl-Penrod, and Greene (2011) maintained that the mission is to provide emotional and social support during professional crises and that police psychologists play an integral role in the development, selection, and training of peer support teams. IACP provides guidelines that include policies and procedures to guide departments in instituting peer support programs (IACP, 2006).

FAMILY SERVICES

As early as 1977, Niederhoffer and Niederhoffer discussed the potential negative impact that law enforcement work has on marriages and families. The development of the federally funded Corrections and Law Enforcement Family Support (CLEF) program subsequently intensified the focus on family issues and funded the development of a number of family programs. To provide guidance to CLEF, Finn and Tomz (1997) interviewed mental health practitioners, law enforcement and corrections officers, and family members in selected sites. Their results suggested that job-related stress affects officers and their family members, who experience:

- Increased cynicism and suspiciousness
- Increased emotional detachment from various aspects of daily life
- Reduced work efficiency, with absenteeism and early retirement
- Excessive aggressiveness, resulting in an increase in citizen complaints
- Substance abuse
- Marital and family problems, compounded by extramarital affairs or domestic violence

- PTSD
- Health problems such as ulcers, weight gain, and cardiac problems
- Suicide

However, in a lessons-learned study of the CLEF program, Delprino (2001) found that, although sources of stress for officers and their family members had been identified, typically the actual use of many of these programs was somewhat limited. Programs that are integrated into the police agency training and behavioral health divisions, such as the Metropolitan Nashville Police daylong training for police trainee spouses and other family members, may have longevity. Initially funded through CLEF, in the late 1990s, it continues to have participants from police trainee family members.

STRESS IN LAW ENFORCEMENT

Law enforcement officers are daily witnesses to man's inhumanity to man and sustain continued exposure to the dark side of life, such as murder, rape, hostage taking, and other violent acts. Yet a fairly common research finding is that the major negative stressors experienced by law enforcement are related more to administrative or routine work factors. In fact, Collins and Gibbs (2003) reported that occupational stressors ranking most highly among police officers "were not specific to policing, but to organizational issues such as the demands of work impinging upon home life, lack of consultation and communication, lack of control over workload, inadequate support and excess workload in general" (p. 256). Similarly, Liberman et al. (2002) found that routine work stressors were more stressful to law enforcement officers than exposure to danger and critical incidents. Routine stressors can range from work schedules and lack of advancement, to media exposure of negative police events and concerns about liability, to a hardening of the emotions that creates communication and attitudinal problems. Several researchers showed how routine stress starts early in a career and subsequently fuels marital conflict (Delprino, 2001; Delprino, O'Quin, & Kennedy, 1997; Eisenberg, 1975; Finn & Tomz, 1997; Toch, 2002). Other consistent findings confirm that seeing a partner or fellow officer killed and responding to abused children and to victims of serious accidents also are major stressors.

Even the summary of the NIJ/CLEF research portfolio that documented the causes and effects of job-related stress, as applied to officers and their families, indicated that, although exposure to violence, suffering, and death is inherent to the profession, other sources of stress have greater impact on officers. These stressors include light sentences for offenders, unfavorable public opinion of police performance, irregular work hours and shift work, dealing with abused children and child homicides, ministering to survivors of vehicle crashes, and organizational stressors, including limited advancement opportunities and excessive paperwork. In contrast to persons in many other occupations, law enforcement personnel view

stress as a normal part of their job but also see themselves as being under more pressure than officers were 10 to 20 years ago (Finn et al., 2000).

Gershon (1999) discussed police stress and its implications for public health. In a video developed by the NIJ, Gershon presented findings from "Project Shield," which examined incidence and prevalence of police occupational stress and the related psychological and physical health outcomes in a specific police department. High scores on a stress inventory were related to self-reports of poor health and spouse abuse. More specifically, Gershon's results documented the most stressful events as identified by police officers: (a) attending a police funeral (the most stressful event), (b) being a target of an internal investigation, (c) sustaining a needle stick injury, (d) making a violent arrest, and (e) having personal knowledge of a victim. These stressors contributed to low energy, headaches, family abuse of both spouse and children, depression, and anxiety. Further, Gershon reported that 1% of the participants admitted to frequent thoughts about suicide.

These findings confirm that, although the job is inherently difficult, law enforcement stress comes from both internal and external sources. However, it is the impact of these issues that typically send law enforcement personnel, and frequently their loved ones, to the psychologist. Moreover, the findings drive the need for a training agenda that calls for the involvement of psychologists who understand law enforcement issues.

LAW ENFORCEMENT TRAINING

Psychologists have used their knowledge base to develop a variety of training programs but have placed strong emphasis on training officers to use stress management tools. The goals have been to make law enforcement officers more resilient and better able to manage the stress in their jobs and to acquaint them with a variety of stress prevention and reduction strategies. In addition to a focus on coping techniques, other training programs have been developed to help supervisors identify signs of stress in their employees and make appropriate and timely referrals for services.

Although training in stress management is critical, White and Honig (1995) discussed how psychologists also apply their expertise to develop training on other subject matters within the domain of psychology. Some examples include enhancing communication skills, responding to persons with mental illness, cross-cultural awareness, hostage negotiations, use of force, domestic violence, dynamics of sexual assault, responding to hate crimes, and practical issues relating to child-rearing practices and preparing for retirement. In-house psychologists have implemented programs to improve field training experiences by providing information on learning preferences and communication styles for both field training officers and trainees.

Over time, psychologists also brought adult learning models into law enforcement training and made greater use of role play and simulation techniques, in contrast

to the talking-head lecture approach of prior years. In this context, they created an approach to training that is less academic and instructor centered in favor of one that focuses on developing competencies in the adult learner. These methods have been used to develop health-related competencies that use disease prevention and wellness methods as appropriate models for responding to police stress (Harpold & Feemster, 2002). Further, there is now greater emphasis on developing resiliency in officers as a means of helping them withstand the stress of their job.

ORGANIZATIONAL CONSULTING

Organizational consulting activities go beyond the delivery of traditional mental health services and focus instead on strengthening the organization. This can be done by building resiliency from within or by acting as a change agent with the goal of improving agency performance. These are newer roles for psychologists in law enforcement, but they can have a substantial impact on the operation of an organization as well as organizational health. In 2002, Sewell contended that law enforcement agencies were in an era of change because of technological advances, environmental and economic factors, and new political influences expressed through grassroots, community-based criminal justice (policing, prosecutors, courts, corrections, and victim services). In his view, such changes affect the structure and policies of organizations and create considerable stress, and he suggested that the organizations need as much help as the individual officers. We believe that similar changes are with us today as law enforcement agencies struggle to provide services at a time of diminished budgets and new challenges. Not the least of these challenges are generational changes in recruits entering law enforcement, including their perspective toward, and the value they place on, work, supervision, accountability, and community, as well as an unprecedented reliance on technology to get the job done. Batts, Smoot, and Scrivner (2012) discussed these challenges in relation to police leadership, but they also have an impact across all practice domains for police psychologists.

Schmuckler's (1995) suggestion that law enforcement agencies need help in directing change efforts may be quite relevant here. He discussed the potential for psychology to help with organizational activities that range from team building to facilitating strategic management. This type of activity was illustrated in an *APA Monitor on Psychology* article by DeAngelis (2002) about an organizational psychologist working with a large urban department to implement a significant and long-term systemic change effort. DeAngelis's work included a focus on systemic change, integrating adult learning models into management training, and designing and implementing a strategic planning initiative to address crime, quality of life, and management issues. This project resulted in new internal and external partnerships, something of a sea change for law enforcement, and structural changes within the agency have evolved due to her work.

Nicoletti et al. (2011) defined internal and external consultation as a growing subspecialty in the field of police psychology that requires a separate set of IACP-PPSS Guidelines that were created in 2006. In their view, consultants have a much wider role than those managing programs and, in essence, become a resource to law enforcement leadership in a wide variety of areas. Although they focus on the consultant's role in police operations, dealing with family issues, and assisting with threat assessment, they also include a broad portfolio of other consulting possibilities that are available to law enforcement agencies. These activities include those listed next:

- Risk assessment and liability mitigation
- Program evaluation
- Applied research
- Targeted interventions
- Psychological autopsies following suspicious deaths
- Consulting on Title VII issues
- Strategies for responding to returning combat veterans
- Conflict management within the agency
- Leadership development through the use of 360-degree feedback models
- Training to prevent bias-based policing
- Training to improve employee performance in highly specific skills such as report writing or pursuit driving
- Responding to traumatic incidents
- Providing training
- Implementing Employee Assistance Programs (EAPs)
- Supervising peer support and personal wellness programs

In essence, consultants can offer a wide range of services, including those that may be performed by an in-house psychologist, provided that they have the demonstrated competency to do so.

These types of activities suggest that psychology has moved far beyond traditional models of clinical services and that law enforcement organizations are fertile ground for the type of expertise that can help them create and adapt to change. In many respects, they have become learning organizations, and police and public safety psychologists are a primary resource to facilitate that learning whether they are serving as consultants or employed by the organization.

OPERATIONAL SUPPORT

Operational support remains a primary domain, but psychologists have become less involved than previously in operational areas such as assisting in criminal investigations and developing a hostage negotiation capacity and barricade call-out consultation in police departments. One cannot help but wonder if this will remain

so, given the increase in mass shootings and threats to communities that may require skills beyond those of standard law enforcement personnel. Police psychologists could be called on to use their skills for assessing and evaluating maladaptive behavior in a very different context. The success of this practical application of psychology to law enforcement operations clearly helped build credibility for psychology in law enforcement. It also strengthened support for developing other niches of specialized services that contributed to investigating and apprehending criminals, and it may well be an area that continues to grow.

The evolving investigative competency incorporates work with crime victims and witnesses through the use of forensic hypnosis or cognitive interviewing as methods to access greater details about specific crimes and to develop criminal profiles and psychological autopsies. Some of this work has stirred professional controversy, and some psychologists regard it as more art than science (Super, 1999). Moreover, some of the investigative processes can result in inaccurate information that could impede an investigation. One only has to recall the profiles that misidentified the snipers who terrorized the Washington, DC, area for 3 weeks in 2002 and were circulated by “experts” who appeared on various media outlets.

Hibler (1995) concluded that, despite only a 12% rate of valid evidentiary contributions resulting for forensic hypnosis, even this level of enhancement could be valuable when a criminal case is stalemated. He argued that it is a potentially valuable tool for law enforcement provided it is used correctly and with the appropriate controls. The same cautions apply when discussing the even more controversial area of psychological autopsies, or what is sometimes referred to as an “equivocal death analysis” (Gelles, 1995). This process is used as an adjunctive investigative aid and as a tool to help clarify manner of death. It presents a model for assessing an individual’s behavior and personality to develop a better understanding of his or her death.

Victim services are a key and vital component for all police community outreach operations. Activities include psychologists working in behavioral health divisions and having oversight for victim interventions, which can include an array of services, such as victim advocacy, psychotherapy, and 24-hour death notifications. Metropolitan Nashville Police received the 2009 Excellence in Victim Services Award among the world’s large police agencies as a result of continually expanding programs to respond to the needs of crime victims, their family members, and others affected by crime within their community (IACP, 2010).

MODELS OF SERVICE DELIVERY

How services actually are delivered in law enforcement agencies vary. Because law enforcement is not a 9-to-5 job, there may be an expectation that a psychologist will provide 24/7 on-call availability. Although the frequency of being called out varies with the size of the agency and the nature of services provided, it is a responsibility that must be factored into the service delivery model, because it

complicates one's ability to deliver services within the traditional framework of client-driven schedules.

To meet service delivery goals, some models have become more prevalent than others. All have advantages and disadvantages. One of the more commonly used models appears to be the professional services contract (Finn & Tomz, 1997; Scrivner, 1994). Using contracts, departments can hire consultants either to provide a range of services or to contract for a specific activity, such as psychological screening or fitness-for-duty evaluation. Other models include developing a link to an employee assistance program, use of a network of clinical referrals, and peer support services. A model that is more prevalent in large departments provides a full range of psychological services to officers and the organization through in-house psychological service units. In 1995, 61 service activities were identified that are now provided by police psychologists. They were categorized into three general areas: individual service activities, program/technical support, and organizational support. These data are another indication of the growth of psychology in law enforcement (Kurke & Scrivner, 1995).

The survey research of Finn and Tomz (1997) laid out a blueprint for how to establish a program of services for law enforcement. Although their survey primarily addressed how law enforcement stress affects families and included mental health professionals and service providers other than psychologists, the cross-cutting issues they defined are critical to establishing and delivering effective services. The issues include:

- Ensuring that services are accessible and private.
- Developing clear guidelines for confidentiality and statements of informed consent that are consistent with state law.
- Developing a record-keeping system that includes procedures for safeguarding confidential or privileged information.
- Developing and circulating written policies and procedures that include clear distinctions between mandatory and voluntary referrals.
- Ensuring that monitoring systems are in place.
- Educating the user community through training, brochures, or publicity.
- Engaging in program planning that includes relevant stakeholders, such as key law enforcement administrators, labor representatives, officers, and family members.
- Availing yourself of ongoing consultation with the jurisdiction's legal counsel.

Their data provide a reasonable protocol for establishing psychological services. However, to be consistent with most recent practices, when providing health services, an additional element is recommended: assuring that all procedures are compliant with the Health Insurance Portability and Accountability Act (HIPAA, 2000).

ISSUES FACING LAW ENFORCEMENT PSYCHOLOGISTS

As previously referenced, the major issue faced by any psychologist who sets out to work with law enforcement is the need to develop credibility with the rank and file and avoid being seen as a shill for management. There is also a need to overcome significant skepticism about the value and credibility of psychology. Law enforcement officers spend considerable time in court and have had numerous experiences hearing expert testimony. It is not unusual for many to have developed a somewhat cynical perception that expert opinions can be bought, including that of psychologists. Were these two hurdles not enough, psychologists also need to deal with the stigma attached to making contact with psychological services and address the fear that services will not be confidential. In the few departments where management actually is the client, this becomes a reality and not a fear, as confidentiality cannot be guaranteed under those conditions. Fortunately, most law enforcement agencies realize that, if the services are to be effective, they must be confidential, and they therefore respect the need to ensure confidentiality.

Psychologists working with law enforcement also need to be careful not to cross ethical lines through out-of-office contact. Conversations when meeting in the hallway or parking lot may be construed as expert psychological opinion and end up affecting a law enforcement officer's private life or, in some instances, career. The same holds true for socializing with the client base, such as attending retirement dinners and promotion celebrations that take place at the local law enforcement hangout. In any of these situations, casual comments can be misconstrued, and ethical standards that govern the practice of psychology can be compromised. To avoid this type of compromise, many psychologists take the position that it is safer to treat all conversations as clinical contacts regardless of where they occur.

Other ethical dilemmas include restricting practice to what one is trained to do and avoiding exceeding one's competence by becoming all things to all people. Law enforcement personnel cannot be expected to understand all the specializations in psychology and often believe that the psychologist they hire is trained to do anything classified as a domain within psychology. Consequently, psychologists must help staff understand the limits of their training and experience and educate the staff as to what they can and cannot do.

Finally, despite tremendous efforts and success in developing a credible program of services that meet professional integrity standards, the police psychologist must be aware that all could be at risk when a new chief executive arrives with very different ideas regarding the needs of the department. Although it may not be viewed as wise to dismantle a successful program, it has long been a problem in law enforcement that innovation and change, even when successful, are not always sustained at the same level in the face of new management. All police psychologists need to be prepared for such events.

NEW AND EMERGING TRENDS

As the presence of psychology in law enforcement continues to expand, some in leadership positions are using psychologists to assist them in addressing significant national law enforcement issues, such as acrimonious interactions between law enforcement officers and citizens and the use of excessive force. Conversely, there is a growing concern about the use of violence against police officers. The IACP has established a specific program to address this issue, and the Bureau of Justice Assistance and the Office of Community Oriented Policing Services (COPS) have established a Police Health and Wellness Working Group.

Most recently, the IACP and the Justice Office on Violence Against Women developed a multidisciplinary working group that included a police psychologist to study sexual misconduct of officers and develop an executive guide, *Addressing Sexual Offenses and Misconduct by Law Enforcement* (IACP, 2011). The guide is designed to explain the complexities of sexual offenses and misconduct involving police officers and help executives in preventing and investigating incidents.

Within a different venue, the series of Police Chief Roundtables that were conducted in conjunction with the APA annual meetings (1998–2000), the police chiefs who met with psychologists identified needs for assistance to end racial profiling, prevent and respond to cases of excessive force, strengthen police integrity, and develop a greater understanding of police officer fear. They also examined alternatives to arresting the homeless, the prevalence of hate crimes, and skill development for officers in the areas of mediation and anger management (Rabasca, 2000). Venues like the roundtables have the potential to generate ideas for research on psychological issues such as how observing violence affects police officers, particularly in relationship to police officer domestic violence, and how psychological research on self-fulfilling prophecies and stereotype change processes could be helpful in designing interventions to deter ethnic profiling.

Some of the issues identified in the roundtables were consistent with a federally funded project, "Hiring in the Spirit of Service," that was pilot-tested at five sites. This project was designed to aggressively market the service characteristics of law enforcement to recruit and select law enforcement candidates who showed a strong service orientation, in contrast to those who were more interested in the spirit of adventure. Moreover, it examined the psychological screening instruments that had the greatest capacity to accomplish this goal, including pilot-tested instruments developed from job-task analyses that incorporated input from community members. This project, *Hiring in the Spirit of Service* (Scrivner, 2006), helped develop new methods to screen-in applicants for law enforcement and remains in place in four of the five pilot sites.

In yet another venue, the Los Angeles Police Department (LAPD) has moved its psychologists from the consulting rooms into the precincts. Department

psychologists were assigned to operational divisions in an effort to make them more accessible and less intimidating and to enhance their capacity to reach department personnel in a proactive manner. It is believed that this development in the psychological services provided to LAPD personnel will provide opportunities for interventions in the field with personnel who would otherwise have little or no contact with department psychologists and will ensure better follow-up with clients (Gelber, 2003). As with any innovation, this process has to be carefully monitored to ensure that psychologist roles are clearly understood.

The blending of law enforcement experience and advanced degrees in psychology has produced what are known as cop docs, psychologists who have the distinction of also having been law enforcement officers. In fact, as of 2004, the unit chief of the Behavioral Sciences Unit of the FBI Training Academy, a prestigious law enforcement training facility, was an FBI agent who also held a doctorate in psychology.

The cop doc influence has stimulated interest in programs that are based on the belief that officers are more comfortable discussing problems with peers who understand the culture, in contrast to professionals. Moreover, some of the labor unions support the peer process provided that peer support officers are well trained. Many psychologists support the notion of peer support but encourage strong supervision of these programs as well as strong familiarity with referral networks.

The measurement of stress has long been wanting in law enforcement. Beyond survey data, there has been little sophisticated analysis that identifies how stress affects behavior and health. Van Hasselt et al. (2008) developed an assessment tool, the Law Enforcement Officer Stress Survey (LEOSS), a 25-item early-warning stress-screening measure for law enforcement officers. Respondents to the LEOSS rate stressful scenarios on two dimensions: the likelihood of encountering the situation described and the difficulty of each situation for a police officer.

Rather than concentrate only on sources of stress, other psychologists are placing greater emphasis on being proactive, developing resilience, and using prevention models that are designed to better prepare law enforcement officers and their families for what to expect. By providing an orientation to police work and to the changing attitudes and behaviors that start to develop early in the career, they hope to prepare family and friends to understand the transition into law enforcement and how it will impact the new officers and their families (Torres, Maggard, & Torres, 2003).

Family orientation programs are examples of the proactive approach advocated by Delprino (2001), who found that many of the programs developed with CLEF funds that were intended to help with family stress actually were diverted to services for law enforcement officers. In his evaluation of these programs, Delprino expressed concern that family services may not be prioritized and advocated for a more holistic approach that would use resources to develop healthy workplaces that minimize the potential negative effects on the officer and family. Greene (1997) provided a logic model to reduce stress and identify protective factors to ensure

career longevity and strong family relations. The model included individual, family, and organizational risk and protective factors as well as measurable objectives and activities that would positively influence outcomes, and it was the first publication that promoted the use of a logic model to address and evaluate complex police issues. Activities such as peer support programs were proposed to address individual risk factors, trainee spouse and family training programs were proposed to address family risk factors, and early-warning tracking systems were proposed to address organizational management risk factors. Others, such as Artwohl and Christenson (1997), encourage law enforcement officers to develop personal resilience by making healthy lifestyle choices, developing support systems, avoiding overcommitment to the job, retaining a positive focus, and determining what is meaningful in their lives. These trends are compatible with the research on the stress hardiness skills and attitudes that promote transformational coping (Maddi, 2002). This line of research could inform proactive approaches to helping officers and their families tolerate both internal and external stress.

The information age presents other vehicles for proactive responses, including but not limited to a proliferation of Web sites that are designed to bring health-related information to officers and their families. Examples include a site developed by the Metropolitan Nashville Police Department (www.policefamilies.com) and the police stress and health program (PSHP) affiliated with the University of California, San Francisco (www.policestressandhealth.med.nyu.edu). The policefamilies.com site initially was funded by the U.S. Department of Justice and provides families of law enforcement officers with mental health information and access to a wide variety of online family support services. In addition, the Web site includes a free curriculum for agencies to deliver a one- or two-day training for spouses or other adult trainee family members. The Web site also provides a curriculum for law enforcement children. The PSHP is funded through grants from the National Institute of Mental Health and focuses primarily on duty-related stress and improving quality of life. These federally funded Web sites are harbingers of how mental health information and psychological knowledge can be managed electronically to reach greater numbers of law enforcement consumers.

CONCLUSIONS

The trends outlined in this chapter are only a small sample of all that is happening in this field, but they demonstrate a level of growth and impact that would have been unbelievable when the collaboration between psychology and law enforcement was first initiated. They also bring an increase in legal challenges by those adversely affected by the decisions of psychologists, and there is a body of case law developing, particularly on selection and clinical issues.

New demands were placed on law enforcement in the aftermath of 9/11 and following some of the mass shootings that have occurred, which have created a new range of stress factors. In all likelihood psychologists, too, will face new

challenges and could become more involved in assisting officers who are working in threat-sensitive environments, such as preventing and deterring terrorism, a part of today's law enforcement executive portfolio. This development has implications for the emergence of new but related issues, such as dealing with bioterrorism and weapons of mass destruction and understanding how law enforcement needs to be delivered without compromising civil liberties. With psychological services better institutionalized, we believe that law enforcement agencies now have a capacity to meet the changing psychological needs of officers and that psychology has an opportunity to contribute its knowledge base, professional expertise, and research capacity to help solve real-world problems that make a difference in ensuring public safety.

REFERENCES

- Americans with Disabilities Act of 1990, Pub. L. No. 101-336, §2, 104 Stat. 328 (1991).
- ADA Amendments Act of 2008, Pub. L. 110-325 (2009).
- Artwohl, A., & Christensen, L. (1997). *Deadly force encounters: What cops need to know to mentally and physically prepare for and survive a gunfight*. Boulder, CO: Paladin Press.
- Batts, A. W., Smoot, S. M., & Scrivner, E. (2012). Police leadership challenges in a changing world. *New Perspectives in Policing Bulletin*. Washington, DC: U.S. Department of Justice, National Institute of Justice. Retrieved from <https://ncjrs.gov/pdffiles1/nij/238338.pdf>
- Ben-Porath, Y. S., Corey, D. M., & Stewart, C. O. (2011, October). *Predicting law enforcement officer performance outcomes using the MMPI-2-RF*. Paper presented at the meeting of the Police Psychological Services Section, International Association of Chiefs of Police, Chicago, IL.
- Ben-Porath, Y. S., Fico, J. M., Hibler, N. S., Inwald, R., Kruhl, J., & Roberts, M. R. (2011, August). Assessing the psychological suitability of candidates for law enforcement positions. *Police Chief*, 78, 64–70.
- Blau, T. (1994). *Psychological services for law enforcement*. New York, NY: Wiley.
- Bohl, N. (1995). Professionally administered critical incident debriefing for police officers. In M. I. Kurke & E. M. Scrivner (Eds.), *Police psychology into the 21st century* (pp. 169–188). Hillsdale, NJ: Erlbaum.
- Carlier, I. V. E., Voerman, A. E., & Gersons, B. P. R. (2000). The influence of occupational debriefing on post traumatic stress symptomatology in traumatized police officers. *British Journal of Medical Psychology*, 73, 87–98.
- Choe, I. (2005). The debate over psychological debriefing for PTSD. *The New School Psychology Bulletin*, 3(2), 71–82.
- Collins, P. A., & Gibbs, A. C. (2003). Stress in police officers: A study of the origins, prevalence, and severity of stress-related symptoms within a county police force. *Occupational Medicine*, 53, 256–265. doi:10.1093/ocmed/kqg061
- Corey, D. M. (2011). Principles of fitness-for-duty evaluations for police psychologists. In J. Kitaeff (Ed.), *Handbook of police psychology* (pp. 263–293). New York, NY: Routledge.
- Corey, D. M., & Borum, R. (2013). Forensic assessment for high-risk occupations. In R. K. Otto (Ed.), *Forensic psychology* (pp. 246–270). Vol. 11 in I. B. Weiner (Ed.-in-Chief), *Handbook of psychology* (2nd ed.). Hoboken, NJ: Wiley.

- Corey, D. M., Cuttler, M. J., Cox, D. R., & Brower, J. (2011, August). Board certification in police psychology: What it means to public safety. *Police Chief, 78*, 100–104.
- Corey, D. M., & Stewart, C. O. (in press). Under the color of authority: Police officers as violent offenders. In C. A. Pietz & C. Mattson (Eds.), *Violent offenders: Understanding and assessment*. New York, NY: Oxford University Press.
- Dantzker, M. L. (2011). Psychological preemployment screening for police candidates: Seeking consistency if not standardization. *Professional Psychology: Research and Practice, 42*(3), 276–283.
- Davis, R. D., & Rostow, C. D. (2008, December). M-PULSE: Matrix-psychological uniform law enforcement selection evaluation. *Forensic Examiner, 19*–24.
- DeAngelis, T. (2002). Revitalizing Boston's police force. *Monitor on Psychology, 33*(6), 62.
- Delprino, R. P. (2001). *Lessons learned from early corrections and law enforcement family support (CLEFS) programs (NCJ 192287)*. Washington, DC: National Institute of Justice.
- Delprino, R. P., O'Quin, K., & Kennedy, C. (1997). *Identification of work and family services for law enforcement personnel (NCJ 171645)*. Washington, DC: National Institute of Justice, U.S. Department of Justice.
- Eisenberg, T. (1975, November). Labor–management relations and psychological stress: View from the bottom. *Police Chief, 54*–58.
- Finn, P., Talucci, V., & Wood, J. (2000, January). On-the-job stress in policing—Reducing it, preventing it. *National Institute of Justice Journal, 19*–24.
- Finn, P., & Tomz, J. E. (1997, March). *Developing a law enforcement stress program for officers and their families (NCJ 163175)*. Washington, DC: U.S. Department of Justice, National Institute of Justice.
- Fischler, G. L., McElroy, H. K., Miller, L., Saxe-Clifford, S., Stewart, C. O., & Zelig, M. (2011, August). The role of psychological fitness-for-duty evaluations in law enforcement. *Police Chief, 78*, 72–78.
- Fisher, M. A. (2009). Replacing “who is the client?” with a different ethical question. *Professional Psychology: Research and Practice, 40*(1), 1–7.
- Flanagan, C. L. (1995). Legal issues regarding police psychology. In M. Kurke & E. M. Scrivner (Eds.), *Police psychology into the 21st century* (pp. 93–107). Hillsdale, NJ: Erlbaum.
- Gelber, C. (2003, September). LAPD bureau psychologists to hit the streets. *Police Chief, 70*(9), 29–31.
- Gelles, M. G. (1995). Psychological autopsy: An investigative aid. In M. Kurke & E. M. Scrivner (Eds.), *Police psychology into the 21st century* (pp. 337–355). Hillsdale, NJ: Erlbaum.
- Genet, B., & Dowling, F. (2011, March 16). *POPPA: Using volunteer peers to prevent warrior suicides and to build resilience*. Paper presented at the 2011 Suicide Prevention Conference, Washington, DC. Retrieved from http://www.dcoe.health.mil/Content/navigation/documents/SPC2011/Wed_1550_Family-PtoP_Genet.ppsx
- Genetic Information Nondiscrimination Act of 2008 (GINA). Pub.L. 110-233, 122 Stat. 881, 42 U.S.C. 2000 (2008).
- Gershon, R. (1999). *Public health implications of law enforcement stress*. [Video]. Washington, DC: National Institute of Justice, U.S. Department of Justice.
- Greene, L. W. (1997, October). Uplifting resilient police families. *Police Chief, 64*, 70–72.
- Harpold, J. A., & Feemster, S. L. (2002). Negative influences of police stress. *Law Enforcement Bulletin, 71*(9), 1–6.

- Health Insurance Portability and Accountability Act (HIPAA). (2000). Title 45, Subtitle A, Subchapter C, Part 160, General Administrative Requirements.
- Hibler, N. S. (1995). Using hypnosis for investigative purposes. In M. Kurke & E. M. Scrivner (Eds.), *Police psychology into the 21st century* (pp. 319–336). Hillsdale, NJ: Erlbaum.
- Hough, L. M. (2000). *Hiring in the spirit of service: Definitions, possibilities, evidence and recommendations*. Unpublished report, available through The Dunnette Group, Ltd., St. Paul, MN.
- Hough, L. M., & Ones, D. S. (2002). The structure, measurement, validity, and use of personality variables in industrial, work, and organizational psychology. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work and organizational psychology, Volume 1: Personnel psychology* (pp. 233–277). Thousand Oaks, CA: Sage.
- International Association of Chiefs of Police, Psychological Services Section. (2006). *Peer Support Guidelines*. http://www.theiacp.org/psych_section/pdfs/PsycAPeerSupportGuidelines.pdf
- International Association of Chiefs of Police. (2009). *Preventing law enforcement officer suicide: A compilation of resources and best practices*. Alexandria, VA: Author.
- International Association of Chiefs of Police. (2010). IACP/Login Excellence in Victim Services Awards, Part III. *Police Chief*, 77.
- International Association of Chiefs of Police. (2011). *Addressing sexual offenses and misconduct by law enforcement: Executive guide*. Alexandria, VA: Author.
- Inwald, R. E., & Brobst, K. E. (1988). *Hilson Personnel Profile/Success Quotient manual*. Kew Gardens, NY: Hilson Research.
- Kamena, M. D., Gentz, D., Hays, V., Bohl-Penrod, N., & Greene, L. W. (2011). Peer support teams fill an emotional void in law enforcement agencies. *Police Chief*, 78, 80–84.
- Kennedy, D. B., Homant, R. J., & Hupp, R. T. (1998). Suicide by cop. *Law Enforcement Bulletin*, 67(8), 21–27.
- Kitaeff, J. (2011). *Handbook of police psychology*. New York, NY: Routledge/Taylor and Francis.
- Kurke, M. I., & Scrivner, E. M. (Eds.) (1995). *Police psychology into the 21st century*. Hillsdale, NJ: Erlbaum.
- Lieberman, A. M., Best, S. R., Metzler, T. J., Fagan, J. A., Weiss, D. S., & Marmar, C. R. (2002). Routine occupational stress and psychological distress in police. *Policing: An International Journal of Police Strategies and Management*, 25(2), 421–439.
- Maddi, S. R. (2002). The story of hardiness: Twenty years of theorizing, research and practice. *Consulting Psychology Journal*, 54(3), 173–185.
- McNally, R. J., Bryant, R. A., & Ehlers, A. (2003). Does early psychological intervention promote recovery from posttraumatic stress? *American Psychological Society*, 4(2), 45–70.
- Mohandie, K., Meloy, J. R., & Collins, P. I. (2009). Suicide by cop among officer-involved shooting cases. *Journal of Forensic Sciences*, 54(2), 456–462.
- National Advisory Commission on Civil Disorder. (1968). *Kerner Commission report*. New York, NY: Dutton.
- Mishara, B. L., & Martin, N. (2012). Effects of a comprehensive police suicide prevention program. *Crisis*, 33(3), 162–168.
- Nicoletti, J., Allen, S. W., Baum, K. M., Deisinger, G., Garrido, S., Gelles, M. G., . . . Kirschman, E. (2011). Police psychologists as consultants. *Police Chief*, 78, 54–62.

- Niederhoffer, A., & Niederhoffer, E. (1977). *The police family: From station house to ranch house*. Lexington, MA: Heath.
- Omnibus Crime Control and Safe Streets Act of 1968. Pub. L. No. 90-351; 82 STAT.197. (H.R., 5037).
- On the front lines: Police stress and family well-being: Hearing before the Select Committee on Children, Youth and Families, U.S. House of Representatives, 102nd Congress, 1st sess.* (1991, May). Washington, DC: U.S. Government Printing Office.
- Ostrov, E. (1995). Legal, psychological, and ethical issues in police-related forensic psychology evaluations. In M. Kurke & E. M. Scrivner (Eds.), *Police psychology into the 21st century* (pp. 133–145). Hillsdale, NJ: Erlbaum.
- Page, K. S., & Jacobs, S. C. (2011). Surviving the shift: Rural police stress and counseling services. *Psychological Services, 8*(1), 12–22.
- President's Commission on Law Enforcement and Administration of Justice. (1967). *Task force report: The police*. Washington, DC: U.S. Government Printing Office.
- Rabasca, L. (2000, October). Psychologist, police chiefs forge an alliance. *APA Monitor on Psychology, 31*(9), 50.
- Reese, J. T. (1987). *A history of police psychological services*. Washington, DC: U.S. Department of Justice, Federal Bureau of Investigation.
- Reese, J. T., Horn, J. M., & Dunning, C. (Eds.). (1991). *Critical incidents in policing*. Washington, DC: U.S. Department of Justice, Federal Bureau of Investigation.
- Reese, J. T., & Scrivner, E. (Eds.). (1994). *Law enforcement families: Issues and answers*. Washington, DC: U.S. Department of Justice, Federal Bureau of Investigation.
- Reiser, M. (1972). *The police department psychologist*. Springfield, IL: Charles C Thomas.
- Roberts, M. (1995). *CPI Form 434 Police and Public Safety Selection Report interpreter's guide: Interim edition*. Palo Alto, CA: Consulting Psychologists Press.
- Schmuckler, E. (1995). Strategic planning. In M. I. Kurke & Ellen M. Scrivner (Eds.), *Police psychology into the 21st century* (pp. 497–510). Hillsdale, NJ: Erlbaum.
- Scrivner, E. M. (1994). *The role of police psychology in controlling excessive force*. Washington, DC: National Institute of Justice Research Report, U.S. Department of Justice.
- Scrivner, E. (2006). *Innovations in recruitment and hiring: Hiring in the spirit of service*. Washington, DC: Office of Community Oriented Policing Services, U.S. Department of Justice.
- Sewell, J. D. (2002, March). Managing the stress of organizational change. *FBI Law Enforcement Bulletin, 71*(3), 14–20.
- Sheehan, D. C., & Warren, J. I. (Eds.). (2001). *Suicide and law enforcement*. Washington, DC: U.S. Department of Justice, Federal Bureau of Investigation.
- Smith, D. (2002, June). Improving the tools to fight terrorism. *APA Monitor on Psychology, 33*(6), 67.
- Stratton, J. G. (1980). Psychological services for police. *Journal of Police Science and Administration, 8*, 38.
- Super, J. T. (1999). Forensic psychology and law enforcement. In A. K. Hess & I. B. Weiner (Eds.), *The handbook of forensic psychology* (2nd ed., pp. 409–439). New York, NY: Wiley.
- Toch, H. (2002). *Stress in policing*. Washington, DC: American Psychological Association.
- Toch, H. (2012). *COP WATCH: Spectators, Social Media, and Police Reform*. Washington, DC: American Psychological Association.

- Torres, S., Maggard, D. L., & Torres, C. (2003, October). Preparing families for the hazards of police work. *Police Chief*, 70(10), 108–114.
- Trompetter, P. (2011, August). Police psychologists: Roles and responsibilities in a law enforcement agency. *Police Chief*, 78, 52.
- Van Hasselt, V. B., Sheehan, D. C., Malcolm, A. S., Sellers, A. H., Baker, M. T., & Couwels, J. (2008). The Law Enforcement Officer Stress Survey (LEOSS): Evaluation of psychometric properties. *Behavior Modification*, 32(1), 133–151.
- Volanti, J. (1996). *Police suicide: Epidemic in blue*. Springfield, IL: Charles C Thomas.
- White, E. K., & Honig, A. L. (1995). The role of the police psychologist in training. In M. Kurke & E. M. Scrivner (Eds.), *Police psychology into the 21st century* (pp. 257–277). Hillsdale, NJ: Erlbaum.

Evaluating and Assisting Jury Competence in Civil Cases

JENNIFER K. ROBBENNOLT, JENNIFER L. GROSCUP, AND STEVEN PENROD

MARK Twain, ever the sharp-tongued critic of all he observed, turned his eye to the jury in his 1872 volume, *Roughing It*:

The jury system puts a ban upon intelligence and honesty, and a premium upon ignorance, stupidity and perjury. It is a shame that we must continue to use a worthless system because it was good a thousand years ago. . . . I desire to tamper with the jury law. I wish to so alter it as to put a premium on intelligence and character, and close the jury box against idiots, blacklegs, and people who do not read newspapers. But no doubt I shall be defeated—every effort I make to save the country “misses fire.” (p. 343)

More than a century later, in an article titled “Juries: They May Be Broken, But We Can Fix Them,” Supreme Court Justice Sandra Day O’Connor (1997) observed:

Juries usually do their job very well. . . . But juries also have the ability to disappoint us, sometimes to the point of forcing us to question whether we should have jury trials at all. One of this country’s great observers of human nature, Mark Twain, once complained that juries had become “the most ingenious and infallible agency for defeating justice that human wisdom could contrive.” (p. 20)

O’Connor and Twain have respectable company in their criticism of the jury. Richard A. Posner (1995), federal court of appeals judge and former University of Chicago law professor, has sounded similar notes of concern about jury decision making:

In recent years, a series of highly publicized criminal trials in which obviously guilty defendants were acquitted by juries . . . has made the American jury a controversial

institution. Civil juries have rendered some astonishing verdicts as well, ladling out billions in other people's money with insouciance and attracting a drumbeat of criticism from the business community. (p. 14)

In recent years, the civil jury, in particular, has come under attack. In civil cases, juries are asked to determine whether the defendant is liable, to award damages intended to compensate the plaintiff for injuries (compensatory damages), and, sometimes, to award damages intended to punish the defendant for engaging in egregious conduct and deter the defendant and others from engaging in such conduct in the future (punitive damages). Large jury verdicts, such as the \$2.7 million verdict against McDonald's when a customer was burned by hot coffee, verdicts in the millions and even billions against tobacco companies, the \$5 billion in punitive damages levied against Exxon following the *Exxon Valdez* oil spill (Broder, 1997; Kozinski, 1995; Sachdev, 2003), and the recent \$1 billion award to Apple, Inc. in its patent suit against rival smart-phone maker Samsung (Wingfield, 2012) have caused some to conclude that juries are not an effective mechanism for determining liability and awarding damages. Critics contend that civil juries are arbitrary, capricious, and unprincipled in the manner in which they award damages, particularly punitive damages. Advocates of reform argue that civil juries are incompetent to decide the cases before them, biased in favor of plaintiffs, and overgenerous. In addition, they contend that huge sums awarded by juries have fueled a "litigation crisis" and contribute to crippling delays in the civil justice system (e.g., Frank, 1949; Quayle, 1992; Viscusi, 1998; see review in Daniels, 1989). Justice O'Connor, dissenting in *Pacific Mutual Life Ins. Co. v. Haslip*, commented, "Recent years . . . have witnessed an explosion in the frequency and size of punitive damages awards" (1991, p. 1066). Large damage verdicts stir incredible controversy and are typically the objects of substantial media attention.

In this chapter, we consider the two themes advanced by Justice O'Connor in the title of her commentary. We consider evidence on the question of just how "broken" the civil jury is, focusing in the first part of the chapter on research examining the factors that influence jury decision making and decision-making processes in civil cases. In the second part of the chapter, we consider research that has examined several of the mechanisms that have been advanced as fixes for jury problems: juror note taking, juror questioning of witnesses, predeliberation jury discussions, access to trial transcripts, and written witness statements.

CIVIL JURY DECISION MAKING

There is little empirical evidence that the civil justice system is "out of control." Reviews of civil jury decision making conclude that, overall, jurors perform relatively well in determining liability and damages (Greene & Bornstein, 2003; Greene et al., 2002; Hans & Reyna, 2011; Robbennolt, 2002a; Vidmar, 1998). Although punitive damage awards, in particular, have garnered much criticism, research

examining the patterns of punitive damage awards indicates that punitive damages are infrequently sought, infrequently awarded, typically not extremely large, and rarely collected in the amounts awarded (see Daniels & Martin, 1990; Eisenberg & Heise, 2011; Eisenberg, Heise, Waters, & Wells, 2010; Landes & Posner, 1986; Peterson, Sarma, & Shanley, 1987; Rustad, 1991; United States General Accounting Office, 1989; see review in Robbennolt, 2002a). Punitive damages, however, are claimed in more cases than they are awarded and thus remain a threat. In addition, large awards secure substantial media attention, whereas reductions, though common, are not as extensively reported (Garber, 1998).

Nonetheless, there are aspects of jury decisions that are cause for some concern, including unpredictability in jury-determined damage awards such that juries may award differing amounts for seemingly similar injuries (studies reviewed in Saks, 1992; but see Baker, Harel, & Kugler, 2003). Thus, although the overall amount of damages awarded by juries is not out of control, there may be large variability in awards made by juries. Moreover, juries tend to overcompensate plaintiffs with relatively small losses and undercompensate plaintiffs with relatively large losses (Conrad & Bernstein, 1964; King & Smith, 1988). In addition, jurors have some difficulty understanding civil jury instructions (Elwork, Sales, & Alfini, 1982; Hastie, Schkade, & Payne, 1998, 1999a; see reviews in English & Sales, 1997; Lieberman & Sales, 1997) and in translating their judgments into dollar awards (Kahneman, Sunstein, & Schkade, 1998; Wissler, Hart, & Saks, 1999).

LEGAL REFORM EFFORTS

Over the past 30 years, the constitutionality of large punitive damage awards has captured the attention of the U.S. Supreme Court (*BMW of North America v. Gore*, 1996; *Browning-Ferris Ind. v. Kelco Disposal, Inc.*, 1989; *Cooper Industries, Inc. v. Leatherman Tool Group, Inc.*, 2001; *Exxon Shipping Co. v. Baker*, 2008; *Honda Motor Co. v. Oberg*, 1994; *Pacific Mutual Life Ins. Co. v. Haslip*, 1991; *Philip Morris USA v. Williams*, 2007; *State Farm v. Campbell*, 2003; *TXO Production Corp. v. Alliance Resources Corp.*, 1993). The Court has held that the traditional method of awarding punitive damages—that is, the determination of the appropriateness and amount of punitive damages by a jury and subsequent review by both trial and appellate courts—is not “so inherently unfair as to deny due process and be *per se* unconstitutional” but has been willing to consider whether specific jury awards are excessive (*Pacific Mutual Life Ins. Co. v. Haslip*, 1991, p. 1043). In 1996, the Court did, for the first time, find a punitive damage award constitutionally excessive (*BMW of North America v. Gore*, 1996) and did so again 7 years later, in *State Farm v. Campbell* (2003).

More recently, in *Exxon Shipping Co. v. Baker* (2008), the Court more explicitly addressed the relationship between punitive and compensatory damages. Eisenberg, Heise, and Wells (2010) noted that the Court relied on studies such as those noted that indicate there is not a problem of runaway punitive damage awards. Nonetheless, the Court expressed concern about variability in the

punitive-compensatory ratio and reduced a \$2.5 billion award for punitive damages against Exxon to \$500 million in order to establish a 1-to-1 punitive-compensatory ratio in cases arising under maritime law. The Court stated that “the constitutional outer limit may well be 1:1.”

Despite this series of decisions, however, the Court has left the primary responsibility for regulating punitive damage awards and civil litigation to the individual states and to the legislative branch. As a consequence, advocates of tort reform have turned to state and federal legislatures to pursue nonjudicial avenues in attempts to restrain what they perceive as out-of-control civil juries. A number of states have enacted and implemented a variety of measures that are aimed at limiting liability or restricting the incidence or the amount of damage awards (see review in Robbennolt, 2002a, and evidence of the effects of these laws in Eisenberg et al., 2010).

For example, a number of states have implemented rules that eliminate liability under certain circumstances (e.g., Texas Civil Practice & Remedies Code §84.003, limiting the liability of volunteers working for charitable organizations). In addition, many states have implemented limitations on joint and several liability (e.g., California Civil Code §1431.2). Also, a number of states have limited the amount of money that may be awarded either for noneconomic compensatory damages (e.g., pain and suffering) or for punitive damages (e.g., Missouri Revised Statutes §538.210, limiting noneconomic damages in medical liability cases; North Dakota Century Code §32-03.2-11, limiting punitive damages to the greater of two times the compensatory damages or \$250,000). In addition, a number of states have begun to require the jury to be more certain in its damage award decision before it may award punitive damages. Thus, many states require juries to conclude that the evidence is “clear and convincing” that punitive damages are appropriately awarded rather than reaching such a conclusion by a “preponderance of the evidence” (e.g., Alaska Statutes §09.17.020; South Carolina Code Annotated §15-33-135), and they may require that the jury reach a unanimous decision to award punitive damages (e.g., Texas Civil Practice & Remedies Code §41.003). Other reform efforts take some decisions out of the hands of the jury altogether—a few states allow judges to assess the amount of punitive damages to be awarded rather than juries (e.g., Connecticut General Statutes Annotated §52-40; Kansas Statutes Annotated §60-3701). All of these reforms raise the dual questions of how juries make decisions in civil cases and whether their decision-making processes signal serious problems.

JURY DECISION-MAKING PROCESSES

A growing body of research focuses on the process by which juries make decisions (for reviews, see Devine, 2012; Greene et al., 2002; MacCoun, 1993a; Vidmar, 1998). There is empirical support for a number of possibilities that have been advanced as methods by which jurors make decisions. In particular, the “story

model” of juror decision making proposes that jurors combine the evidence that is presented into a narrative story, learn the verdict options, and choose the one that best fits the story they have constructed (Pennington & Hastie, 1993). Robbennolt, Darley, and MacCoun (2003) suggested that jurors operate as “goal managers” as they attempt to use the available verdict options to satisfy a number of goals simultaneously. They proposed that juror decision making operates through a process of parallel constraint satisfaction in which jurors seek to maximally satisfy, in parallel, a variety of potentially competing goals, such as achieving appropriate compensation, effecting deterrence, exacting retribution, and expressing symbolic values.

Investigators have focused on the processes by which jurors determine damage awards (see review in Greene & Bornstein, 2003). One hypothesis is that jurors anchor on an initial value and then adjust this value as they become aware of more and more new facts; this is termed “anchoring and adjustment” (Tversky & Kahneman, 1982, p. 14). Consistent with this view, jurors are influenced by attorney damage award recommendations (Chapman & Bornstein, 1996; Hinsz & Indahl, 1995), exposure to high damage awards in the press (Greene, Goodman, & Loftus, 1991; Viscusi, 2001b), and caps on damage awards (Robbennolt & Studebaker, 1999; Saks, Hollinger, Wissler, Evans, & Hart, 1997). Some jurors who were interviewed after reaching verdicts reported that they arrived at their compensatory damage award by deciding on an amount for each component of damages and then summing to get a total award amount (Goodman, Greene, & Loftus, 1989; Mott, Hans, & Simpson, 2000). Goodman, Greene, and Loftus (1989) found that 27% of jury-eligible adults directed to read written vignettes and award damages reported arriving at their damage award merely by “picking a fair number” rather than engaging in any calculations that would be required by the additive or anchoring and adjustment methods (see also Mott et al., 2000). Interviews with jurors also provide some evidence that final punitive damage awards represent a compromise between high and low amounts advocated by different factions of the jury. In addition, Greene (1989) noted that many punitive damage awards are rounded numbers (e.g., \$1 million, \$500,000), suggesting that minute calculations are not taking place.

Kahneman et al. (1998) found that jurors have difficulty translating into a dollar award their outrage at the defendant’s conduct and their intent to punish the defendant accordingly. Across a variety of cases, mock jurors were relatively consistent in their ratings of outrage and evaluations of the degree of punishment required. However, the dollar amounts of their punitive damage judgments were less consistent. Wissler et al. (1999) demonstrated a similar difficulty for jurors, attorneys, and judges in translating noneconomic damages into dollar awards. While evaluations of injury severity were highly predictable from participants’ ratings of the specific aspects of the injury, damage awards were less predictable.

INFLUENCES ON JURY DECISIONS

Empirical research has identified a variety of factors that play a role in jury decisions about liability or damages, including the nature of the parties' conduct, the severity of the harm, the defendant's wealth, and individual differences among jurors. On the whole, this research indicates that jurors perform their tasks quite well.

NATURE OF DEFENDANT'S (AND PLAINTIFF'S) CONDUCT

As a general rule, conduct is considered legally blameworthy when it is intended to cause harm or involves an undue risk of harm (Prosser & Keeton, 1984). Thus, the intentional or unreasonable nature of the defendant's conduct ought to have some bearing on determinations of civil liability. Consistent with this premise, Greene, Johns, and Bowman (1999) found that, when presented with an automobile accident case, mock jurors and juries were more likely to determine that a defendant was negligent when the defendant's conduct was unreasonable than when it was reasonable. In a related study, Greene, Johns, and Smith (2001) found that mock jurors were more likely to determine that the defendant was negligent when the defendant's conduct was more careless than in a version of the case in which the defendant's behavior was less careless. Other researchers have found that defendants who take more (and more likely effective) precautions are less likely to be found liable than those who have taken fewer (Karlovak & Darley, 1988; see also Wiener et al., 1994).

It is less clear that an offender's conduct should be related to the amount of compensatory damages awarded to an injured party. Nonetheless, there is some evidence that compensatory damages are greater when information about an offender's negligent conduct is available (Greene, Johns, & Smith, 2001; Smith & Greene, 2005).

In contrast, in their discussions of what factors ought to influence punitive damage awards, commentators invariably note that the reprehensibility of the conduct complained of is one such relevant factor (Owen, 1994). Indeed, punitive damages cannot be awarded unless the defendant's conduct is "outrageous, because of the defendant's evil motive or his reckless indifference to the rights of others" (Restatement [Second] of Torts, §908(2)). Prosser and Keeton (1984) have noted that for punitive damages to be awarded, there

must be circumstances of aggravation or outrage, such as spite or "malice," or a fraudulent or evil motive on the part of the defendant, or such a conscious and deliberate disregard of the interests of others that the conduct may be called willful or wanton. (pp. 9–10)

One of the guideposts identified by the U.S. Supreme Court in *BMW* (1996) was the degree of reprehensibility of the defendant's conduct. The Court noted that the reprehensibility of the defendant's conduct was "perhaps the most important

inducium of the reasonableness of a punitive damages award” because a punitive damages award should reflect the “enormity” of the defendant’s offense (p. 575).

Consistent with the legal theory, research suggests a positive relationship between the reprehensibility of the defendant’s actions and the size of the punitive damage award. Cather, Greene, and Durham (1996) found that participants awarded greater punitive damages in response to high-reprehensibility scenarios than they did in response to low-reprehensibility scenarios. In another study, using 768 jury-eligible adults and an audiotaped trial, Horowitz and Bordens (1990) found that the reprehensibility of a defendant manufacturer’s conduct (operationalized as the length of time the defendant was aware of the harmful effects of its product) was not correlated with compensatory damages but was significantly correlated with punitive damages. Relatedly, Robbennolt (2002b) found that jury-eligible citizens and trial court judges who rated a defendant’s conduct as more offensive awarded higher amounts of punitive damages.

One concern related to jurors’ evaluation of plaintiff and defendant conduct is hindsight bias (Fischhoff, 1975; Guilbault, Bryant, Brockway, & Posavac, 2004; Hawkins & Hastie, 1990). The law requires assessments of conduct to be made from an *ex ante* perspective—that is, the tortiousness of the conduct should be judged without taking into account any consequences that might have resulted. As Prosser and Keeton noted (1984, §31):

The actor’s conduct must be judged in the light of the possibilities apparent to him at the time, and not by looking backward “with the wisdom born of the event.” The standard is one of conduct, rather than of consequences. It is not enough that everyone can see now that the risk was great, if it was not apparent when the conduct occurred. (p. 170)

However, psychological research has shown the difficulties jurors experience when trying to make such *ex ante* judgments. Once the outcome of an action is known, other information about the action is interpreted in light of its outcome, making it hard to see how anyone could have expected things would turn out otherwise (Fischhoff, 1982). Thus, knowing the outcome of an actor’s conduct can influence judgments of the foreseeability or risks, the likelihood that harm would occur, and the likely severity of any resulting harm. Experimental studies that have examined how mock jurors assess conduct and its risks in hindsight find evidence of hindsight bias at work—when mock jurors are informed of a bad outcome, they tend to evaluate the actor’s conduct as being less reasonable, to be more likely to think that precautions ought to have been taken, and to see the bad outcome as having been more foreseeable (see Hastie, Schkade, & Payne, 1999b; Kamin & Rachlinski, 1995; LaBine & LaBine, 1996; Lowe & Reckers, 1994). At the same time, however, hindsight bias can be weaker following negative outcomes, in real-world cases, and in tasks that do not require numerical probability estimates—all characteristics of the civil context within which jurors operate (Guilbault, Bryant, Brockway, & Posavac, 2004).

OUTCOME SEVERITY

One factor thought to be an important consideration in the awarding of damages, but not liability, is the severity of the outcome to the plaintiff (*BMW of North America v. Gore*, 1996). However, the expected relationship between injury severity and damage awards is complex. Compensatory damages logically should be greater when the injuries and other damages are more severe, because the resulting medical bills, lost wages, and pain and suffering are increased. However, this is not necessarily the case with punitive damages, which are aimed not at compensating the plaintiff and making him or her whole, as are compensatory damages, but at punishing and deterring the defendant. In fact, many have argued that punitive damages ought to be scaled to the heinousness of the conduct and *not* to the magnitude of the harm (Galanter & Luban, 1993). Nonetheless, the U.S. Supreme Court has concluded that punitive damages ought to have some reasonable relationship to the damage suffered by the plaintiff (*BMW*, 1996; *State Farm v. Campbell*, 2003; *TXO*, 1993).

The empirical findings regarding the influence of injury severity on liability determinations have been mixed. Several studies have demonstrated a relationship between the severity of the injury to the plaintiff and determinations of liability (Bornstein, 1998; Greene et al., 1999; Van der Keilen & Garg, 1994); others have found no relationship (Green, 1968; Greene, Johns, & Bowman, 1999; Peterson, 1984; Taragin, Willett, Wilczek, Trout, & Carson, 1992). In a meta-analysis, Robbennolt (2000) found that injury severity had only a small effect ($r = .03$) on whether the plaintiff received a payment (including civil liability verdicts and settlements).

A number of studies have shown that more compensatory damages are awarded when injuries are more severe. In a study of 8,231 medical malpractice cases, Taragin et al. (1992) found that the likelihood of a plaintiff obtaining a payment and the amount of that payment (either settlement or jury verdict) increased with the severity of the injury (see also Peterson, 1984). Experimental research has produced similar findings. For example, Wissler, Evans, Hart, Morry, and Saks (1997) found that pain and suffering awards were strongly influenced by information about the nature, characteristics, and consequences of the injury (see also Greene, Woody, & Winter, 2000; Robbennolt, 2000).

The evidence with respect to the influence of injury severity on punitive damages is mixed. Several studies of actual cases have shown a correlation between the severity of the injury to the plaintiff and punitive damage awards (Eisenberg, Goerdt, Ostrom, Rottman, & Wells, 1997; Rustad, 1992). Experimental studies have also demonstrated this relationship. Cather, Greene, and Durham (1996) investigated the influence of the severity of the injury to the plaintiff on the amounts of compensatory and punitive damages awarded by jury-eligible adults in response to written vignettes. Overall, damage awards in a personal injury case were higher when the plaintiff was more severely injured than when the plaintiff was only mildly injured, but this relationship was not observed in other types of cases (product liability and insurance bad-faith cases). When they examined

punitive damages in particular, the investigators did not find significant differences in the amounts awarded to severely injured and mildly injured plaintiffs (see also Kahneman et al., 1998).

The extent of the actual injury suffered by the plaintiff is not the only important factor related to the severity of the harm inflicted by the defendant. Equally important are the injuries that could have resulted from the defendant's conduct. In *TXO* (1993), the Court recognized that the relationship between injury severity and punitive damages could not be quantified in a numerical ratio and cited a common example of circumstances in which punitive damages many times the compensatory damages would be appropriate:

For instance, a man wildly fires a gun into a crowd. By sheer chance, no one is injured and the only damage is to a \$10 pair of glasses. A jury reasonably could find only \$10 in compensatory damages, but thousands of dollars in punitive damages to teach a duty of care. We would allow a jury to impose substantial punitive damages in order to discourage future bad acts. (p. 459)

The Court determined that it was appropriate to take into account the harm that could have occurred due to the defendant's actions along with harm that did indeed occur and to take into account the "possible harm to other victims that might have resulted if similar future behavior were not deterred" (p. 460). This approach to the relationship between the severity of the injury to the plaintiff and the punitive damage award was echoed in *BMW* (1996), where the Court found that an important guide for the review of punitive damage awards is the ratio of the punitive damage award to the "harm or potential harm" caused by the defendant (see also *State Farm v. Campbell*, 2003).

Karlovac and Darley (1988, p. 289) noted that, in determining an actor's negligence, the legal system takes into account not only the severity of the actual outcome but also "the severity of all the harms that *could* foreseeably have eventuated from a risky action." In a series of studies, Karlovac and Darley investigated the influence of the severity of the potential harms risked by an actor on the judgments of participants. Using undergraduate participants and tape-recorded stories accompanied by slides, they examined the effect on judgments of varying the degree of the maximum possible harm that could have resulted from an actor's risky action. Consistent with legal theory, judgments of negligence were influenced by the severity of the harm risked. Moreover, judgments of the degree of punishment that was perceived as appropriate were determined by the severity of the harm risked. These judgments were similarly affected both before and after participants were informed of the actual outcome.

In another experimental study, Robbennolt (2002b) examined the influence of both the actual severity of the injury to the plaintiff and the severity of the potential harm. She found that the severity of the actual injury influenced mock jurors' compensatory damage awards but that both the actual and potential injury influenced punitive damage awards.

DEFENDANT'S WEALTH

Another factor thought by some to be influential in juror decision making is the wealth of the defendant. Because the purpose of compensatory damages is to “make the plaintiff whole”—that is, to compensate the plaintiff for his or her losses—the wealth of the defendant should play no role in the amount of compensatory damages awarded. However, the purposes of punitive damages are different from those of compensatory damages. To punish or deter a wealthy defendant, the amount of punitive damages awarded must be sufficient to make an impact on him or her (Simpson, 1996; see also Abraham & Jeffries, 1989; and Arlen, 1992; see generally *Pacific Mutual Life Ins. Co. v. Haslip*, 1991; *State Farm v. Campbell*, 2003; *TXO*, 1993).

Hans and Ermann (1989) found that mock jurors differentiated between the financial resources available to an individual defendant (i.e., “Mr. Jones”) as compared to a defendant corporation (i.e., the “Jones Corporation”) and awarded a plaintiff suing the corporation more compensation than a plaintiff suing the individual. However, regression analysis indicated that there was not a consistent effect of the defendant’s presumed resources on awards. Rather, awards were more strongly linked to judgments about the defendant’s recklessness, with participants attributing more recklessness to the corporation than to the individual.

A subsequent study attempted to delineate the distinction between the impact of a “corporate identity” on juror decisions and the impact of the defendant’s wealth. Using written case materials, MacCoun (1996; see also MacCoun, 1993b) found that jury-eligible adults treated corporations differently from individual defendants, such that larger compensatory damage awards were assessed against the corporate defendant than against the wealthy individual defendant. However, MacCoun found that the compensatory damages awarded against the wealthy individual were no greater than those awarded against the poor individual.

Thus, it appears that there is little evidence for a deep-pockets effect, at least in terms of the impact of the wealth of the defendant on compensatory damage awards. This is as it should be; as noted earlier, the wealth of the defendant does not impact the extent of the plaintiff’s damages or the amount of money appropriate to compensate the plaintiff. In contrast, wealth arguably should influence punitive damage awards. Results of several experimental studies indicate that the wealth of the defendant does influence the punitive damages awarded. Robbennolt (2002b) found that both jury-eligible citizens and trial court judges awarded more in punitive damages against a wealthier defendant than against a less wealthy defendant. Across three different cases, Greene, Woody, and Winter (2000) found that higher amounts of punitive damages were assessed against wealthier defendants. Similarly, Kahneman et al. (1998) found that the size of a corporate defendant’s annual profit influenced punitive damage awards.

COMPLEX CIVIL TRIALS

On occasion, civil trials can be lengthy, complex, or both. The legal decisions required can be complicated, and the evidence can be voluminous, technical,

or ambiguous. Some cases turn on difficult statistical, epidemiological, or other scientific evidence. Jurors work hard to understand complex evidence, are assisted by jury members who do understand it, and use generally sound approaches to make sense of complicated evidence (see, e.g., Diamond, Rose, Murphy, & Smith, 2006; Hans, Kaye, Dann, Farley, & Albertson, 2011; Vidmar & Diamond, 2001). Jurors do tend to find the evidence in longer trials more challenging but still believe that they understand it reasonably well (Cecil, Lind, & Bermant, 1987). When evidence is ambiguous, jurors may examine it in more nuanced ways (Horowitz, Bordens, Victor, Bourgeois, & ForsterLee, 2001). At the same time, jurors can have difficulty understanding and using complicated evidence about scientific causation (Sanders, 1998), evaluating the validity of scientific evidence (Kovera, McAuliff, & Herbert, 1999; McAuliff, Kovera, & Nunez, 2009), and understanding statistical evidence (Thompson & Schumann, 1987). In some instances, the ways in which experts and attorneys present evidence to jurors can compound these difficulties (see Sanders, 1998; Lempert, 1993).

INDIVIDUAL CHARACTERISTICS OF DECISION MAKERS

A variety of individual difference variables have been explored in an effort to determine their relationship to legal judgments (Ford, 1986; Litigation Sciences, 1993). In general, demographic variables such as age, gender, and social class are of limited value in predicting judgments. However, some personality and attitudinal variables have proven to be somewhat more useful. Ellsworth (1993) attempted to determine which components of the juror decision-making process are influenced by juror attitudes. Ellsworth noted that legal decisions are inherently imprecise and require that the decision maker resolve numerous ambiguities and engage in a great deal of interpretation. Thus, there is ample room for juror attitudes to influence juror decisions. Ellsworth proposed that attitudes might influence verdicts in three distinct ways:

1. Attitudes may influence jurors' evaluation of the credibility of witnesses.
2. The inferences drawn by jurors—inferences that are based in part on the jurors' attitudes—may influence their construction of a narrative summary of the evidence.
3. Attitudes may influence how jurors apply the judge's instructions regarding the law to the facts as they have constructed them.

Ellsworth found support for the conclusion that attitudes influence verdicts in all three of these ways.

In their investigation of attitudes toward the police and toward due process, Casper, Benedict, and Perry (1989) hypothesized that attitudes might influence damage awards in a civil rights action through their role in shaping the processing of the testimony to which jurors are exposed. They found that attitudes operated to influence damage awards to some extent through their influence on jurors'

interpretation of trial testimony but that the attitudes also retained an independent effect on awards.

In a series of studies, Hans and Lofquist (1992, 1994) investigated jurors' attitudes toward civil litigation. They found that jurors in actual tort cases had strong negative views of both the frequency and the legitimacy of civil lawsuits and believed that civil damage awards are too high. However, jurors also agreed that jurors generally do a good job and found their own jury experience to be positive. Hans and Lofquist (1994) found that their 7-item scale measuring juror attitudes toward civil litigation comprised two separate factors, one measuring attitudes toward the worth of civil litigation and a second measuring beliefs about the abilities of civil juries. Moreover, Hans and Lofquist (1992) found a significant correlation between the jury members' average scores on the civil litigation scale and the jury's damage awards, such that the more strongly the jurors believed there was a litigation crisis, the lower the damages awarded.

Other researchers investigating the relationship between attitudes toward the civil litigation system and legal decisions have reported similar findings. In telephone interviews, Moran, Cutler, and De Lisa (1994) found that attitudes toward tort reform predicted mock juror verdicts in both civil and criminal fictional cases. Similarly, Greene et al. (1991) found that the scores of jury-eligible adults on a scale measuring attitudes toward tort reform and damages (e.g., whether there is an insurance crisis, the influence of media on attitudes about civil lawsuits, and beliefs about attorney credibility and damage requests) were significantly correlated with damage awards, such that those most supportive of tort reform gave lower damage awards. In addition, the researchers found a significant positive correlation between participants' estimates of the frequency of large damage awards and the amount of damages they awarded. Moreover, they found that such attitudinal measures were more reliable predictors than demographic variables (see review in Robbennolt & Studebaker, 2003).

Finally, the schemas that people hold for certain kinds of accidents and injuries can also influence their decisions in civil cases (see Hans & Dee, 2003; Hart, Evans, Wissler, Feehan, & Saks, 1997). Such schemas can influence the extent to which a particular account tells a story that is perceived as plausible and coherent and can influence how the evidence is understood and remembered (see Davies, 2009; Smith & Studebaker, 1996).

JURORS VERSUS JUDGES

Many believe that judges would engage in qualitatively different kinds of decision making than jurors. For example, in his concurrence in *BMW of North America v. Gore* (1996), Justice Breyer noted that one cannot "expect jurors to interpret law like judges, who work within a discipline and hierarchical organization that normally promotes roughly uniform interpretation and application of the law" (p. 596). However, there is a paucity of research regarding the comparison between the

decision making of jurors and the decision making of judges. What we do know suggests that judges and jurors may engage in similar decision-making processes.

One of the earliest comparisons of judges and jurors was conducted by Kalven and Zeisel (1966; Kalven, 1964). They asked judges to report, for cases tried before them, how the jury decided the case and how they would have decided it had it been a bench trial. While judges knew the juries' verdict and, thus, their reports were not made independent of this knowledge, this was an important foray into comparisons of judge and jury decision making. Across 4,000 civil cases, the researchers found that judges and juries agreed 78% of the time as to the liability of defendants. In terms of the amount of damage awards, they found that, when both the judge and jury decided in favor of the plaintiff, juries awarded more damages 52% of the time and judges would have awarded more damages 39% of the time, with approximate agreement in 9% of the cases. On average, Kalven and Zeisel found that juries awarded 20% more in damages than judges report they would have awarded (see also Eisenberg et al., 2005; Heuer & Penrod, 1994b). Studies have found that rates of judge–jury agreement are not associated with the complexity of the trial (Eisenberg et al., 2005; Heuer & Penrod, 1994b; see also Kalven & Zeisel, 1966).

Judges and jurors are also similar in a number of additional aspects of decision making. For example, judges and jurors react similarly to potentially biasing, but inadmissible, evidence (Landsman & Rakos, 1994), are similarly susceptible to cognitive biases (Guthrie, Rachlinski, & Wistrich, 2001, 2007), and have similar difficulties in assessing statistical (Wells, 1992) and scientific (Kovera et al., 1999; Kovera & McAuliff, 2000) evidence.

Evidence regarding how jurors and judges compare in their assessments of damages has been somewhat mixed. Archival research and some experimental research has found many similarities in the damage award decision making of judges and jurors (Clermont & Eisenberg, 1992; Eisenberg, LaFountain, Ostrom, & Rottman, 2002; Eisenberg et al., 2006; Robbennolt, 2002b; Vidmar, 1995; Wissler et al., 1999; see Robbennolt, 2002a, for a review). However, the results of other experimental research have suggested some differences (Hastie & Viscusi, 1998; Hersch & Viscusi, 2004; Viscusi, 2001a). To the extent that there are differences, it is difficult to exclude the possibility that litigants channel different kinds of cases to judges and juries (see Eisenberg & Heise, 2011).

SUMMARY

On balance, existing research on jury decision making in civil cases suggests that the process is, if not perfect, at least orderly. Jurors seem to give systematic consideration to factors such as the severity of outcomes and the reprehensibility of the alleged acts and do not seriously misuse information about a defendant's wealth. Furthermore, the decisions of juries seem to stack up reasonably well against the decisions of other, arguably more expert, decision makers. These conclusions,

however, must be qualified insofar as the body of scientific research on which they are based is not large. Most of the research on civil jury decision making is of relatively recent vintage, and it is easy to imagine that our understanding of these processes will be much richer as more research is conducted.

AIDS TO JURY DECISION MAKING

Although our survey of research on jury decision making in civil cases suggests that the civil jury is probably not as broken as some critics would like us to believe,

[t]here is a little disagreement among social scientists (or even among most lawyers and judges) that traditional trial procedures have largely failed to take account of how jurors process new information. Recognition of this shortcoming, and the need to modify existing trial procedures, is the first step toward improving the effectiveness of the jury as an integral component of our justice system. (Munsterman, Hannaford-Agor, & Whitehead, 2006, p. 5)

Commentators have been quite inventive in advancing recommendations for jury aids. For instance, in her 1997 article, Justice O'Connor recommended:

In my view, the first level for reform is in the courtroom. . . . Jurors should be allowed, and encouraged, to take notes at trial. I frankly cannot understand the resistance to this practice. . . . Taking notes is a way for a person to make sense of the information being received. . . . and perhaps most importantly for the juror, to take an active, rather than a passive, part in what is going on. (pp. 23–24)

Attorney Kenneth Adamo (1996) recommended:

Let Them Take Notes. . . . Allowing note-taking is almost de rigeur if juror comprehension and interest are to be maximized. . . . Allow the Jury to Ask Questions. . . . If you want an interested and knowledgeable jury, especially as trial proceeds, you need to provide for juror questions. (pp. 354–355)

Among the recommendations in the Final Report of the Blue Ribbon Commission on Jury System Improvement for the State of California (Kelso, 1996) are these:

Adopt a Rule of court which requires the trial court to inform jurors of their right to take written notes [and] . . . adopt a [rule] recommending that judges permit jurors to submit written questions to the court which, subject to the discretion of the trial judge and the rules of evidence, may be asked of witnesses who are still on the stand. (p. 1504)

In their volumes on the jury, Abramson (1994) and Adler (1994) also advocated use of these procedures and drew approval from Judge Posner (1995):

For complex modern cases, both Abramson and Adler propose a series of reforms to make the jury's task easier: allowing jurors to take notes and ask questions. (p. 16)

Regarding predeliberation jury discussions, the Honorable B. Michael Dann (1993) observed:

The “rules for getting the floor” during trial ought to be modified to permit at least limited discussions of the evidence among jurors who wish to participate, thereby establishing a form of “speaking rights” for the decision makers. (p. 1265, citations omitted)

Persuaded by studies of group psychology and their own experiences, legal commentators argue that the restriction on predeliberation discussions is anti-educational, nondemocratic, and unnecessary to ensure, at least in its present form, an orderly or otherwise fair trial. In what has been noted as one of the early court decisions to positively evaluate jury discussions during the trial, Judge Ditter in *United States v. Wexler* (1987) observed:

The duty of a juror involves complex thought processes: assimilating and comprehending the evidence, determining credibility issues, recalling the evidence, putting it all into context and relative degrees of reliability, participating in discussions, and making informed decisions. Jurors need all the help they can get and their only source of untainted information and assistance is from those who share with them the responsibility for making the ultimate decisions. (p. 969)

He further noted that predeliberation jury discussions

will make them more attentive, more apt to be interested and involved, more likely to focus on the issues as they unfold. Jurors who have been told, figuratively, to clap their hands over their mouths, who cannot share their ideas and impressions, may tend to clap their hands over their minds as well. (p. 969)

Despite the enthusiasm for jury aids expressed by such authorities as Justice O’Connor and the California Commission, these procedures are rather controversial and not universally endorsed. Arguments for and against such aids have been advanced by the courts, legal scholars, and social scientists alike, and the debate over these procedures is far from new. Appellate decisions concerning juror questions date back to as early as 1825 and decisions about note taking to at least 1900. Contemporary commentary is also abundant.

Although many appellate courts have addressed these issues, there is no clear consensus on their advantages and disadvantages. There is some consensus on how the procedures should be evaluated, at least insofar as the same criteria appear across cases repeatedly. Unfortunately, the appellate judges writing these decisions draw on their own experiences as the principal evidence concerning the strengths and weaknesses of the methods. Of course, until recently, there was little in the way of systematic evidence about the impact of the procedures for judges to rely on. Although late to the scene, the social science community has generated some discussion and research on a variety of jury aids in the past 25 years.

JUROR QUESTIONS

In an early review of the case law on jury questions, Purver (1970) noted that most courts concluded it is not improper but is a matter within the discretion of the trial judge. Courts have disagreed about whether juror questions should be encouraged or discouraged (see *State v. Graves*, 1995; *Williams v. Commonwealth*, 1997). However, the general conclusion remains the same as that advanced by Purver: Many courts are reluctant to encourage or to discourage juror questions. Some jurisdictions do discourage the procedure, and Texas has prohibited it (Cano, 2001; Wolff, 1990). Other states provide for juror questions by state law (e.g., *Lawson v. State*, 1996) or by court rule (e.g., *Cohee v. State*, 1997; *State v. Greer*, 1997).

Overall, appellate decisions reflect some disagreement among judges regarding the propriety of this procedure, but it is not difficult to find cases in which courts advise caution. The Second Circuit (*United States v. Douglas*, 1996) takes a firmly skeptical view:

In three recent cases, we have considered the issue of juror questioning of witnesses. . . . All three decisions expressed varying degrees of disapproval of juror questioning, though only [*United States v. Ajmal*] concluded that the questioning that occurred warranted reversal of the conviction. (p. 326)

In *United States v. Ajmal* (1995), the court chastised:

The district court's decision to invite juror questioning was not necessitated by the factual intricacies of this banal drug conspiracy, nor was it prompted by the urging of the jurors themselves. . . . Not surprisingly, the jurors took extensive advantage of this opportunity to question witnesses, including Ajmal himself. Such questioning tainted the trial process. (pp. 14–15)

The Seventh Circuit in *United States v. Feinberg* (1996) also expressed reservations:

Whether to permit jurors to ask questions is a decision best left to the discretion of the district judge. . . . However, implicit in his exercise of discretion is an obligation to weigh the potential benefit to the jurors against the potential harm to the parties, especially when one of those parties is a criminal defendant. . . . In the vast majority of cases the risks outweigh the benefits. (p. 336)

Empirical Research. There is some research examining the potential advantages and disadvantages of juror questions. In addition to the Purver (1970) piece (which has been updated with cases through 1995), particularly thorough discussions of the impact of juror questions can be found in McLaughlin (1982), in Wolff (1990), and in the Eighth Circuit opinion in *United States v. Johnson* (1989). A field study by Sand and Reiss (1985) allowed jurors in 26 trials to submit questions to the judge to be asked of witnesses. Unfortunately, there was no nonquestion control group. A pilot field experiment in Dane County, Wisconsin, by Penrod, Linz, and

Rios (1983) randomly assigned criminal trials in one courtroom to question versus no-question conditions. Trials in the no-question control group were supplemented with trials from a second courtroom in which questions were not permitted (for a total of 31 trials), which created a partial confound between judges and question asking.

More recent field research on juror questions has investigated the type of juror questions posed and jurors' reactions to the courts' responses during deliberations (Diamond, Rose, Murphy, & Smith, 2004; Diamond et al., 2006). Questions from 50 civil cases were collected as part of the Arizona jury filming project investigating predeliberation discussion (Diamond, Vidmar, Rose, Ellis, & Murphy, 2003). Diamond et al. (2006) reported on all 829 questions collected (76% of which were answered by the courts). On average, 16.6 questions were asked per trial. These questions covered a range of topics, including legal elements, fact confirmation, and witness credibility.

Diamond et al. (2004) focused on the 197 questions (24%) that were not answered by the courts, at least one of which occurred in 39 of the 50 trials. Unanswered questions also covered a range of topics, including legal standards, witness credibility issues, and financial issues (Diamond et al., 2004). Mott (2003), who summarized the frequency and content of 2,271 juror questions from both civil and criminal cases, including those analyzed by Diamond et al. (2004) and Diamond et al. (2006), found that jurors asked a median of seven questions per case and asked twice as many questions in criminal cases as in civil cases. The most common questions were about general, nonexpert witnesses (Mott, 2003). In these three field studies, question-asking procedures were not manipulated, juries were not randomly assigned to question-asking procedures, and there were no no-question control groups. Jurors also were not asked to respond to any questions probing their reactions to questions.

Heuer and Penrod (1988, 1989, 1994a, 1994b) conducted two courtroom field experiments that examined the consequences of permitting jurors to take notes and direct questions to witnesses during trial. Data for the first experiment were obtained from 550 jurors, 29 judges (sitting in 63 different trials), and 95 lawyers—all of whom participated in the same 67 Wisconsin state court trials (Heuer & Penrod, 1988, 1989). Data for the second experiment included 75 civil and 85 criminal trials in courtrooms from 33 states; there were 1,229 jurors, 103 judges, and 220 lawyers (Heuer & Penrod, 1994a, 1994b). The procedures in the experiments were similar and included approximately equal numbers of criminal and civil trials. In both studies, judges received packets of materials including (a) instructions about the combination of questioning and note-taking procedures they were to employ in their next jury trial; (b) suggestions about how to administer the procedures; and (c) questionnaires to be completed by the judge, the jurors, and the lawyers at the conclusion of the trial. All respondents were questioned about demographic information and asked their general evaluations of the trial, the trial participants, and the experimental procedures. Judges and lawyers were asked to complete

questionnaires while the jury was deliberating. In most trials, questionnaires were completed before participants left the courtroom.

Proponents and critics of jury questions have advanced a number of proposals for questioning procedures and advanced numerous hypotheses about the impact of juror questions. These ideas guided the development of the Heuer and Penrod (1988, 1989, 1994a, 1994b) procedures and dependent measures. A number of courts have stated their preference about the procedures to be employed if juror questions are permitted. In *United States v. Polowichak* (1986), the court disapproved allowing juror questions to be stated within the hearing of other jurors and suggested that the district court require jurors to submit questions in writing, without revealing the question to other jurors, at which point the court could pose the question after determining that it was proper. Similar procedures have been approved in state and federal cases, such as *Commonwealth v. Urena* (1994), *State v. Alexander* (1997), *State v. Greer* (1997), *United States v. Bush* (1995), *United States v. Richardson* (2000), *United States v. Stierwalt* (1994), and *United States v. Feinberg* (1996; where the court disallowed permitting jurors to ask their questions orally but did not overturn the defendant's conviction because the jury asked only "innocuous" questions). Courts and commentators have also suggested that both attorneys be allowed to make any objections to a juror's written question at a bench conference and that the judge's ruling on these objections be made outside the hearing of the jury (Cano, 2001; Dann, 1996; *DeBenedetto v. Goodyear*, 1985; *State v. Howard*, 1987).

In the Heuer and Penrod (1988, 1989, 1994a, 1994b) studies, for trials randomly assigned to permit juror questions, judges received instructions much like those outlined earlier, and they generally followed the recommendations. In trials assigned not to include juror questions, judges were asked to disallow direct questions to witnesses. In the Wisconsin study (Heuer & Penrod, 1988, 1989), jurors were permitted to pose questions in 33 trials and asked a total of 88 questions (2.3 questions per trial). Two-thirds were directed to prosecution witnesses and one-third to defense witnesses. Fifteen of the 88 questions (17%) drew objections from the prosecution, the defense attorney, or both. There was considerable agreement about which questions were objectionable, for both attorneys typically objected to the same questions. These questions frequently concerned evidence that both attorneys knew was inadmissible (e.g., questions about insurance in civil cases).

In the national study (Heuer & Penrod, 1994a, 1994b), questions were permitted in 71 trials, although questions were posed in only 51 (a finding that suggests that jurors do not necessarily act on their license). Not counting questions submitted but not asked (due to lawyer objections or screening by the judge), jurors asked an average of 4.4 questions per criminal trial (median = 1.3) and 5.1 questions per civil trial (median = 1.8). In both civil and criminal trials, questions were asked at the rate of about 1 question per 2 hours of trial time (the median was only .25 questions per hour, with a mode of 0).

In the national study (Heuer & Penrod, 1994a, 1994b), the majority of jury questions were directed to prosecution or plaintiff witnesses (79% in civil trials,

77% in criminal trials). Although this may suggest some disparity in the rate of questions directed to opposing sides, when the amount of time that prosecution and defense witnesses spent on the stand was considered, the rate was fairly evenly distributed: Questions were submitted to prosecution witnesses at a rate of approximately 0.7 questions per hour of testimony compared to approximately 0.5 per hour for defense witnesses. Twenty-four percent of the jurors' questions were objected to by one or both attorneys. As in the Wisconsin study, the attorneys in trials in the national study largely agreed about which questions were objectionable: 44% of the questions that were objected to were challenged by both lawyers. Defense attorneys reported that 81% of their objections were sustained, compared to 79% for prosecutors. The high levels of co-objection and judges' sustaining of those objections reflect the fact that many questions concerned evidence that was inadmissible (e.g., questions that would have called for hearsay testimony).

Evaluation of the Major Possible Advantages of Juror Questions. These studies make it possible to evaluate a variety of possible advantages of allowing jurors to ask questions.

Do Juror Questions Promote Juror Understanding of the Evidence and Issues? Scully (1996) argued, "[One] method of improving juror understanding is to allow the jurors to ask questions of expert witnesses. This would be helpful because an expert may overlook information that the jurors believe is crucial to making a decision" (pp. 650–651). In *Williams v. Commonwealth* (1997), the court similarly observed, "[A] juror may, and often does, ask a very pertinent and helpful question in furtherance of the investigation" (p. 155, citation omitted). Similar arguments were advanced by the courts in *Schaefer v. St. Louis & Suburban R. Co.* (1895), *Krause v. State* (1942), and *Ratton v. Busby* (1959).

Heuer and Penrod's (1988, 1989, 1994a, 1994b) findings generally support the proposition that juror questions enhance juror understanding. In the Wisconsin cases, jurors permitted to ask questions were more satisfied that the questioning of witnesses had been thorough, seldom believed that a witness needed to be further questioned, and were more convinced that they had sufficient information to reach a responsible verdict. In the national study, jurors in question-asking trials were asked how helpful their questions were for clarifying the evidence, clarifying the law, and getting to the truth. Overall, the answers indicated modest but positive appraisals. Jurors in trials in which questions were permitted also indicated that they were somewhat better informed by the evidence and were more confident that they had sufficient information to reach a responsible verdict in trials.

Do Juror Questions Help Jurors Get to the Truth? Some advocates of juror questions believe they can do more than aid understanding. The Supreme Court of Massachusetts observed in *Commonwealth v. Urena* (1994), "Indeed, there are

asserted benefits to juror questioning of witnesses, such as the opportunity for jurors to more fully understand the evidence, . . . enhanced attentiveness of jurors, and furtherance of the truth-seeking ideal” (p. 1205). McLaughlin (1982) observed, “Rather than an indifferent battle of legal minds with jurors as mere spectators, a trial is above all a search for truth. . . while justice is blind, jurors need not also be” (pp. 697–698). In *State v. Kendall* (1907), the court held that there was nothing improper in a juror asking a question with the apparent purpose of discovering the truth. The court pointed out that jurors ask often pertinent questions that help in advancing the investigation. In other cases (e.g., *Hudson v. Markum*, 1997; *Louisville Bridge & Terminal Co. v. Brown*, 1925; *State v. Graves*, 1995; *United States v. Callahan*, 1979; *United States v. Thompson*, 1996; *White v. Little*, 1928), courts have observed that juror questions might aid the jury in finding the truth.

Heuer and Penrod’s (1988, 1989, 1994a, 1994b) findings do not offer much support for this proposition. In both the Wisconsin and the national studies, judges and attorneys were asked whether they believed juror questions helped get to the truth. Their answers indicated that they did not expect juror questions to help get to the truth, and, after participating in a trial in which questions were permitted, both groups reported that the questions were not very helpful.

Do Juror Questions Increase Juror, Attorney, or Judge Satisfaction With the Trial or the Verdict? As Judge B. Michael Dann (1996) put it, “The more active jurors are at trial, the more attentive they are to the proceedings. And juror satisfaction with the whole experience is enhanced” (p. 6). Jurors’ overall satisfaction with their trials was assessed in both the Wisconsin (Heuer & Penrod, 1988, 1989) and the national studies (Heuer & Penrod, 1994a, 1994b). In both, the conclusion was that jurors were quite satisfied with their experiences, and their assessments were not influenced by the availability or use of juror questions. Jurors’ satisfaction with their verdicts and attitudes toward jury service were similarly unaffected by their opportunity to ask questions. The lawyers and judges in the national trial were also asked how satisfied they were with the jury’s verdict. Overall, lawyers and judges indicated that they were reasonably satisfied (with judges somewhat more satisfied than attorneys); these assessments were also not influenced by the presence or absence of juror questions.

Do Juror Questions Alert Counsel to Issues That Require Further Development? In *United States v. Callahan* (1979), the court observed, “If a juror is unclear as to a point in the proof, it makes good common sense to allow a question to be asked about it. If nothing else, the question should alert trial counsel that a particular factual issue may need more extensive development” (p. 1086). In Heuer and Penrod’s (1988, 1989, 1994a, 1994b) studies, lawyers and judges were asked whether questions had signaled juror confusion about the law or case evidence. Lawyers and judges expected juror questions to provide useful information about the jury’s

thinking, but after participating in a trial in which questions were allowed, judges and lawyers agreed that questions did not yield these benefits.

Evaluation of Possible Disadvantages of Juror Questions. Several possible disadvantages of allowing jurors to ask questions were also examined in these studies.

When Jurors Are Allowed to Ask Questions, Do They Become Advocates Rather Than Neutrals? In *United States v. Johnson* (1989), Chief Judge Donald Lay observed,

The fundamental problem with juror questions lies in the gross distortion of the adversary system and the misconception of the role of the jury as a neutral factfinder in the adversary process.... The neutrality and objectivity of the juror must be sacrosanct. (p. 713)

The Second Circuit raised the same concern in several cases, including *United States v. Thompson* (1996) and *United States v. Bush* (1995, p. 515), "Although we reaffirm . . . that juror questioning of witnesses lies within the trial judge's discretion, we strongly discourage its use. The most troubling concern is that the practice risks turning jurors into advocates, compromising their neutrality." McLaughlin (1982) described this phenomenon as the "'twelve angry men' syndrome," (p. 702) in which jurors lose their objectivity and begin to direct accusatorial questions to the witness.

Diamond et al. (2006) addressed the question of jurors becoming advocates through asking questions in two ways. First, they examined whether the questions posed by jurors were facially argumentative or phrased as if they were a cross-examination, and they found that none of the jurors' questions was asked using adversarial language. Second, they examined whether the likely answer to a question could be easily predicted or favored one of the parties to the case. Using this definition, only 69 (8.3%) of the questions were classified as argumentative.

Heuer and Penrod (1988, 1989, 1994a, 1994b) examined several types of evidence that indirectly address this concern. One was the pattern of jury decisions. The verdict pattern in the national study (Heuer & Penrod, 1994a, 1994b) indicated that jury questions did not have a significant effect on the verdicts. Heuer and Penrod also asked the judges what their preferred verdict would have been in those trials. This allowed the researchers to examine the rate of judge and jury agreement. The agreement rate was not affected by juror questions; judges and jurors agreed on the verdict in 69% of the cases. Although the agreement was slightly higher in cases in which questions were permitted (74% versus 65%), this difference was not statistically significant. In addition, there was no evidence that either lawyer was perceived less favorably as a result of the questioning procedure (a result that might be expected if jurors lost their neutrality). In fact, attorneys on both sides were perceived somewhat more favorably in trials where questions were permitted. All of this research indicates that jurors are not likely to become advocates and lose neutrality by asking questions.

Do Jurors Ask Improper Questions? One concern of trial attorneys is that jurors, because they are untutored in the law, will ask impermissible questions and should therefore be discouraged from asking any questions at all. Chief Judge Donald Lay in *United States v. Johnson* (1989) observed, "Because lay jurors will not understand the rules of evidence, they may well ask impermissible questions, such as those directed at the defendant's character" (p. 713). The court in *Day v. Kilgore* (1994) expressed the concern this way: "Questions from a jury, untrained in the rules of evidence, may be improper or may solicit information that is either irrelevant or outside of the evidence presented" (p. 518). Examples of jurors asking impermissible questions can be found in the case law. For example, in *Maggart v. Bell* (1931), one juror asked the defendant whether he was covered by accident insurance.

Despite these sorts of reservations, Heuer and Penrod (1988, 1989, 1994a, 1994b) found that although jurors do not know the rules of evidence, they nonetheless ask appropriate questions. In the Wisconsin study, both lawyers and judges reported that they did not expect juror questions to be inappropriate or inept, and they did not find them to be so. Lawyers and judges in the national study who participated in a trial with juror questions reported that improper questions were not a problem.

Diamond et al. (2006) concluded that jurors' questions were largely about understanding the facts and evidence presented in the cases, meaning that the large majority of questions were not improper. They also concluded that discussion of these questions did not improperly dominate deliberations. Even the questions that the courts declined to answer were not always considered to be improper. There were a variety of reasons why those questions were unanswered, only some of which related to jurors asking about legally improper information (Diamond et al., 2004). Mott (2003) also found that jurors' questions probed topics relating to the evidence and legal procedures, indicating that questions are an attempt to clarify testimony and information already presented to them and not an attempt to garner additional or inadmissible information.

Do Juror Questions Interfere With Attorney Trial Strategies? Attorneys in the Wisconsin study were also asked whether juror questions brought up information that they had deliberately omitted; this question was asked because preliminary questioning of trial attorneys revealed a fear that juror questions would play havoc with attorney trial strategies: "'Trials should continue to be what parties deem to present to jurors,' not an extended search by those jurors for an underlying truth," as one attorney quoted by Tripoli (1997, pp. 104–105) put it. However, attorneys who participated in trials in which questions were permitted reported that this was not a problem.

Is Trial Counsel Reluctant to Object to Inappropriate Juror Questions? Numerous courts have refused to reverse when counsel did not object, during trial, to permitting jurors to ask questions (e.g., *Chicago Hansom Cab Co. v. Havelick*, 1869) or

to improper juror questions (e.g., *Louisville Bridge & Terminal Co. v. Brown*, 1925). In considering whether counsel should be *required* to object to improper juror questions in order to preserve the point for appeal, the court in *State v. Sickles* (1926) asked whether this standard was appropriate when objections raise the risk of offending the juror.

In *Day v. Kilgore* (1994), the South Carolina Supreme Court expressed the concern and noted the actions of its brethren in Texas:

When either the judge or the jury departs from their assigned roles, the lawyer is confronted with the dilemma of whether to object and risk alienating the judge or jury, or remain silent and risk waiving the issue for appeal purposes. . . . Confronted with a barrage of appeals where the jury departed from its normal role of passive listeners, the Texas Supreme Court issued an absolute prohibition on the procedure. (*Morrison v. State*, 1992; pp. 517–518)

Heuer and Penrod's (1988, 1989, 1994a, 1994b) studies show that lawyers are not immobilized by such fears. In the national and Wisconsin studies, lawyers objected to 20% and 17% (respectively) of questions submitted by jurors. In the national study, lawyers objected to at least one question in 40% of the trials in which at least one question was asked. Of course, Heuer and Penrod's practice of suggesting that jurors submit questions in writing (a procedure now formally adopted in some jurisdictions) so that attorneys may object in private offered some protection to an objecting attorney. Furthermore, if an objection was sustained, judges were asked to explain the ruling to the jury to minimize the possibility that jurors would draw an adverse inference.

Do Jurors React Negatively When Their Questions Are Not Answered? Judges may be reluctant to allow juror questions because they fear jurors may have a negative reaction when some of their questions are unanswered due to legal reasons (Diamond et al., 2004). Jurors could be offended by the lack of an answer, they may become unduly focused on the topic of the question or on the reason for the lack of an answer, and they may attempt to generate their own answers to these questions.

Diamond et al. (2004) obtained records of the types of unanswered questions jurors asked in 50 civil cases and the jurors' discussions of the unanswered questions in the deliberation room. Judges did not answer 24% of the 820 questions asked in these trials. Unanswered questions were present in 39 of the 50 trials, and they covered a range of topics including legal standards, insurance coverage, attorney fees, and witness character. The majority of unanswered questions (62%) were never discussed by jurors at all, and when unanswered questions were discussed, that discussion was limited except for questions about insurance coverage and attorney fees, about which discussion was longer. Jurors also did not appear to react negatively to the lack of answers. When an unanswered question was mentioned,

the most common response was explicit acceptance that the question was not answered (49%), and explicit negative reactions were rare (4%). Finally, jurors were not likely to generate their own answers in the absence of a response from the court. Jurors attempted to generate an answer to 16% of the questions, and many of these “answers” were culled from the trial evidence. Overall, Diamond et al. concluded that jurors do not have extremely negative reactions to unanswered questions, and she recommends that courts instruct jurors that some of their questions will not be answered to help prevent any negative reaction.

General Conclusions About Juror Questioning of Witnesses. The research reviewed above permits a number of conclusions regarding the effect of juror questioning of witnesses:

- Jury questioning promotes juror understanding of the facts and issues.
- Juror questions do not clearly help get to the truth.
- Juror questions do not alert trial counsel that issues require more extensive development.
- Juror questions do not increase participants’ satisfaction with the trial, the judge, or the verdict.
- Jurors do not become advocates rather than neutrals.
- Although jurors do not know the rules of evidence, they ask appropriate questions.
- Juror questions do not interfere with attorney trial strategies.
- Counsel are not reluctant to object to inappropriate juror questions.
- Jurors do not react negatively when their questions are not answered.

JUROR NOTE TAKING

The courts have frequently considered the merits of permitting jurors to take notes during trials (an exhaustive, 50,000-word review of the case law can be found in Larsen, 1996). Traditionally, courts were cool to the idea, because juror literacy was far from uniform and there were reservations about allowing some jurors to rely on memory and others on notes. As the court in *Sligar v. Bartlett* (1996) observed, “The common law rule grew from a suspicion that a ‘lettered’ juror would be revered, and thus excessively persuasive to the other jurors who could not read or write. To guard against this note taking was prohibited” (p. 1385).

The illiteracy objection has largely disappeared (though see *State v. Triplett*, 1992, for an expression of concern on this matter), but there are other objections to note taking, and the courts in many jurisdictions have not resolved fully their stance on the question. Thus, in *United States v. Darden* (1995) the court took a disapproving or cautious approach: “Note taking by jurors is not a favored procedure. As we have stated, trial courts are properly concerned that the juror with the most detailed notes, whether accurate or not, may dominate jury deliberations” (p. 1537). In contrast, in

Crum v. State (1997), the court took a neutral to approving stance on note taking: "The decision to allow jurors to take notes and consult them during deliberation is within the sound discretion of the trial judge" (p. 15).

Empirical Research. Several investigators have examined the advantages and disadvantages of juror note taking. The research methods employed in these studies have varied widely. In a field study, Flango (1980) assigned one civil trial and one criminal trial to a note-taking condition and compared them to two non-note-taking control trials. A field study by Sand and Reiss (1985) permitted jurors to take notes in 14 criminal and 18 civil trials. Neither of these field studies used random assignment of cases. At the other end of the methodological spectrum is a laboratory study by Hastie (1983), who randomly assigned six-person simulated juries to a note-taking or non-note-taking condition, presented a videotape of an actual armed robbery trial, and had them deliberate to a verdict.

Rosenhan, Eisner, and Robinson (1994) also conducted a laboratory experiment, in which 144 jury-eligible college students and jurors were randomly assigned to note-taking or no-notes conditions, viewed a 75-minute videotaped simulation of a civil trial, and were tested for recall and comprehension of trial material immediately afterward. Horowitz and colleagues have conducted several studies using a similar experimental methodology (ForsterLee & Horowitz, 1997; ForsterLee, Horowitz, & Bourgeois, 1994; Horowitz & Bordens, 2002; Horowitz & ForsterLee, 2001). In this line of research, the participants watched a videotaped mock trial of a toxic tort case with multiple and differentially worthy plaintiffs. The presence or absence of note taking was manipulated, in addition to other trial procedures, and liability judgments, compensation awards, and measures of cognitive performance were recorded. ForsterLee et al. (1994) and ForsterLee and Horowitz (1997) presented the mock trial to individual jurors, and Horowitz and ForsterLee (2001) and Horowitz and Bordens (2002) tested the effect of note taking on juries. The two Heuer and Penrod (1988, 1989, 1994a, 1994b) field experiments described earlier also manipulated the opportunity for jurors to take notes.

Note-Taking Procedures. In the Heuer and Penrod (1988, 1989, 1994a, 1994b) studies, when a trial was assigned to the note-taking condition, judges were asked to permit jurors to take notes during all phases of the trial and instruct the jurors about this permission as soon as practicable after the jury was impaneled. Judges were also provided suggested instructions about note taking. In trials assigned to non-note taking, judges were asked to bar notes.

Across Heuer and Penrod's (1988, 1989, 1994a, 1994b) two studies, juror note taking was allowed in 135 trials. When jurors were allowed to take notes, most did so (66% in the Wisconsin study, 87% in the national study), but they did not take extensive notes. In the Wisconsin study (Heuer & Penrod, 1988, 1989), where trials lasted an average of 2.3 days, jurors took an average of 5.4 pages of notes. In the

national study, the juror averages for civil trials (which lasted an average of nearly 10 days) were 14.4 pages of notes, and for criminal trials (which lasted an average of nearly 6 days), 7.1 pages of notes. In the national study, Heuer and Penrod (1994a, 1994b) estimated that jurors in both types of trials took an average of 0.6 pages of notes per hour of trial time. ForsterLee and Horowitz (1997) found that jurors' notes ranged in length from a couple of sentences to 7 pages. However, that study was based on an hourlong video, not a full trial, which could explain the less extensive note taking.

Evaluation of the Possible Advantages of Juror Note Taking. Juror note taking is thought to have a number of advantages. Several of the studies examining the practice have examined these potential benefits.

Does Juror Note Taking Improve Decision Making? The primary method of measuring the quality of juror and jury decision making used by Horowitz and colleagues was to present the mock jurors with four differentially worthy plaintiffs and determine if their compensatory damage awards appropriately distinguished among these plaintiffs. ForsterLee et al. (1994) and Horowitz and ForsterLee (2001) found that note-taking jurors were better able to distinguish among these plaintiffs than non-note-taking jurors. ForsterLee and Horowitz (1997) also found that note taking improved jurors' ability to distinguish among plaintiffs with varying levels of injury, especially when the evidence was less complex and pre-instructions on the law were provided. Overall, this research indicates that note taking may help improve the quality of jury decision making, at least as measured by the legal appropriateness of compensatory awards.

Does Juror Note Taking Serve as a Memory Aid? Some earlier studies (e.g., Flango, 1980; Sand & Reiss, 1985) reported that jurors found the note-taking procedure helpful as a memory aid, and some courts (e.g., *Marbley v. State*, 1984; *Reece v. Simpson*, 1983; *State v. Trujillo*, 1994; *United States v. Carlisi*, 1940) have endorsed this seemingly reasonable proposition, arguing that there is no reason why notes should not be made by jurors, given that judges and lawyers make notes and jurors' notes might aid their memories and enable them to consider the evidence more intelligently. In *Densen v. Stanley* (1919), the court concluded that note taking can assist the jurors in arriving at a correct and fair verdict. As the Oklahoma Court of Criminal Appeals (*Cohee v. State*, 1997) observed,

We find that jurors may benefit from notes in several ways: (1) jurors may follow the proceedings more closely and pay more attention as they take notes for later use; (2) jurors' memories may be more easily and reliably refreshed during deliberations; (3) jurors may make fewer requests to have portions of trial transcript read back during deliberations; and (4) the ability to use their notes may result in increased juror morale and satisfaction. (p. 2)

In both of the Heuer and Penrod (1988, 1989, 1994a, 1994b) studies, jurors were asked a variety of questions about their recall of the evidence. In the Wisconsin study (Heuer & Penrod, 1988, 1989), jurors even completed a multiple-choice test of their understanding of the judge's instructions. Heuer and Penrod's conclusion from both studies (1988, 1989, 1994a, 1994b) was that there was no evidence to suggest that note taking produced better recall. Although we believe the evidence from these field studies is more compelling than findings from prior, but weaker, field research, it is still difficult to argue that there is no memory advantage to juror note taking. As in the other field studies, the measures used by Heuer and Penrod may not have been sufficiently sensitive to detect memory benefits. The researchers relied on quite general measures of recall rather than measures tailored to the facts of each case.

In assessing memory enhancement effects, the benefits of experiments in controlled environments (e.g., mock trials) are clear: Such studies are much more powerful test settings because they can control the content of the trial, vary the complexity of the trial, and directly measure juror performance as a function of their opportunity to take notes. The Rosenhan et al. (1994) laboratory experiment on note taking did test jurors' recall and comprehension. Jurors were asked questions tailored to the case they had observed, and they had their notes available for reference while answering the questions. On a measure of recall, note takers outperformed non-note takers by a modest but significant margin. The authors reported that 7 of the 10 highest scores on the recall measure were attained by note takers whereas 8 of the 10 lowest scores were attained by non-note takers. Among note takers, the authors found a positive relationship between the quantity of notes taken and recall and between the degree of organization in notes and recall. The authors found no effect for notes on jurors' verdict preferences.

Horowitz and colleagues also measured recall of trial evidence in a laboratory experiment by having the participants free-recall the evidence they could remember and later coding that evidence as probative, nonprobative, and evaluative. Generally, note taking resulted in an increase in the amount of information recalled (ForsterLee et al., 1994). Note takers reported more probative evidence than non-note takers, and they reported less nonprobative and evaluative evidence than non-note takers (ForsterLee & Horowitz, 1997; ForsterLee et al., 1994; Horowitz & Bordens, 2002). However, this difference was more pronounced when the jurors were pre-instructed and when the case was less ambiguous, indicating that notes are more helpful when jurors are given a framework for the evidence and it is easily understood (ForsterLee & Horowitz, 1997). Horowitz and ForsterLee (2001) measured the effect of note taking on memory by presenting mock juries with recognition items containing trial facts and plausible lures of facts not presented in trial. Note-taking juries made fewer recognition errors and fewer false alarms for both pro-plaintiff and pro-defendant lures.

Given these results, it appears that note taking has the potential to increase accurate recall of trial evidence, but ForsterLee et al. (1994) also investigated *how*

note taking improves memory. Jurors' access to their notes was manipulated, in addition to whether they were permitted to take notes at all. Although the researchers did observe an improvement in recall for the note-taking jurors, they observed no difference in memory between the note takers with and without access to their notes. Based on this finding, the authors concluded that the aid to memory provided by note taking may occur at the encoding stage.

Does Note Taking Increase Juror Satisfaction With the Trial or the Verdict? Dann (1996) concurred with the Oklahoma court in *Cohee* that note taking can have a salutary effect on jurors: "The more active jurors are at trial, the more attentive they are to the proceedings. And juror satisfaction with the whole experience is enhanced" (p. 6). In the Wisconsin experiment, Heuer and Penrod (1988, 1989, 1994a, 1994b) detected a slight increase in juror satisfaction with trials, but the finding was not replicated in the national experiment, nor did note taking influence jurors' verdict confidence in the national study. Of course, jurors were already quite satisfied with their verdicts and the procedures in their trials (on 9-point scales, with higher scores indicating greater satisfaction, jurors' mean satisfaction with the verdict was 7.0, and their mean satisfaction with the trial procedure was 7.2), so there may be a ceiling effect in operation. In the laboratory experiments measuring satisfaction, Horowitz and ForsterLee (2001) also found that note-taking juries reported more satisfaction with the deliberations, and note-taking juries believed they were more efficient as a group than non-note-taking juries.

Evaluation of the Possible Disadvantages of Juror Note Taking. A variety of possible disadvantages of allowing juror note taking have also been examined.

Do Jurors' Notes Produce a Distorted Record of the Case? According to the majority in *Thornton v. Weaber* (1955), jurors are unable to distinguish important from unimportant evidence and will therefore miss the important evidentiary points while noting the unimportant ones. These biased notes, according to this analysis, will then distort the jurors' evaluation of the trial evidence. Similarly, the defendant in *State v. Triplett* (1992) argued (unsuccessfully) that juror notes had distorted the evidence. In *United States v. Davis* (1900), the court considered whether it was appropriate for a judge, upon noticing that two jurors had occasionally taken notes, to direct them to discontinue and turn their notes over to the marshal. In ruling that note taking was improper, the appeals court stated, "Without corrupt purpose, [the juror's] notes may be inaccurate, or meager or careless, and loosely deficient, partial, and altogether incomplete" (p. 839). This critique suggests that juror note taking will interfere with the accurate transmission of information from the courtroom to the deliberation room.

In both of Heuer and Penrod's (1988, 1989, 1994a, 1994b) studies, they concluded that notes tended to be a fair and accurate record of the trial proceedings. With respect to the most important trial outcome, Heuer and Penrod found no evidence

that verdicts were affected by note taking. In the experimental studies that measured liability verdicts, there was also no evidence that verdicts were affected by note taking (ForsterLee et al., 1994; Horowitz & ForsterLee, 2001). The absence of a main effect for note taking on verdicts clearly indicates that note taking does not systematically favor the defense or prosecution/plaintiff.

There have been mixed results for the effects of note taking on compensatory awards. Some studies have found no effect (ForsterLee et al., 1994); other studies have found that note takers made reduced awards (Horowitz & ForsterLee, 2001). ForsterLee and Horowitz (1997) found that note takers awarded higher compensatory damages than non-note takers. However, they concluded that this was actually a reflection of better discrimination between the plaintiffs and was a positive effect.

Heuer and Penrod (1988, 1989, 1994a, 1994b) also asked jurors whether their notes tended to be valuable records of the trial or mostly doodles, and they reported that they were considerably more likely to be accurate records. More impressive perhaps are the comments from one of the participating judges in this experiment, who was initially quite skeptical about jurors' note-taking abilities. Upon reviewing the notes from eight trials, his report included these comments:

Approximately one-third of all the jurors . . . took surprisingly detailed notes. The notes were so clearly written and organized that I had little trouble determining what went on in the case. . . . Many of the notes were extremely articulate and well organized. I concluded that jurors have far better notetaking capacity than I had realized. (as cited in Heuer & Penrod, 1988, p. 250)

ForsterLee and Horowitz (1997) also measured the quality of jurors' notes and observed that, as the length of the notes increased, the jurors' ability to correctly recall probative trial evidence and ability to distinguish between plaintiffs also increased.

Is Note Taking Distracting? In *Fischer v. Fischer* (1966), the court concluded that jurors should not be allowed to take notes, because poor note takers are likely to be distracted. A similar argument was made in *Matthews v. Commonwealth Edison Co.* (1995): "Unless a case is complex note taking by jurors is unwarranted and may even interfere with the jurors' ability to observe the witness and attend the testimony" (p. 7). Hastie (1983) similarly suggested that note takers might be distracted from assessing witness credibility. Flango (1980) suggested that note takers may distract non-note takers or themselves by doodling whereas McLaughlin (1982) suggested that jurors making notes on a trivial point will miss important evidence. The majority in *Thornton v. Weaber* (1955) and Flango (1980) similarly suggested that note takers could not keep pace with the trial and would therefore miss important points. Jurors in note-taking trials in both Heuer and Penrod (1988, 1989, 1994a, 1994b) experiments overwhelmingly reported that the trial did not proceed too quickly for

them to keep pace with the proceedings: 85% of the jurors in the Wisconsin study and 87% of the jurors in the national study said this was not a problem. In both studies, note takers and non-note takers in note-taking trials agreed that they were not distracted by note takers. In the Wisconsin study, moreover, the judges and attorneys said they neither expected nor found note taking to be distracting.

Do Note Takers Unduly Influence Non-Note Takers? Several decisions have expressed concern that more prolific note takers might have inappropriate influence on other jurors. The court in *Fischer v. Fischer* (1966) concluded that jurors should not be allowed to take notes because skilled note takers will gain a marked influential advantage over other jurors. In *Thornton v. Weaber* (1955), the court cleverly speculated that note takers might have more influence because they might seem more alert and informed than non-note takers. And in *United States v. Davis* (1900), the court speculated that a juror who can refer to notes could have undue influence in conflicts of memory.

In both of Heuer and Penrod's (1988, 1989, 1994a, 1994b) studies, note takers and non-note takers agreed that note takers should not and did not have an advantage over non-note takers during deliberations. In addition, in the Wisconsin experiment, Heuer and Penrod found no evidence that better-educated jurors participated more in the jury's deliberations when aided by trial notes.

Do Juror Notes Favor One Side or the Other? Flango (1980) suggested that note taking might favor the prosecution or plaintiff if jurors take notes early in the trial but lose their enthusiasm and take fewer notes later in the trial. Neither of the Heuer and Penrod (1988, 1989, 1994a, 1994b) studies found jurors were more diligent note takers during earlier phases of a trial. In the Wisconsin study, jurors in note-taking trials did report slightly less favorable impressions of the defense attorney, but the effect was small and the pattern was not reproduced in the national study. As noted earlier, the national study also revealed no effect of note taking on verdicts. However, ForsterLee and Horowitz (1997) concluded that the increased compensatory awards given by note takers suggested a pro-plaintiff bias. Overall, the majority of the research indicates that note taking does not favor either the prosecution or the defense.

Does Juror Note Taking Consume Too Much Trial Time? Several appellate decisions have indicated that note taking is acceptable only if it does not require substantial court time (e.g., *Cahill v. Baltimore*, 1916; *Tift v. Towns*, 1879). Hastie (1983) speculated that note taking might lengthen jury deliberations as jurors try to resolve discrepancies in their notes. However, Hastie came to the same conclusion as Heuer and Penrod (1988, 1989, 1994a, 1994b): In none of the studies was deliberation time affected by juror note taking. In the Wisconsin study, the jurors in note-taking trials did not report any increase in the difficulty of agreeing on the meaning of the law

on the application of the judge's instructions to trial facts or in acrimonious debate. The jurors in the national study indicated that little deliberation time was devoted to discussions of notes (the median estimate was 1%; the mode was 0%).

General Conclusions About Juror Notes. The research reviewed permits a number of conclusions regarding the effect of jurors' note taking:

- Note taking may improve the quality of jury decision making.
- Juror notes probably are a minor memory aid, perhaps at the encoding stage.
- Juror note taking may increase juror satisfaction with the trial, the judge, or the verdict.
- Jurors' notes do not produce a distorted view of the case.
- Note takers can keep pace with the trial.
- Note-taking jurors do not distract other jurors.
- Note takers do not have an undue influence over non-note takers.
- Juror note taking does not favor either the prosecution or the defense.
- Juror note taking does not consume too much time.

PREDELIBERATION JURY DISCUSSIONS

Traditionally, discussion among jury members has been prohibited until both parties rest and the jury is sent to deliberate. Arguments against permitting jurors to discuss the case prior to deliberation include that jurors will be more likely to judge the case prior to the presentation of all evidence, shared biases among the jurors will be created, and discussions will result in a pro-plaintiff or pro-prosecution bias. The potential benefits include increased comprehension; increased ability to recall evidence; a reduction in individual biases due to comparison of views with other jurors; and a reduction in inappropriate discussions with other jurors, friends, and family members.

In 1995, Arizona made several reforms to the jury system. One of these reforms made it permissible to instruct civil jurors that they may discuss the case prior to deliberation if all the jury members are present (Arizona Rules of Civil Procedure 39(f)). Jurors are also instructed to avoid making premature judgments about the verdict. A few other states are also considering such reforms (for a review of these issues, see Diamond et al., 2003; Hannaford, Hans, & Munsterman, 2000; Hans, Hannaford, & Munsterman, 1999).

Research has been conducted on the effects of this procedure. Hannaford et al. (2000) studied the effects of predeliberation jury discussions on jury decision making with real juries in Arizona. One strength of the study was that juries were randomly assigned to either a discussion or no-discussion condition. In the discussion condition, juries were given the Arizona instruction permitting discussions when all jury members were present, and in the no-discussion condition, juries were instructed not to discuss the case prior to deliberation. At the close of

each case, all of the trial participants were given questionnaires, including the judge, the attorneys, the parties to the case, and the jurors. Overall, there were few beneficial effects of discussions on jury decision making, but there were also few harmful effects. Jurors reported that they believed the discussions were helpful, but the other trial participants noted few differences between the discussion and no-discussion juries.

Although discussions have the potential to increase comprehension of trial evidence and improve decision-making quality, discussions did not significantly improve jurors' comprehension of evidence, nor did they improve the quality of their ultimate decisions. Some results indicated that jurors permitted to discuss the case with other jurors engaged in less discussion of the case with friends and family members, which is a positive effect of permitting discussions. However, the discussion jurors were also more likely to engage in impermissible informal discussions with other jurors than jurors instructed not to discuss the case at all.

As for the potential negative effects of discussions, the main concern over allowing discussions is the potential for prejudice, but there were no significant differences observed between the discussion and no-discussion juries on their level of prejudice. A pro-plaintiff bias was observed, but only in one of many counties studied. From these findings, the authors concluded that permitting predeliberation jury discussions did not have a dramatic positive or negative impact on jury decision making (Hannaford et al., 2000).

More recently, the same jury reform was evaluated by observing jury deliberations in actual Arizona trials, in addition to measuring the effect of discussions by surveying the trial participants (Diamond et al., 2003). Diamond and her colleagues randomly assigned 50 juries to either a discussion or no-discussion procedure. The trials were videotaped, and all of the jurors' interactions in the jury room during breaks and deliberation were videotaped. All trial participants were also surveyed about their perceptions of the trial. Jurors permitted to have discussions spent significantly more time discussing the cases than no-discussion jurors, and this discussion tended to improve recall of the evidence.

In general, these discussions were appropriately focused on the trial evidence. Statements of prejudice were made by some jurors, but no jury reached a verdict prior to deliberations and the prejudices were not necessarily related to the final verdict. Jurors in both conditions were equally likely to discuss the cases with people outside the jury. As was observed by Hannaford and her colleagues (2000), discussion jurors did not strictly adhere to the rule requiring all members of the jury to be present for discussions. The authors suggested that modifications in the procedure permitting jury discussions might improve their effectiveness, including appointment of a temporary leader for the jury to ensure that the discussion topics were appropriate, provision of preliminary instructions on the use of jury questions, and provision of more extensive written instruction regarding discussions.

Overall, permitting jurors to engage in predeliberation discussions did not have the extremely positive or negative effects that were predicted. There is evidence

that discussions increase the jurors' perceptions that they understand the evidence and are performing their duties efficiently. There is also some evidence that it may increase comprehension of the trial evidence. However, there is little evidence that it results in prejudgment of the case, reduces the amount of improper discussion, or improves overall decision-making quality. Additional procedures related to the jury discussions may improve the effectiveness of this aid to jury decision making.

ACCESS TO TRIAL TRANSCRIPTS

Providing jurors access to the trial transcripts could improve juror decision making, and it is generally within the broad discretion of the judge to determine if the jury will have access to transcripts (see, e.g., *Commonwealth v. Richotte*, 2003). One of the judicial concerns in reviewing testimony for the jury upon their specific request is that the jury will unduly rely on the reviewed testimony in reaching a verdict (*Commonwealth v. Bacigalupo*, 2000). However, permitting jurors to review the transcript of the entire case would allow them to refresh their recollection of the testimony and to clarify points of argument over the testimony.

It has been suggested that trial transcript access could increase the chances that jurors will process the trial evidence systematically (Horowitz & ForsterLee, 2001). Prior research indicates that access to the trial transcript prevents jurors from relying on biased heuristic cues in their decision making. Bourgeois, Horowitz, and ForsterLee (1993) manipulated access to trial transcripts in a medical malpractice case. Mock jurors with access to transcripts made legally correct decisions whereas mock jurors without the transcripts were more likely to rely on heuristic cues and to make legally incorrect decisions.

Horowitz and ForsterLee (2001) investigated the effects of trial transcript access on jury decision making. It was hypothesized that access to trial transcripts would result in systematic evidence processing, better discrimination among plaintiffs who should be differentially compensated, and improved recall of probative evidence. Mock juries in a toxic tort case were or were not allowed to access the trial transcript on a computer in the deliberation room. Four differentially worthy plaintiffs were presented to the jurors, and note taking was also manipulated. Access to transcripts did not significantly affect liability verdicts but did decrease the compensation awarded to all of the plaintiffs. However, transcript access did not result in better discrimination among the differentially worthy plaintiffs in terms of the compensatory awards. Juries with access to trial transcripts were not more likely to distinguish among the plaintiffs than the juries without access. It is unclear if the jurors who were given access to the transcripts actually reviewed the transcripts, and if they did, to what extent they relied on the transcripts; this could account for the transcripts' lack of influence in this area. As for positive outcomes, juries with transcript access made fewer recognition errors for the trial evidence. Additionally, jurors given access to transcripts reported being more satisfied with

the trial process and were more likely to believe that their jury worked efficiently than jurors without transcript access (Horowitz & ForsterLee, 2001).

Overall, the results of the existing research indicate that transcripts may help jurors process the evidence, as evidenced by fewer errors in recognizing trial evidence. However, the results are mixed as to whether decision making was improved with the access to transcripts.

WRITTEN WITNESS SUMMARIES

Written summaries of witnesses' testimony have the potential to help jurors process the information presented during testimony by providing them with an appropriate schema with which to interpret the evidence (ForsterLee, Horowitz, Athaide-Victor, & Brown, 2000). This may be especially so when the information is complex. However, there are several procedural problems with the use of written witness summaries, including court monitoring to ensure accuracy and prevent improper conclusions from being presented to the jury (ForsterLee et al., 2000). Although providing summary testimony is not a common courtroom procedure, some states have suggested the use of similar procedures, such as the court or the attorneys providing summaries of the evidence presented throughout a case (see, e.g., Michigan Supreme Court Order 2005-19).

Very little research has examined the effectiveness of evidence summaries. ForsterLee et al. (2000) examined the effectiveness of giving jurors written summaries of expert testimony. The mock trial presented to participants was a toxic tort case that included complex expert testimony from a physician. The written summary was three pages long and included the credentials of the expert and a short summary of his testimony, but it did not include his conclusions. This summary was presented to the jurors prior to the trial, after the trial, or not at all. To determine the effectiveness of the expert, four plaintiffs were involved in the lawsuit who had varying degrees of injury about which the expert provided information. Participants who were not given a written summary were unable to reliably distinguish between the injury levels of the four plaintiffs, whereas participants receiving the pretrial summary were able to distinguish among all four of the plaintiffs. The participants receiving the posttrial summary were able to distinguish between only the highest and the lowest level of injury. Other differences were observed between the pretrial and posttrial groups. Recall of relevant expert information was better for the pretrial group than both the posttrial and no-summary groups, and ratings of the technicality of the expert's testimony were lower for the pretrial group than for both the posttrial group and the no-summary group. Overall, jurors' memories, information processing, and decision making were improved by the presentation of the summary prior to the evidence. The authors predicted that the observed positive effects on cognitive processing would be even greater in real trials, which include jury deliberations.

In a follow-up study including deliberations, ForsterLee, Kent, and Horowitz (2005) investigated the effectiveness of written summaries in combination with juror note taking. Participants were exposed to the same case materials, and the presence of pre-evidence summaries and note taking were manipulated. Results after deliberations in this study were similar to those without deliberations. Access to expert witness summaries prior to hearing the expert testify increased the jury's ability to discriminate among differentially injured plaintiffs and improved recall of the evidence. Both of these positive effects were increased when note taking was permitted. Jurors were also more satisfied and believed the evidence was less complex when provided summaries. Witness summaries also increased jury damage awards overall, primarily for the most severely injured plaintiffs, indicating a potential negative effect of the summaries (ForsterLee et al., 2005). Despite this potentially negative effect, which requires further research, the overall impact of allowing expert witness summaries was positive for evidence processing and decision making.

CONCLUSIONS FROM EMPIRICAL RESEARCH

These procedures—juror questions, juror notetaking, predeliberation jury discussion, access to trial transcripts, and written witness statements—deserve consideration as ways to assist jurors with their often complicated task. Commentators, scholars, attorneys, and judges have long complained about jury performance. It is noteworthy that criticisms have been offered and jury reforms advanced without relying on relevant systematic data. This situation is beginning to change as studies such as those discussed in this chapter provide new insights into the strengths and weaknesses of jury decision making and allow us to identify procedural reforms and decision aids that will optimize jury performance.

REFERENCES

- Abraham, K. S., & Jeffries, J. C. (1989). Punitive damages and the rule of law: The role of defendant's wealth. *Journal of Legal Studies*, 18, 415–425.
- Abramson, J. (1994). *We, the jury: The jury system and the ideal of democracy*. New York, NY: Basic Books.
- Adamo, K. R. (1996). Reforming jury practice in patent cases: Suggestions towards learning to love using an eighteenth century system while approaching the twenty-first century. *Journal of the Patent and Trademark Office Society*, 78, 345–358.
- Adler, S. J. (1994). *The jury: Trial and error in the American courtroom*. New York, NY: Times Books.
- Alaska Statutes §09.17.020 (1994).
- Arizona Rules of Civil Procedure, Rule 39(f).
- Arlen, J. H. (1992). Should defendants' wealth matter? *Journal of Legal Studies*, 21, 413–429.
- Baker, T., Harel, A., & Kugler, T. (2003). *The virtues of uncertainty in law: An experimental approach*. Available from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=380302

- BMW of North America, Inc. v. Gore, 517 U.S. 559 (1996).
- Bornstein, B. (1998). From compassion to compensation: The effect of injury severity on mock jurors' liability judgments. *Journal of Applied Social Psychology, 28*, 1477–1502.
- Bourgeois, M., Horowitz, I., & ForsterLee, L. (1993). The effects of technicality and access to trial transcripts on verdicts and information processing in a civil trial. *Personality and Social Psychology Bulletin, 19*, 200–227.
- Broder, J. M. (September 10, 1997). Stares of lawyerly disbelief at a huge civil award. *New York Times*, p. C-1.
- Browning-Ferris Ind. v. Kelco Disposal, Inc., 492 U.S. 257 (1989).
- Cahill v. Baltimore, 98 A. 235 (Md. 1916).
- California Civil Code §1431.2.
- Cano, E. (2001). Speaking out: Is Texas inhibiting the search for truth by prohibiting juror questioning of witnesses in criminal cases? *Texas Tech Law Review, 32*, 1013–1051.
- Casper, J. D., Benedict, K., & Perry, J. L. (1989). Juror decision making, attitudes, and the hindsight bias. *Law and Human Behavior, 13*, 291–310.
- Cather, C., Greene, E., & Durham, R. (1996). Plaintiff injury and defendant reprehensibility: Implications for compensatory and punitive damage awards. *Law and Human Behavior, 20*, 189–205.
- Cecil, J., Lind, E., & Bermant, G. (1987). *Jury service in lengthy trials*. Washington, DC: Federal Judicial Center.
- Chapman, G. B., & Bornstein, B. H. (1996). The more you ask for, the more you get: Anchoring in personal injury verdicts. *Applied Cognitive Psychology, 10*, 519–540.
- Chicago Hansom Cab Co. v. Havelick, 22 N. E. 797 (Ill. 1869).
- Clermont, K. M., & Eisenberg, T. (1992). Trial by jury or judge: Transcending empiricism. *Cornell Law Review, 77*, 1124–1177.
- Cohee v. State, 942 P.2d 211 (Okla. 1997).
- Commonwealth v. Bacigalupo, 731 N. E.2d 559 (Mass. App. Ct., 2000).
- Commonwealth v. Richotte, 796 N. E.2d 890 (Mass. App. Ct., 2003).
- Commonwealth v. Urena, 632 N. E.2d 1200 (Mass. 1994).
- Connecticut General Statutes Annotated, 52–40 (West 1987).
- Conrad, A. F., & Bernstein, H. (1964). *Automobile accident costs and payments: Studies in the economics of injury reparations*. Ann Arbor: University of Michigan Press.
- Cooper Industries, Inc. v. Leatherman Tool Group, Inc., 532 U.S. 424 (2001).
- Crum v. State, 946 S.W.2d 349 (Tex. App. 1997).
- Daniels, S. (1989). The question of jury competence and the politics of civil justice reform: Symbols, rhetoric, and agenda-building. *Law & Contemporary Problems, 52*, 269–298.
- Daniels, S., & Martin, J. (1990). Myth and reality in punitive damages. *Minnesota Law Review, 75*, 1–64.
- Dann, B. M. (1993). “Learning lessons” and “speaking rights”: Creating educated and democratic juries. *Indiana Law Journal, 68*, 1229–1279.
- Dann, B. M. (1996). Free the jury. *Litigation, 23*, 5–6, 64–66.
- Davies, G. M. (2009). Estimating the speed of vehicles: The influence of stereotypes. *Psychology, Crime, and Law, 15*, 293–312.
- Day v. Kilgore, 444 S. E.2d 515 (S.C. 1994).
- DeBenedetto v. Goodyear, 754 F.2d 512 (4th Cir. 1985).
- Densen v. Stanley, 84 So. 770 (Ala. Ct. App. 1919).

- Devine, D. J. (2012). *Jury decision making: The state of the science*. New York, NY: New York University Press.
- Diamond, S. S. (2006). Beyond fantasy and nightmare: A portrait of the jury. *Buffalo Law Review*, 54, 717–763.
- Diamond, S. S., Rose, M. R., Murphy, B., & Smith, S. (2004). Juror questions during trial: A window into juror thinking. *Vanderbilt Law Review*, 59, 1927–1972.
- Diamond, S., Rose, M., Murphy, B., & Smith, S. (2006). Jurors' unanswered questions. *Court Review*, 41, 20–29.
- Diamond, S., Vidmar, N., Rose, M., Ellis, L., & Murphy, B. (2003). Juror discussions during civil trials: Studying an Arizona innovation. *Arizona Law Review*, 45, 1–79.
- Eisenberg, T., Goerdt, J., Ostrom, B., Rottman, D., & Wells, M. T. (1997). The predictability of punitive damages. *Journal of Legal Studies*, 26, 623–660.
- Eisenberg, T., Heise, M., & Wells, M. T. (2010). Variability in punitive damages: Empirically assessing Exxon Shipping Co. v. Baker. *Journal of Institutional and Theoretical Economics*, 116, 5–26.
- Eisenberg, T., LaFountain, N., Ostrom, B., & Rottman, D. (2002). Juries, judges, and punitive damages: An empirical study. *Cornell Law Review*, 87, 743–780.
- Eisenberg, T., Hannaford-Agor, P. L., Hans, V. P., Waters, N. L., Munsterman, G. T., Schwab, S. J., & Wells, M. T. (2005). Judge-jury agreement in criminal cases: A partial replication of Kalven and Zeisel's *The American Jury*. *Journal of Empirical Legal Studies*, 2, 171–206.
- Eisenberg, T., Hannaford-Agor, P. L., Heise, M., LaFountain, N., Munsterman, G. T., Ostrom, B., & Wells, M. T. (2006). Juries, judges, and punitive damages: Empirical analyses using the civil justice survey of state courts 1992, 1996, and 2001 data. *Journal of Empirical Legal Studies*, 3, 263–295.
- Eisenberg, T., Heise, M., Waters, N. L., & Wells, M. T. (2010). The decision to award punitive damages: An empirical study. *Journal of Legal Analysis*, 2, 577–620.
- Eisenberg, T. & Heise, M. (2011). Judge-jury difference in punitive damages awards: Who listens to the Supreme Court? *Journal of Empirical Legal Studies*, 8, 325–357.
- Ellsworth, P. C. (1993). Some steps between attitudes and verdicts. In R. Hastie (Ed.), *Inside the juror: The psychology of juror decision making* (pp. 42–64). Cambridge, England: Cambridge University Press.
- Elwork, A., Sales, B. D., & Alfini, J. J. (1982). *Making jury instructions understandable*. Charlottesville, VA: Michie.
- English, P., & Sales, B. (1997). A ceiling or consistency effect for the comprehension of jury instructions. *Psychology, Public Policy, and Law*, 3, 381–401.
- Exxon Shipping Co. v. Baker, 554 U.S. 471 (2008).
- Fischer v. Fischer, 142 N. W. 2d 857 (Wis. 1966).
- Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 288–299.
- Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.) *Judgment under uncertainty: Heuristics and biases* (pp. 422–444). New York, NY: Cambridge University Press.
- Flango, V. E. (1980). Would jurors do a better job if they could take notes? *Judicature*, 63, 436–443.

- Ford, M. C. (1986). The role of extralegal factors in jury verdicts. *Justice System Journal, 11*, 16–39.
- ForsterLee, L., & Horowitz, I. (1997). Enhancing juror competence in a complex trial. *Applied Cognitive Psychology, 11*, 305–319.
- ForsterLee, L., Horowitz, I., Athaide-Victor, E., & Brown, N. (2000). The bottom line: The effect of written expert witness statements on juror verdicts and information processing. *Law and Human Behavior, 24*, 259–270.
- ForsterLee, L., Horowitz, I., & Bourgeois, M. (1994). Effects of notetaking on verdicts and evidence processing in a civil trial. *Law and Human Behavior, 18*, 566–578.
- ForsterLee, L., Kent, L., & Horowitz, I. (2005). The cognitive effects of jury aids on decision-making in complex civil litigation. *Applied Cognitive Psychology, 19*, 867–884.
- Frank, J. (1949). *Courts on trial*. Princeton, NJ: Princeton University Press.
- Galanter, M., & Luban, D. (1993). Poetic justice: Punitive damages and legal pluralism. *American University Law Review, 42*, 1393–1463.
- Garber, S. (1998). Product liability, punitive damages, business decisions, and economic outcomes. *Wisconsin Law Review, 1998*, 237–295.
- Goodman, J., Greene, E., & Loftus, E. F. (1989). Runaway verdicts or reasoned determinations: Mock juror strategies in awarding damages. *Jurimetrics, 29*, 285–309.
- Green, E. (1968). The reasonable man: Legal fiction or psychosocial reality? *Law and Society Review, 2*, 241–257.
- Greene, E. (1989). On juries and damage awards: The process of decisionmaking. *Law and Contemporary Problems, 52*, 225–246.
- Greene, E., & Bornstein, B. H. (2003). *Determining damages: The psychology of jury awards*. Washington, DC: American Psychological Association.
- Greene, E., Chopra, S., Kovera, M. B., Penrod, S., Rose, V. G., Schuller, R., & Studebaker, C. (2002). Jurors and juries: A review of the field. In J. Ogloff (Ed.), *Taking psychology and law into the twenty-first century* (pp. 225–284). New York: Kluwer Academic/Plenum Press.
- Greene, E., Goodman, J., & Loftus, E. F. (1991). Jurors' attitudes about civil litigation and the size of damage awards. *American University Law Review, 40*, 805–820.
- Greene, E., Johns, M., & Bowman, J. (1999). The effects of injury severity on jury negligence decisions. *Law and Human Behavior, 23*, 675–693.
- Greene, E., Johns, M., & Smith, A. (2001). The effects of defendant conduct on jury damage awards. *Journal of Applied Psychology, 86*, 228–237.
- Greene, E., Woody, W. D., & Winter, R. (2000). Compensating plaintiffs and punishing defendants: Is bifurcation necessary? *Law and Human Behavior, 24*, 187–205.
- Guilbault, R. L., Bryant, F. B., Brockway, J. H., & Posavac, E. J. (2004). A meta-analysis of research on hindsight bias. *Basic and Applied Social Psychology, 26*, 103–117.
- Guthrie, C., Rachlinski, J. J., & Wistrich, A. J. (2001). Inside the judicial mind. *Cornell Law Review, 86*, 777–830.
- Guthrie, C., Rachlinski, J. J., & Wistrich, A. J. (2007). Blinking on the bench: How judges decide cases. *Cornell Law Review, 93*, 1–43.
- Hannaford, P., Hans, V., & Munsterman, T. (2000). Permitting jury discussions during trial: Impact of the Arizona reform. *Law and Human Behavior, 24*, 359–382.
- Hans, V. P., & Dee, J. (2003). Whiplash: Who's to Blame? *Brooklyn Law Review, 68*, 1093–1120.
- Hans, V. P., & Ermann, M. D. (1989). Responses to corporate versus individual wrongdoing. *Law and Human Behavior, 13*, 151–166.

- Hans, V. P., Hannaford, P., & Munsterman, T. (1999). The Arizona jury reform permitting civil jury trial discussions: The views of trial participants, judges, and jurors. *University of Michigan Journal of Law Reform*, 32, 349–377.
- Hans, V. P., Kaye, D. H., Dann, B. M., Farley, E. J., & Albertson, S. (2011). Science in the jury box: Jurors' comprehension of mitochondrial DNA evidence. *Law and Human Behavior*, 35, 60–71.
- Hans, V. P., & Lofquist, W. S. (1992). Jurors' judgments of business liability in tort cases: Implications for the litigation explosion debate. *Law and Society Review*, 26, 85–115.
- Hans, V. P., & Lofquist, W. S. (1994). Perceptions of civil justice: The litigation crisis attitudes of civil jurors. *Behavioral Science and the Law*, 12, 181–196.
- Hans, V. P. & Reyna, V. F. (2011). To dollars from sense: Qualitative to quantitative translation in jury damage awards. *Journal of Empirical Legal Studies*, 8, 120–147.
- Hart, A. J., Evans, D. L., Wissler, R. L., Feehan, J. W., & Saks, M.J. (1997). Injuries, prior beliefs, and damage awards. *Behavioral Science and the Law*, 15, 63–82.
- Hastie, R. (1983). *Final report to the National Institute for Law Enforcement and Criminal Justice*. Unpublished manuscript, Northwestern University.
- Hastie, R., Schkade, D., & Payne, J. (1998). A study of juror and jury judgments in civil cases: Deciding liability for punitive damages. *Law and Human Behavior*, 22, 287–314.
- Hastie, R., Schkade, D. A., & Payne, J. W. (1999a). Juror judgments in civil cases: Effects of plaintiff's requests and plaintiff's identity on punitive damage awards. *Law and Human Behavior*, 23, 445–470.
- Hastie, R., Schkade, D. A., & Payne, J. W. (1999b). Juror judgments in civil cases: Hindsight effects on judgments of liability for punitive damages. *Law and Human Behavior*, 23, 597–614.
- Hastie, R., & Viscusi, W. K. (1998). What juries can't do well: The jury's performance as a risk manager. *Arizona Law Review*, 40, 901–921.
- Hersch, J., & Viscusi, W. K. (2004). Punitive damages: How judges and juries perform. *Journal of Legal Studies*, 33, 1–36.
- Heuer, L., & Penrod, S. D. (1988). Increasing jurors' participation in trials: A field experiment with jury notetaking and question asking. *Law and Human Behavior*, 12, 409–430.
- Heuer, L., & Penrod, S. D. (1989). Instructing jurors: A field experiment with written and preliminary instructions. *Law and Human Behavior*, 13, 231–162.
- Heuer, L., & Penrod, S. D. (1994a). Juror notetaking and question asking during trial: A national field experiment. *Law and Human Behavior*, 18, 121–150.
- Heuer, L., & Penrod, S. D. (1994b). Trial complexity: A field investigation of its meaning and its effects. *Law and Human Behavior*, 18, 29–52.
- Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, 107, 311–327.
- Hinsz, V. B., & Indahl, K. E. (1995). Assimilation to anchors for damage awards in a mock civil trial. *Journal of Applied Social Psychology*, 25, 991–1026.
- Honda Motor Co. v. Oberg, 512 U.S. 415 (1994).
- Horowitz, I., & Bordens, K. S. (1990). An experimental investigation of procedural issues in complex tort trials. *Law and Human Behavior*, 14, 269–285.
- Horowitz, I., & Bordens, K. S. (2002). The effects of jury size, evidence complexity, and note taking on jury process and performance in a civil trial. *Journal of Applied Psychology*, 87, 121–130.

- Horowitz, I. A., Bordens, K. S., Victor, E., Bourgeois, M. J., & ForsterLee, L. (2001). The effects of complexity on jurors' verdicts and construction of evidence. *Journal of Applied Psychology, 86*, 641–652.
- Horowitz, I., & ForsterLee, L. (2001). The effects of notetaking and trial transcript access on mock jury decisions in a complex civil trial. *Law and Human Behavior, 25*, 373–391.
- Hudson v. Markum, 948 S. W.2d 1 (Tex. App. 1997).
- Kahneman, D., Sunstein, C., & Schkade, D. (1998). Shared outrage and erratic awards: The psychology of punitive damages. *Journal of Risk & Uncertainty, 16*, 49–86.
- Kalven, H. (1964). The dignity of the civil jury. *Virginia Law Review, 50*, 1055–1075.
- Kalven, H., & Zeisel, H. (1966). *The American jury*. Boston: Little, Brown.
- Kamin, K. A., & Rachlinski, J. J. (1995). Ex post ≠ ex ante: Determining liability in hindsight. *Law and Human Behavior, 19*, 89–104.
- Kansas Statutes Annotated §60–3701 (1994).
- Karlovac, M., & Darley, J. M. (1988). Attribution of responsibility for accidents: A negligence law analogy. *Social Cognition, 6*, 287–318.
- Kelso, J. C. (1996). Final report of the Blue Ribbon Commission on Jury System Improvement. *Hastings Law Journal, 47*, 1433–1592.
- King, E. M., & Smith, J. P. (1988). *Economic loss and compensation in aviation accidents*. Santa Monica, CA: RAND Corporation.
- Kovera, M. B., & McAuliff, B. D. (2000). The effects of peer review and evidence quality on judge evaluations of psychological science: Are judges effective gatekeepers? *Journal of Applied Psychology, 85*, 574–586.
- Kovera, M. B., McAuliff, B. D., & Herbert, K. S. (1999). Reasoning about scientific evidence: Effects of juror gender and evidence quality on juror decisions in a hostile work environment case. *Journal of Applied Psychology, 84*, 362–375.
- Kozinski, A. (1995, Jan. 19). The case of punitive damages v. democracy. *Wall Street Journal*, p. A-19.
- Krause v. State, 132 P.2d 179 (Okla. Crim. App. 1942).
- LaBine, S. J. & LaBine, G. (1996). Determinations of negligence and the hindsight bias. *Law and Human Behavior, 20*, 501–516.
- Landes, W. M., & Posner, R. A. (1986, September/October). New light on punitive damages. *Regulation, 10*, 33–36, 54.
- Landsman, S., & Rakos, R. F. (1994). A preliminary inquiry into the effects of potentially biasing information on judges and jurors in civil litigation. *Behavioral Science and the Law, 12*, 113–126.
- Larsen, S. (1996). Taking and use of trial notes by jury. *American Law Reports, 36*, 1–254.
- Lawson v. State, 664 N. E.2d 773 (Ind. Ct. App. 1996).
- Lempert, R. O. (1993). Civil juries and complex cases: Taking stock after twelve years. In R. E. Litan (Ed.), *Verdict: Assessing the civil jury system* (pp. 181–247). Washington, DC: Brookings Institution.
- Lieberman, J., & Sales, B. (1997). What social psychology teaches us about the jury instruction process. *Psychology, Public Policy, and Law, 3*, 589–644.
- Litigation Sciences. (1993). Psychological characteristics of punitive damage jurors. In *Jury Research*. PLI Corp. Law and Prac. Course Handbook Series No. 833 (p. 59). New York, NY: Practising Law Institute.
- Louisville Bridge & Terminal Co. v. Brown, 277 S.W. 320 (Ky. 1925).

- Lowe, D. J., & Reckers, P. M. J. (1994). The effects of hindsight bias on jurors' evaluations of auditor decisions. *Decision Science*, 25, 401–426.
- MacCoun, R. (1993a). Inside the black box: What empirical research tells us about decisionmaking by civil juries. In R. E. Litan (Ed.), *Verdict: Assessing the civil jury system* (pp. 137–180). Washington, DC: Brookings Institution.
- MacCoun, R. (1993b). *Is there a "deep-pocket" bias in the tort system? The concern over biases against deep-pocket defendants*. Issue Paper, Santa Monica, CA: RAND Institute for Civil Justice.
- MacCoun, R. (1996). Differential treatment of corporate defendants by juries: An examination of the "deep-pockets" hypothesis. *Law and Society Review*, 30, 121–161.
- Maggart v. Bell, 2 P.2d 516 (Cal. App. 1931).
- Marbley v. State, 461 N. E.2d 110 (Ind. 1984).
- Matthews v. Commonwealth Edison Co., WL 478820 (N.D. Ill. 1995).
- McAuliff, B. D., Kovera, M. B., & Nunez, G. (2009). Can jurors recognize missing control groups, confounds, and experimenter bias in psychological science? *Law and Human Behavior*, 33, 247–257.
- McLaughlin, M. A. (1982). Questions to witnesses and notetaking by the jury as aids in understanding complex litigation. *New England Law Review*, 18, 687–713.
- Michigan Supreme Court Order 2005-19.
- Missouri Revised Statutes §538.210 (June 29, 2011). Amendment of Rules 2.512, 2.513, 2.514, 2.515, 2.516, and 6.414 of the Michigan Court Rules.
- Moran, G., Cutler, B. L., & DeLisa, A. (1994). Attitudes toward tort reform, scientific jury selection, and juror bias: Verdict inclination in criminal and civil trials. *Law and Psychology Review*, 18, 309–328.
- Morrison v. State, 845 S.W.2d 882 (Tex. 1992).
- Mott, N. (2003). "To ask or not to ask, that is the question." *Chicago-Kent Law Review*, 78, 1099–1122.
- Mott, N. L., Hans, V. P., & Simpson, L. (2000). What's half a lung worth? Civil jurors' accounts of their award decision making. *Law and Human Behavior*, 24, 401–419.
- Munsterman, G. T., Hannaford-Agor, P. L., & Whitehead, G. M. (2006). *Jury trial innovations* (2nd ed.). Williamsburg, VA: National Center for State Courts.
- North Dakota Century Code §32–03.2–11.
- O'Connor, S. D. (1997). Juries: They may be broken, but we can fix them. *Federal Lawyer*, 44, 20–25.
- Owen, D. G. (1994). A punitive damages overview: Functions, problems and reform. *Villanova Law Review*, 39, 363–413.
- Pacific Mutual Life Ins. Co. v. Haslip, 498 U.S. 1306 (1991).
- Pennington, N., & Hastie, R. (1993). The story model for juror decision making. In R. Hastie (Ed.), *Inside the juror: The psychology of juror decision making* (pp. 192–221). New York, NY: Cambridge University Press.
- Penrod, S., Linz, D., & Rios, P. (1983, June). *Allowing jurors to question witnesses: A courtroom experiment*. Paper presented at the Law and Society Association annual meeting, Denver, CO.
- Peterson, M. A. (1984). *Compensation of injuries: Civil jury verdicts in Cook County*. Santa Monica, CA: RAND Institute for Civil Justice.

- Peterson, M. A., Sarma, S., & Shanley, M. (1987). *Punitive damages: Empirical findings*. Santa Monica, CA: RAND Institute for Civil Justice.
- Philip Morris USA v. Williams, 549 U.S. 346 (2007).
- Posner, R. A. (1995, March 1). Juries on trial. *Commentary*, 99, 49.
- Prosser, W. L., & Keeton, P. (1984). *Prosser and Keeton on the law of torts* (5th ed.). St. Paul, MN: West.
- Purver, J. M. (1970). Propriety of jurors asking questions in open court during course of trial. *American Law Reports*, 31(Series 3), 872–892.
- Quayle, D. (1992). Civil justice reform. *American University Law Review*, 41, 559–569.
- Ratton v. Busby, 326 S.W.2d 889 (Ark. 1959).
- Reece v. Simpson, 437 So. 2d 68 (Ala. 1983).
- Restatement (Second) of Torts, §908(2).
- Robbennolt, J. K. (2000). Outcome severity and judgments of “responsibility”: A meta-analytic review. *Journal of Applied Social Psychology*, 12, 2575–2609.
- Robbennolt, J. K. (2002a). Determining punitive damages: Empirical insights and implications for reform. *Buffalo Law Review*, 50, 103–203.
- Robbennolt, J. K. (2002b). Punitive damage decision making: The decisions of citizens and trial court judges. *Law and Human Behavior*, 26, 315–341.
- Robbennolt, J. K., Darley, J. M., & MacCoun, R. J. (2003). Symbolism and incommensurability in civil sanctioning: Legal decision-makers as goal managers. *Brooklyn Law Review*, 68, 1121–1158.
- Robbennolt, J. K., & Studebaker, C. A. (1999). Anchoring in the courtroom: The effects of caps on punitive damages. *Law and Human Behavior*, 23, 353–373.
- Robbennolt, J. K., & Studebaker, C. A. (2003). News media reporting on civil litigation and its influence on civil justice decision making. *Law and Human Behavior*, 27, 5–27.
- Rosenhan, D. L., Eisner, S. L., & Robinson, R. J. (1994). Notetaking can aid juror recall. *Law and Human Behavior*, 18, 53–61.
- Rustad, M. (1991). *Demystifying punitive damages in products liability cases: A survey of a quarter century of trial verdicts*. Washington, DC: Roscoe Pound Foundation.
- Rustad, M. (1992). In defense of punitive damages in products liability: Testing tort anecdotes with empirical data. *Iowa Law Review*, 78, 1–88.
- Sachdev, A. (2003, March 22). \$10 billion smoking judgment: Philip Morris deception found over light cigarettes. *Chicago Tribune*, p. 1.
- Saks, M. J. (1992). Do we really know anything about the behavior of the tort litigation system: And why not? *University of Pennsylvania Law Review*, 140, 1147–1289.
- Saks, M. J., Hollinger, L. A., Wissler, R. L., Evans, D. L., & Hart, A. J. (1997). Reducing variability in civil jury awards. *Law and Human Behavior*, 21, 243–256.
- Sand, L. B., & Reiss, S. A. (1985). A report on seven experiments conducted by district court judges in the Second Circuit. *New York University Law Review*, 60, 423–497.
- Sanders, J. (1998). *Bendectin on trial: A study of mass tort litigation*. Ann Arbor: University of Michigan Press.
- Schaefer v. St. Louis & Suburban R. Co., 30 S.W. 331 (Mo. 1895).
- Scully, J. (1996). Markman and Hilton Davis: The Federal Circuit strikes an awkward balance: The roles of the judge and jury in patent infringement suits. *Hastings Communications and Entertainment Law Journal*, 18, 631–655.

- Simpson, J. T. (1996). Discovery of net worth in bifurcated punitive damages cases: A suggested approach after *Transportation Insurance Co. v. Moriel*. *Southern Texas Law Review*, 37, 193–229.
- Sligar v. Bartlett, 916 P.2d 1383 (Okla. 1996).
- Smith, A. C., & Greene, E. (2005). Conduct and its consequences: Attempts at debiasing jury judgments. *Law and Human Behavior*, 29, 505–526.
- Smith, V. L., & Studebaker, C. A. (1996). What do you expect?: The influence of people's prior knowledge of crime categories on fact-finding. *Law and Human Behavior*, 5, 517–532.
- South Carolina Code Annotated §15–33–135 (Law Co-op. 1993).
- State v. Alexander, 1997 WL 116903 (Ohio App. 1997).
- State v. Graves, 907 P.2d 963 (Mont. 1995).
- State v. Greer, 948 P.2d 945 (Ariz. Ct. App. 1997).
- State v. Howard, 360 S.E.2d 790 (N.C. 1987).
- State v. Kendall, 57 S.E. 340 (N.C. 1907).
- State v. Sickles, 286 S.W. 432 (Mo. App. 1926).
- State v. Triplett, 421 S.E.2d 511 (W. Va. 1992).
- State v. Trujillo, 869 S.W.2d 844 (Mo. App. 1994).
- State Farm v. Campbell, 538 U.S. 408 (2003).
- Taragin, M. I., Willett, L. R., Wilczek, A. P., Trout, R., & Carson, J. L. (1992). The influence of standard of care and severity of injury on the resolution of medical malpractice claims. *Annals of Internal Medicine*, 117, 780–784.
- Texas Civil Practice & Remedies Code §41.003.
- Texas Civil Practice & Remedies Code §84.003.
- Thompson, W. C., & Schumann, E. L. (1987). Interpretation of statistical evidence in criminal trials: The prosecutor's fallacy and the defense attorney's fallacy. *Law and Human Behavior*, 11, 167–187.
- Thornton v. Weaber, 112 A.2d 344 (Pa. 1955).
- Tift v. Towns, 63 Ga. 237 (1879).
- Tripoli, L. (1997). Precipice of change: Professional groups urge striking changes in trial. *Inside Litigation*, 3, 1–6.
- Tversky, A., & Kahneman, D. (1982). Judgment under uncertainty: Heuristics and biases. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 3–20). Cambridge, England: Cambridge University Press.
- Twain, Mark (Samuel L. Clemens). *Roughing it*. Hartford, CT: American Publishing Co., 1872.
- TXO Production Corp. v. Alliance Resources Corp., 509 U.S. 443 (1993).
- United States General Accounting Office. (1989). *Product liability: Verdicts and case resolution in five states*. Washington, DC: Author.
- United States v. Ajmal, 67 F.3d 12 (2d Cir. 1995).
- United States v. Bush, 47 F.3d 511 (2d Cir. 1995).
- United States v. Callahan, 588 F.2d 1078 (5th Cir. 1979).
- United States v. Carlisi, 32 F. Supp. 479 (E.D. N.Y. 1940).
- United States v. Darden, 70 F.3d 1507 (8th Cir. 1995).
- United States v. Davis, 103 F. 457 (W.D. Tenn), aff 'd, 107 F. 753 (6th Cir. 1900).
- United States v. Douglas, 81 F.3d 324 (2d Cir. 1996).

- United States v. Feinberg, 89 F.3d 333 (7th Cir. 1996).
- United States v. Johnson, 892 F.2d 707 (8th Cir. 1989).
- United States v. Polowichak, 783 F.2d 410 (4th Cir. 1986).
- United States v. Richardson, 233 F.3d 1285 (11th Cir. 2000).
- United States v. Stierwalt, 16 F.3d 282 (8th Cir. 1994).
- United States v. Thompson, 76 F.3d 442 (2d Cir. 1996).
- United States v. Wexler, 657 F.Supp. 966 (E.D. Pa. 1987).
- Van der Keilen, M., & Garg, R. (1994). Moral realism in adults' judgments of responsibility. *Journal of Psychology, 128*, 149–156.
- Vidmar, N. (1995). *Medical malpractice and the American jury: Confronting the myths about jury incompetence, deep pockets, and outrageous damage awards*. Ann Arbor: University of Michigan Press.
- Vidmar, N. (1998). The performance of the American civil jury: An empirical perspective. *Arizona Law Review, 40*, 849–899.
- Vidmar, N., & Diamond, S. S. (2001). Juries and expert evidence. *Brooklyn Law Review, 66*, 1121–1180.
- Viscusi, W. K. (1998). Why there is no defense of punitive damages. *Georgetown Law Journal, 87*, 381–395.
- Viscusi, W. K. (2001a). Jurors, judges, and the mistreatment of risk by the courts. *Journal of Legal Studies, 30*, 107–136.
- Viscusi, W. K. (2001b). The challenge of punitive damages mathematics. *Journal of Legal Studies, 30*, 313–350.
- Wells, G. L. (1992). Naked statistical evidence of liability: Is subjective probability enough? *Journal of Personality and Social Psychology, 62*, 739–752.
- White v. Little, 268 P. 221 (Okla. 1928).
- Wiener, R. L., Gaborit, M., Pritchard, C. C., McDonough, E. M., Staebler, C. R., Wiley, D. C., & Goldkamp, K. S. (1994). Counterfactual thinking in mock-juror assessments of negligence: A preliminary investigation. *Behavioral Sciences and the Law, 12*, 89–102.
- Williams v. Commonwealth, 484 S.E.2d 153 (Va. Ct. App. 1997).
- Wingfield, N. (August 24, 2012). Jury awards \$1 billion to Apple in Samsung patent case. *New York Times*. Retrieved from <http://www.nytimes.com/2012/08/25/technology/jury-reaches-decision-in-apple-samsung-patent-trial.html>
- Wissler, R. L., Evans, D. L., Hart, A. J., Morry, M. M., & Saks, M. J. (1997). Explaining “pain and suffering” awards: The role of injury characteristics and fault attributions. *Law and Human Behavior, 21*, 181–207.
- Wissler, R. L., Hart, A. J., & Saks, M. J. (1999). Decisionmaking about general damages: A comparison of jurors, judges, and lawyers. *Michigan Law Review, 98*, 751–826.
- Wolff, M. A. (1990). Juror questions a survey of theory and use. *Missouri Law Review, 55*, 817–873.

Evaluating Eyewitness Testimony of Adults

STEPHEN J. ROSS, COLIN G. TREDoux, AND ROY S. MALPASS

IN the middle of the morning of January 11, 1993, a 36-year-old housekeeper in La Center, Washington, was cleaning her clients' property down a secluded, rural driveway. Two men (one dark-haired, the other blond) barged in and assaulted her. Although she put up a valiant defense, she was unable to ward them off. The blond man grabbed her wrists and held her hands up while the smaller, dark-haired man pulled her to the floor. The dark-haired assailant restrained her, while the blond man found an electric cord in the garage and used it to tie her arms and one ankle to the legs of the dining table. He also placed electrical tape over the victim's eyes. The blond assailant held her arms while the dark-haired assailant cut off the victim's shirt and bra and cut a hole in her pants. As he rubbed the box-cutter blade across the victim's chest, stomach, and thighs, the dark-haired assailant sexually assaulted the victim. After a few minutes, a car horn honked and the men argued briefly before they fled. As quickly as it began, they were gone, leaving the victim blindfolded and bound to the kitchen table. After several minutes of struggling, the victim was able to free herself and dial 911.

In the 911 call, she described the assailants as young White men, in their 20s. She was able to provide only limited details about the blond man—describing him as a “big guy, like a football kinda guy” who was tall and had long blond hair. She described the dark-haired assailant as having shorter dark hair, medium build, but shorter and smaller than the blond male.

Although the investigators collected physical evidence from the scene of the crime and the sexual assault kit, the case ultimately hinged on the victim's memory of her assailants. Within days of the crime, the victim sat with an artist to produce a composite sketch. Although she felt that her memory of the blond attacker was too weak to help create a sketch, she had enough exposure to the dark-haired man to produce a composite. During the course of the composite production, the victim

had difficulty providing a good description of the man's nose and mouth, so the sketch artist included "a nose that fits with [the] face" (Innocence Project Northwest, 2010, p. 9). The composite was widely distributed, and tips soon began coming in that a local logger named Alan Northrop resembled the sketch. A few days after the assault, police constructed a photo montage including photographs of Northrop. Later that day, the montage was shown to the victim, who did not identify anyone as the perpetrator. (See Figure 17.1.)

The lack of a positive identification stalled the investigation. Weeks passed and investigators were not receiving any solid leads on who the blond assailant might be; however, tips kept coming in that Alan Northrop resembled the composite sketch. With this as their only lead, investigators began looking to Northrop's acquaintances for individuals who might fit the description of the blond assailant. On February 2, 3 weeks after the crime, investigators questioned Steven Shade, whom they thought fit the description of the blond man. Shade denied any knowledge of the crime but did agree that the composite resembled Northrop, and pointed police to another of Northrop's friends, Larry Davis, as fitting the description of the blond assailant. The investigators created a second photo montage including Davis's photograph. (See Figure 17.2.) According to the police report, when viewing this montage later the same day, the victim positively identified Davis as her assailant saying, "That's not him, but that's the one" while pointing to the photograph of Larry Davis. However, in subsequent testimony, the victim stated that she did not say that—rather, she testified that she said the individual may have been the one of the two who assaulted her, but she could not identify his face; however, his neck looked familiar and that she "fe[lt] like [she] kn[e]w this guy" (Innocence Project Northwest, 2010, p. 11). Regardless of what she actually said, investigators viewed her statements

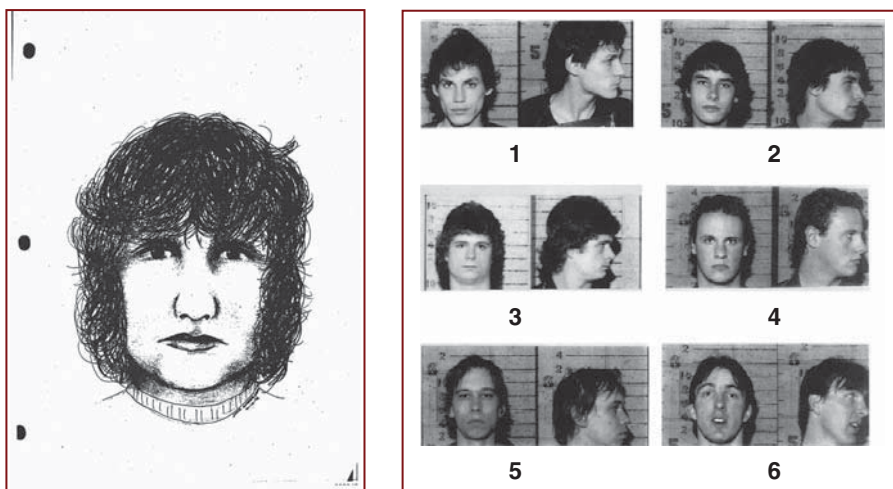


Figure 17.1 Composite Sketch Constructed of Perpetrator (left) and Northrop Photo Montage (right; Alan Northrop is in position #3)

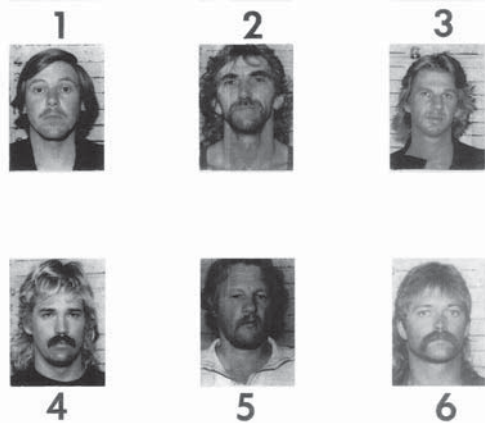


Figure 17.2 Photo Montage for Larry Davis (Davis is in position #3)

as a positive identification, and Larry Davis was arrested and booked into jail on suspicion of burglary.

After Davis's arrest, the victim received a phone call from a friend who worked at the jail to inform her that Larry Davis had been arrested in connection with her crime and that a second individual, Alan Northrop, was suspected of being involved but had not yet been arrested. The victim called the jail and was informed that an individual by the name of Larry Davis was in jail on suspicion of burglary. Assuming that the dark-haired man—the one who sexually assaulted her—would have been charged with a more serious offense than burglary, the victim concluded that Davis must have been the blond man and that Northrop must be suspected as being the dark-haired assailant. Shortly after the confirmation of Davis's arrest, investigators contacted the victim and informed her that they had arrested an individual and that she should come down to the station to view a live lineup. Later that day she viewed the lineup and identified Larry Davis as one of her assailants.

A short time later, investigators informed the victim that they had a suspect whom they thought might be the dark-haired assailant and that they had not yet made an arrest because she did not identify him in the photo montage she saw earlier. They told her that they were trying to put a live lineup together and, if they could create it, they would have her come to the station. Later that day, the victim was called back to the station where she viewed a live lineup and immediately identified Alan Northrop as the second assailant. (See Figure 17.3.)

Based on these identifications, both Alan Northrop and Larry Davis were arrested in connection with the crime. Northrop was charged with burglary, kidnapping, and sexual assault. Davis was charged with burglary, kidnapping, and accomplice to sexual assault. In separate trials conducted in the spring and early summer of 1993, they were convicted on all counts. Davis received a sentence of 20.5 years and Northrop was sentenced to 23.5 years.



Figure 17.3 Live Lineup for Alan Northrop (Northrop is in position #4)

In 2000, Davis and Northrop contacted the Innocence Project Northwest in an effort to gain access to DNA testing of the crime scene evidence. After Washington legislators changed the state law, leaving the decision to allow postconviction DNA testing to judges rather than to prosecutors, DNA testing was granted in 2006. Of the 27 pieces of evidence tested, none produced DNA matches to Northrop or Davis; however, the DNA profiles of two unidentified males were present. Further DNA testing excluded the victim's boyfriend. Based on these results, the Clark County Superior Court vacated Northrop and Davis's convictions in April 2010, concluding that the newly discovered DNA evidence would likely lead the jury to determine that neither Davis nor Northrop committed the crime. A few months later in July 2010, prosecutors officially dismissed all charges against Davis and Northrop.

EXTENT OF THE PROBLEM

Egregious errors of eyewitness identification, of the kind illustrated by the Northrop/Davis case, have plagued our criminal justice system for many years. The extent of the problem has come into public attention most clearly in the last 20 years, though, based largely on the development of DNA testing of identity during the 1980s and the use of this technology in criminal investigations. DNA testing gave us, for the first time, the ability to assess the validity of prior convictions from cases in which biological material had been preserved. Since 1989—the first year DNA evidence was admitted in U.S. courts—postconviction DNA testing has led to the exoneration of 301 individuals in the United States alone (Innocence Project, n.d.). Analysis of these exonerations reveals that erroneous eyewitness testimony

was present in 71% of these cases, making eyewitness misidentification the single greatest contributor to these miscarriages of justice (Garrett, 2012).

Although DNA testing has most clearly shown the scale of the problem, knowledge that eyewitness errors are on occasion implicated in the conviction of the innocent is not news. For over 100 years, scientists have written about the problems with eyewitness memory and the potential for its influence on errors in the criminal justice system. Münsterberg (1908) was among the first to describe the problems with eyewitness memory, and research in the intervening century has continued to reveal that erroneous eyewitness testimony plays a significant role in the conviction of innocent individuals (for a review of this research, see Malpass, Ross, Meissner, & Marcon, 2009). More recently, Gross and Shaffer (2012) affirmed the significant role of inaccurate eyewitness testimony in 873 cases of wrongful conviction in the United States since 1989.

Given the scale of the problem, legal and law enforcement communities should consider drawing on the substantial body of research on eyewitnesses conducted over many decades. This work can assist directly with the way that we think about the benefits and limitations of eyewitness evidence, and it can also provide practical benefits in developing “best practices” for collecting eyewitness evidence during the course of the investigation and for evaluating the quality of eyewitness testimony and identification after it has been obtained. We provide an overview of the tools available and the research findings regarding factors that are relevant to assessing eyewitness testimony, and we suggest a conceptual framework for organizing and interpreting this information. Later, we complete our discussion by demonstrating a practical application to the eyewitness evidence provided in the Northrop/Davis case.

INFORMATIONAL AND DECISIONAL CONSIDERATIONS IN RELATION TO EYEWITNESS MEMORY

Later in this chapter (in the Evaluating Eyewitness Testimony section), we discuss the factors that influence the reliability of eyewitness testimony by focusing on the underlying cognitive processes that influence eyewitness memory and response decision making. Although there are many theories and methodologies for assessing eyewitness testimony, we rely on the package of theory and procedure derived from signal detection theory (SDT). SDT was first used in legal psychology in the study of the “own and other race” effect (Malpass & Kravitz, 1969), and the decisional aspects were later elaborated in work on lineup instructional bias (Malpass & Devine, 1984). Recently, more complete applications of SDT were used by Meissner, Tredoux, Parker, and MacLin (2005) and Gronlund et al. (2012) to study contrasts among various kinds of simultaneous lineups, sequential lineups, and showups. Given our reliance on SDT as a framework for understanding how the factors we discuss later in the chapter influence eyewitness testimony, in this section we

present a conceptual structure derived from SDT for a forensic analysis of eyewitness testimony.

The distinction between the *information* available to a witness and the level of risk or uncertainty at which the witness is prepared to make a *decision* has proved to be highly beneficial for thinking about human perception and memory. The distinction comes from work by statisticians and psychologists during World War II who had the difficult task of deciding whether a sonar/radar signal represented an enemy craft (e.g., submarine, aircraft, etc.), some other object (e.g., a school of fish, a swarm of birds, etc.), or was due to some source of artifact (e.g., noise from electronic apparatus). Sonar/radar operators had information at their disposal (the signal on the sonar/radar screen) but could not be absolutely certain what it represented; they nevertheless had to make a decision.

The information at their disposal and the level of subjective certainty at which they would be prepared to make a decision were conceptualized as two separate dimensions, with operators likely differing on either or both of these dimensions. An operator might be fairly confident that the blip on the radar screen is that of an enemy bomber but cannot be certain. If it is identified as a bomber, there will be serious consequences, as ground and air resources will likely be marshaled to shoot it out of the sky. If it is not identified as a bomber but turns out to be one, the consequences are also potentially serious, leading to loss of life. The operator may weigh the consequences when making the decision, and the relative balance of damage to civilians or soldiers may be decisive in deciding what to identify the blip on the radar as being. The decision clearly rests on two types of thing: information available to the operator and the subjective threshold for making the decision. These are informational and decisional factors, respectively (see Gigerenzer et al., 1989, for a brief review of SDT and its connection to probability theory).

An eyewitness is in an analogous situation. There is a memory for the event in question, and this memory can vary in accuracy. The event may have been well encoded, or it may not have been. The witness may be able to retrieve the information effortlessly or may struggle to do so. Independent of this, the witness will be prepared to make a decision—an identification or reporting of information about the crime event—only at a level of certainty that is subjectively satisfying. That is, one witness may think that a crime is very important to solve and identify someone from the lineup, even if the match to her memory is not very good, whereas another may not be prepared to make an identification, even though the match to his memory is very good, reasoning that the small degree of doubt is too much to warrant jeopardizing a potentially innocent person. Just as in the case of the radar operator, the eyewitness has (a) information about an event (the memory of the perpetrator, and what happened), and (b) a subjective threshold for making an identification decision.

An important additional point in our modeling of eyewitness testimony through SDT (for more complete accounts of SDT, see MacMillan & Creelman, 2005; Wickens, 2002) is set out in Table 17.1. In this table, the columns refer to the true

Table 17.1
Outcomes of a Witness's Decision, Cross-Tabulated With Possible States
of Fact

	True State of the World	
Witness decision	Suspect <i>is</i> the perpetrator	Suspect is <i>not</i> the perpetrator
Witness identifies suspect	Correct identification (HIT)	Incorrect identification (FALSE ALARM)
Witness does not identify suspect	Incorrect rejection (MISS)	Correct rejection

status of things—in the case of a police lineup, whether the suspect is in fact the perpetrator or not the perpetrator—and the rows refer to the decision that is made—in the case of the police lineup, whether the witness does or does not identify the suspect.

Two of the decisions in the table are errors, and two are correct. The correct decisions are (1) correctly identifying a guilty suspect as the perpetrator and (2) correctly rejecting a lineup containing an innocent suspect. One type of error is when the eyewitness identifies an innocent suspect as a perpetrator. In this case, the liberty of an innocent person is put in jeopardy. A second kind of error is when the witness fails to identify the suspect as the perpetrator when this is in fact true. In this case, a guilty and perhaps dangerous criminal may escape apprehension and may continue to perpetrate crimes. The situation depicted in the table has been neatly wrapped up in SDT. It is useful to briefly discuss the statistical rudiments of SDT, as it demonstrates the power of separating out the informational and decisional aspects of decision making.

We start in the context of a lineup or photospread by assuming that the strengths of the memory for all the parties—perpetrator, innocent suspect, and fillers (i.e., the nonsuspects included in a lineup)—are distributed over a population of possible eyewitness events in the shape of a normal distribution, as represented by the curves in Figure 17.4. Of course, any given lineup will contain either the guilty suspect (perpetrator) or an innocent suspect. We include both cases by superimposing them.

If the witness's memory strength for the suspect (innocent or guilty) is not much greater than for the most familiar filler, then the two distributions will overlap to a considerable degree. But if the memory strength for the suspect is high in relation to the most familiar filler, then the two distributions will show less overlap. Figure 17.4 also shows the point on the strength of memory scale at which the witness is willing to identify the suspect as the perpetrator. It is called the decision criterion and is represented by a vertical line. Any decision criterion has consequences for the decision's accuracy. If the criterion is set to be moderate (i.e., below the middle on the memory strength scale [lower panel of Figure 17.5]), then we would have in this example 56% hits, 46% false alarms, and 36% filler identifications. If the criterion

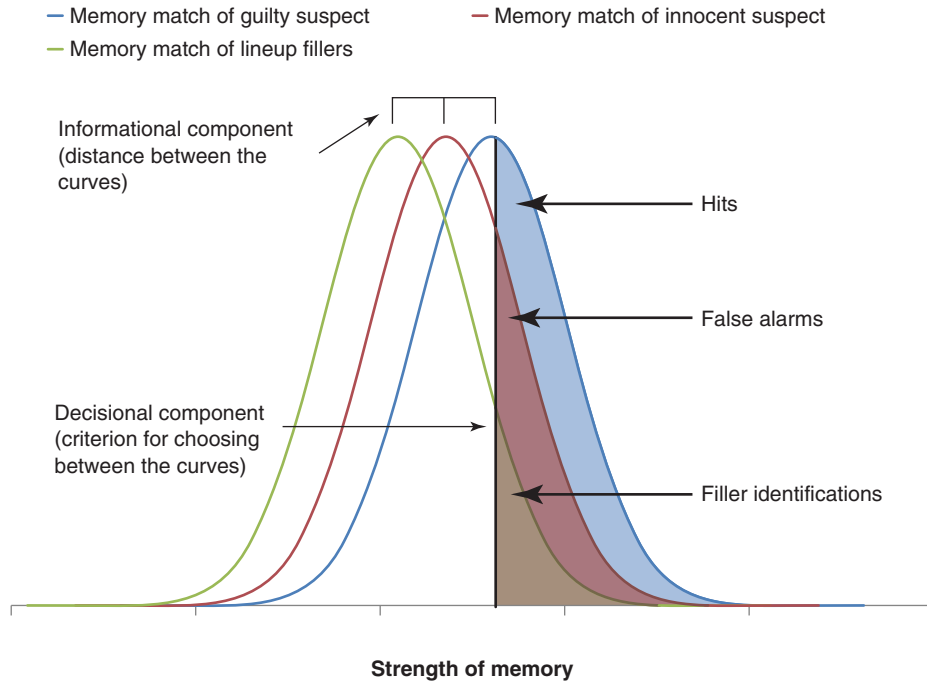


Figure 17.4 Informational and Decisional Components in a Joint Distribution of Perpetrator, Innocent Suspect, and Filler Match-to-Memory Strength

is set higher, on the strength of memory scales (upper panel), then we would have 46% hits, 36% false alarms, and 27% filler identifications. As shown in Figure 17.5, making the witness's identification judgment more conservative by moving the criterion upward on the memory strength scale results in fewer identifications overall (compare areas under the curves), fewer false identifications ($36\% < 46\%$), and fewer hits ($46\% < 56\%$).

This basic concept is of vital importance to medical and decisional sciences. We will use it as the framework through which we discuss the factors that should be considered when assessing eyewitness testimony. The evaluator of eyewitness testimony will find it beneficial to view the factors that influence witness testimony in relation to whether they are affecting the informational or decisional component of memory reporting. Factors related to the *informational* aspects are those that influence the form and strength of memory the witness has, the faces offered for possible identification, and the witness's evaluation of how similar any of the images presented in the identification procedure is to the information in memory. Many of these factors will be present during the criminal event, as that is the point at which the memory is being formed.

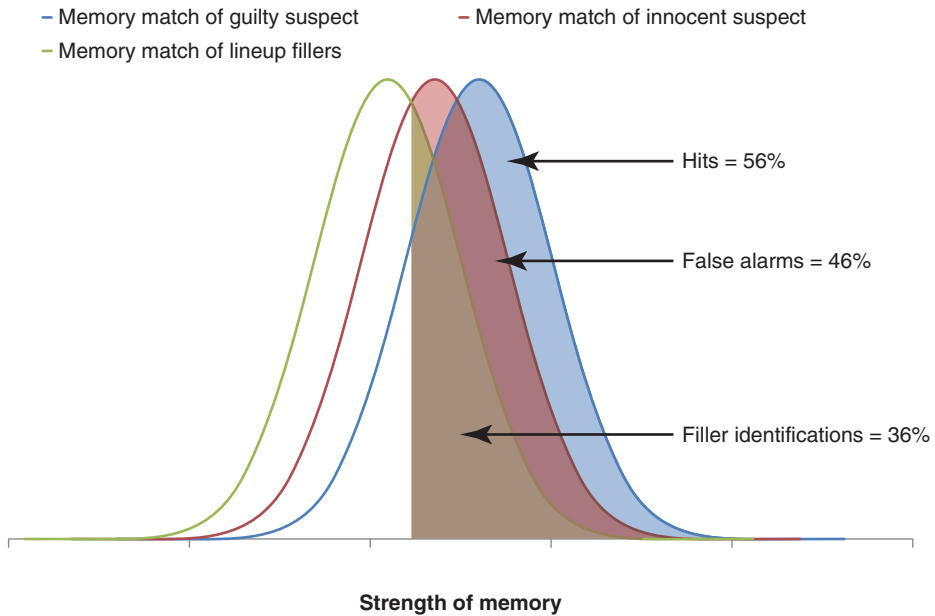
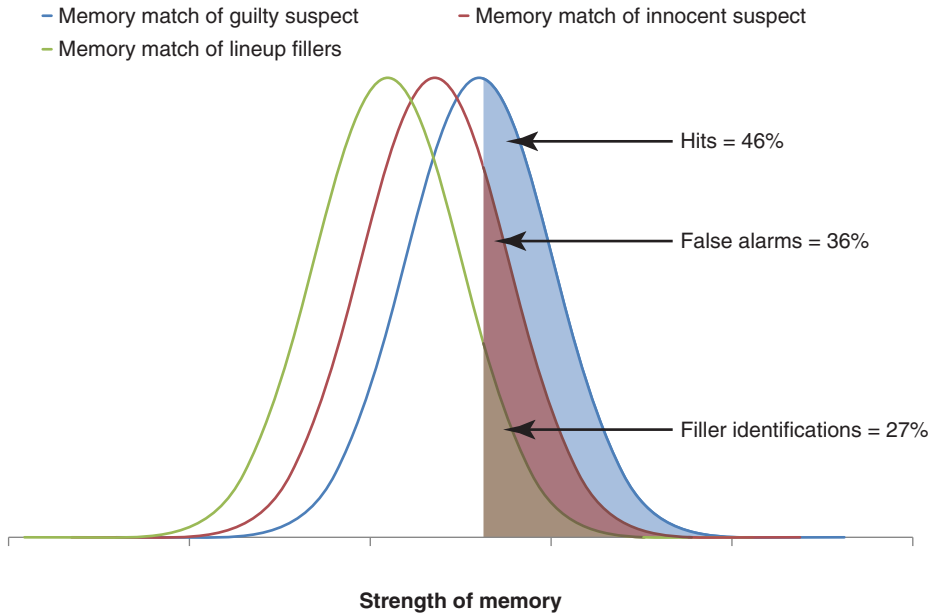


Figure 17.5 Effect of Changing the Decision Criterion on the Relative Proportion of Hits and False Alarms

For example, factors such as the distance of the observer from the perpetrator's face, the level of illumination, the angle of regard, the speed at which the face is moving in reference to the observer, and the degree to which facial features are obscured by clothing, accessories, or disguise all have an effect on the witness's ability to extract information from the face that will be useful in a subsequent identification procedure. Other factors influencing the informational component of memory may transpire in the retention phase between the criminal event and the witness providing information to law enforcement. For example, memory may decay over time, thereby reducing the amount of information the witness has available, and/or the witness may be exposed to other information that can alter his or her memory for the event (e.g., media coverage, discussions with other witnesses). In addition, some of the factors influencing the information component may also be produced during the witness's interaction with law enforcement (e.g., questioning style, lineup composition).

Some factors, however, influence the witness's *decision* about whether and how to report the information that he or she has retrieved. In SDT terms, these factors relate to the placement of the decision criterion. They do not influence the quality of the witness's memory or his or her evaluation of how strong a memory he or she has; rather, they influence the willingness to report the information. The factors that influence the willingness of persons to say, for example, that a particular photograph represents a perpetrator, include the value the observer places on the various outcomes that he or she anticipates might arise and a judgment of the likelihood that each of these outcomes would occur.

Thus, the witness may fear reprisals by the perpetrator and decide not to express his or her identification. Another witness may be fearful of the consequences of not cooperating with law enforcement and, as a result, attempt to identify the person from a lineup (or in court) whom he or she thinks is the person officials want to be identified. Or the witness may believe that he or she will be perceived as less credible if he or she can testify about many details of the crime and yet be unable or unwilling to identify the perpetrator. Some witnesses are very enthusiastic about identifying a perpetrator whereas others agonize over not being able to make an identification or identifying the wrong person. There are many motives possible, and all can influence the witness's decision to report information and/or make a positive identification. By nature, the decisional aspects exist at the time of the reporting decision. Some may be introduced by actions of law enforcement officers (e.g., the preidentification procedure instructions), whereas others may be the product of the witness's long-standing beliefs, emanating from personal values (e.g., a due process perspective toward crime, a desire to "help" law enforcement, etc.). Using this distinction between informational and decisional processes, we identify factors that influence eyewitness testimony and provide practical recommendations

for applying this knowledge to the assessment of eyewitness testimony in a given case.

EVALUATING EYEWITNESS TESTIMONY

Formally, the role of experimental psychologist as expert witness is to assist the jury in determining whether the eyewitness evidence identifies the defendant as the perpetrator or is indeterminate. "Assist" means to explain the science concerning factors that can influence identification accuracy and the complexities of interpreting that information in the case at hand.

Many other people, from police, to prosecutors, to policy makers, will have an interest in evaluating eyewitness evidence and may think themselves in a position to accurately assess the validity of the eyewitness memory and identification decision. Their ability to do so, however, is often limited, as demonstrated by the great many incorrect identifications made by eyewitnesses in the exoneration cases. Although this chapter is oriented toward the expert witness, the information we provide is beneficial for any individual attempting to evaluate the validity of the eyewitness evidence (e.g., law enforcement investigators, prosecuting/defense attorneys, judges, jurors).

Many relevant factors are inside the witness's head, a fact that hinders those seeking to share this information with others (police, prosecutors, juries). As such, identifying "ground truth" (i.e., what exactly occurred during commission of the criminal event) and producing an unequivocal answer regarding the accuracy of the witness's testimony is extremely difficult to achieve. As we discuss throughout the chapter, there is a gray area between memory for the perpetrator and the appearance of the suspect and factors beyond the memory match that drive the identification report. We have known of this gray area for a long time, but we know of it today in some detail. The identification is literally accurate or not depending in part on (a) how the witness thinks about his or her memory and (b) the consequences of the various ways in which the witness reports the information. Under these conditions of uncertainty, it is inappropriate for experts to produce *unconditional* ultimate opinions. The key point here, for nonjurors, is that expert testimony has to be conditional.

We turn now to the "conditions" attending assessment of eyewitness testimony. But we warn prospective evaluators (i.e., any individuals who might be in the position of assessing the accuracy of eyewitness evidence) that what follows is not always based on knowledge and skills available to those untrained in cognitive and social psychology, and it surely will not flow from common knowledge.

Important factors affecting eyewitness memory and identification are associated with three regions of the timeline for the criminal event and the subsequent

eyewitness identification procedure. These include (1) the attributes of the witness prior to observing the criminal event, (2) the attributes of the criminal event, and (3) postevent factors, including the actions of law enforcement and the court system.

ATTRIBUTES OF THE WITNESS PRIOR TO OBSERVING THE CRIMINAL EVENT

Many witness attributes can influence memory formation, retention, and testimony. Some may exert their influence on subsequent eyewitness testimony through informational processes and some through decisional processes. Attributes that influence informational processing include the witness's prior experiences with faces (at both the individual and the group level) and the witness's gender and age.

Experience With Faces. People achieve stable levels of face recognition accuracy by the time they are in their early to mid-teens. There is little evidence that any form of special experience advances face recognition ability above that of the normal population. In particular, studies show that, contrary to folk beliefs about the abilities of law enforcement officers, they are no better at face recognition than the general population. In addition, face recognition appears to be very difficult to train, particularly so that the enhanced performance lasts beyond training (Malpass, Susa, & Meissner, 2008).

Individuals readily infer personality traits and other global characteristics (e.g., attractiveness, baby-facedness, criminality) from faces (Hassin & Trope, 2000; Noor & Evans, 2003), and these personological attributions are used to mark faces in ways that are useful for subsequent recognition (see our later discussion of Doob & Kirshenbaum, 1973, in the section on lineup composition). However, personological memory tags are somewhat unique to the observer, and mapping personological attributions onto facial characteristics is inconsistent. For example, criminality is an important tag (e.g., "he looks like a criminal"), and a person in a police lineup who receives the "criminal" tag by agreement among observers is more likely to be identified as the suspect than others who are not tagged as "criminal" (MacLin & MacLin, 2004).

Familiar/Stranger Faces. Recognition for familiar faces is far more accurate than for faces of strangers seen only once. However, familiarity varies by degree, and the extent of the claimed familiarity should be investigated. Identifications are most likely to be correct when the person has been very well known to the witness over a long period of time, but, even then, difficult observation conditions (e.g., poor lighting, brief durations of exposure) may impede correct identification. As such, a high degree of familiarity does not guarantee correct identification.

Although familiar others are more likely to be correctly identified than strangers, eyewitnesses sometimes make errors when identifying individuals who are not well known but to whom they have had exposure (e.g., a regular customer in a store, someone from the neighborhood, etc.). These errors occur when the witness correctly recognizes the individual as familiar but forgets or confuses the source of that memory (Ross, Ceci, Dunning, & Toglia, 1994). As a result, the witness may infer that the experienced familiarity is because the individual is the perpetrator of the crime. For example, in a recent Seattle case, two eyewitnesses identified the defendant as the perpetrator without realizing he was familiar to them because they had briefly worked together in the past. Only after incontrovertible evidence of the defendant's innocence was produced did these misidentifications come to light and the prosecutor's office dismiss the charges. This familiarity can also arise out of exposure to an individual across multiple identification procedures throughout the course of the investigation (e.g., Deffenbacher, Bornstein, & Penrod, 2006)—a topic we discuss in more detail in a later section on repeated identification procedures.

Faces from Unfamiliar Regional Populations (e.g., Cross-Race Recognition). It is well established that members of one ethnic group or race have greater difficulty recognizing members of another ethnic group than they have recognizing members of their own ethnic group. Malpass and Kravitz (1969) were the first to show this in a laboratory, and there have been many replications over the years. A meta-analytic summary of these studies can be found in Meissner and Brigham (2001b), who estimated that individuals are 1.4 times more likely to correctly identify own group members and 1.56 times more likely to falsely identify out-group members. The practical significance of this so-called own group bias (OGB) is reflected in the DNA postconviction exonerations of prisoners in the United States. Of the 301 cases (as of 2012), approximately 71% involved mistaken eyewitness identification, and of these, approximately 50% were cases in which a White eyewitness mistakenly identified an African American suspect in a lineup or similar identification task.

In earlier research literature, the OGB was often called the own race bias or the cross-race recognition effect, but it has become clear that the focus on race is not a boundary condition of the effect, with several researchers reporting similar effects for age groups (Anastasi & Rhodes, 2005), gender (Wright & Sladden, 2003), and subgroups within broader ethnic or race categories (Chiroro, Tredoux, Radaelli, & Meissner, 2008). Why should human observers be better at recognizing members of their own group? Most researchers agree that differential perceptual experience must somehow underlie the own group recognition advantage. For extended theoretical accounts of the OGB, see Hugenberg, Young, Bernstein, and Sacco (2010) and Sporer (2001).

Gender. The gender of the witness is a factor that may play a role in the accuracy of the witness's report or the later identification, but differences in either respect are

not particularly significant, and little is known about why these differences may exist. Women are better than men at laboratory-style face recognition tasks (e.g., Rehnman & Herlitz, 2007) and at a variety of general tasks that involve memory (e.g., recalling names, episodic memories, object recognition; for a review, see Areh, 2011). There are some notable exceptions, however. Valentine and Mesout (2009), for instance, in an innovative study of visitors to a “horror labyrinth,” showed that female visitors were more likely to develop high levels of anxiety under stress and demonstrate poorer performance in a subsequent identification task.

Age. The effect of age on witness’s abilities has interested researchers, particularly the abilities of the very young and very old. Studies on child witnesses, particularly those below the age of 10 years, indicate that child witnesses are likely to make as many correct identifications as adults (i.e., when the suspect is in fact the perpetrator) but are more likely than adults to make false-positive identifications. These differences seem to disappear when children reach approximately 10 years, at which point their performance appears to be indistinguishable from adults between the ages of 20 and 50 (Parker & Ryan, 1993). Similar results have been reported by a number of authors for elderly adults, especially those over 70 years of age—that is, older witnesses are disposed to make more false alarms than younger witnesses. It is not clear what the reasons are for the effects of age on eyewitness performance, and some authors have speculated that, in the case of children, it could simply be the adult-controlled nature of the task that elicits the higher number of false alarms—in particular, a perceived pressure to comply with the experimenter or police official by choosing someone from the lineup (see Hobbs et al., Chapter 11, this volume).

ATTRIBUTES OF THE CRIMINAL EVENT AND TRANSIENT ATTRIBUTES OF THE WITNESS

The eyewitness typically is not afforded ideal opportunities for observation, and this should be considered when evaluating the identification or testimony. The event may have been of brief duration, or the eyewitness may not have had an unobstructed or clear view of the perpetrator. The duration of the event, the ambient lighting, the distance of the witness from the perpetrator, the physical layout of the environment, the involvement of the witness in the event (e.g., as victim or as bystander), and the possible demands on the witness’s attention should all be considered. The perpetrator’s appearance at the time of the event may have been markedly altered, either through deliberate disguise or some alteration to impermanent facial features, such as facial or skull hair.

The research literature on these questions is not well developed, however, and, given the complexity of some of the considerations, this is perhaps not surprising. For instance, a careful parametric calibration of the effects of lighting, particularly as these effects interact with the distance, duration, and nature of the event, would involve the construction of perhaps several hundred conditions and would leave

many other factors unstudied. Although studies of this scope have not yet been conducted, there is research on these important factors, as we describe next.

Distance. The distance between the witness and the event is inversely proportional to the accuracy of a later identification of the perpetrator, especially for previously unfamiliar faces (but not for familiar faces, which are robustly identifiable at greater distances; Greene & Fraser, 2002). There is also evidence suggesting that distance affects correct identifications of guilty perpetrators and incorrect identifications of innocent suspects, differentially. As distance increases, correct identifications of guilty perpetrators decrease but incorrect identifications of innocent suspects remain constant (Lindsay, Semmler, Weber, Brewer, & Lindsay, 2008).

Lighting. Angle of direction and degree of illumination have been systematically studied in face recognition research, showing that matches in the angle across encoding and retrieval lead to more accurate identification and that, as we would expect, greater levels of illumination lead to higher rates of identification (de Jong, Wagenaar, Wolters, & Verstijnen, 2005). Research on eyewitnesses supports these general trends (e.g., Wagenaar & van der Schrier, 1996), but there are some notable failures to replicate these patterns (e.g., Valentine, Pickering, & Darling, 2003).

Duration. Face recognition research indicates that longer exposure to a target face ("encoding duration" in the face recognition paradigm) is associated with improved recognition accuracy at a later point in time (Shepherd, Gibling, & Ellis, 1991), and research using eyewitness simulations supports this general trend (Meissner & Brigham, 2001b). However, it should be noted that this relationship is probabilistic and that many highly significant mistaken identifications of innocent suspects resulted from events of very long duration (e.g., the cases of Johnny Briscoe and Ronald Cotton; see Innocence Project, n.d.).

Disguise. Perpetrators may wear disguises or change their appearance in significant ways before or after the crime, and it is clear from the literature that this may be a significant impediment to later identification. Alterations to or disguises of the eyes (e.g., wearing sunglasses), mouth, and hair decrease identification accuracy (McKelvie, 1976; Wright & Sladden, 2003), and this is not surprising from a face recognition theoretical point of view, as human observers pay most attention to these features.

Alcohol/Drug Intoxication. Law enforcement interactions with intoxicated witnesses are fairly common (Evans, Schreiber Compo, & Russano, 2009), and laboratory research consistently demonstrates that alcohol and marijuana intoxication at the time of encoding significantly decreases quantity and accuracy of recall (e.g.,

Ranganathan & D'Souza, 2006; Steele & Josephs, 1990) and recognition accuracy (e.g., Maylor, Rabbitt, & Kingstone, 1987). However, a limitation of much of this laboratory research is that the memory tasks do not accurately represent the demands that are present in an eyewitness situation.

A smaller body of literature using more ecologically valid designs indicates that alcohol intoxication at the time of the event increases the likelihood of a false identification (e.g., Hilliar, Kemp, & Denson, 2010; Yuille & Tollestrup, 1990). The effects of alcohol consumption on recall accuracy are a bit less clear, with some studies indicating that intoxicated witnesses recall fewer details about an event (Sporer, 1992; Yuille & Tollestrup, 1990) and others indicating that alcohol intoxication has no effect on recall accuracy or susceptibility to misinformation (Schreiber Compo et al., 2011). Results of the one study that evaluated the effects of marijuana consumption on memory in an eyewitness context indicated that witnesses under the influence of marijuana at the time of the event produced less complete descriptions when questioned immediately after the event; however, this difference disappeared after a week's delay (Yuille, Tollestrup, Marxsen, Porter, & Herve, 1998). To our knowledge, there is no research on the effects of other illicit drugs on eyewitness memory.

Stress and Weapon Presence. Stress influences the quantity and type of information that is encoded by witnesses. However, research on the effects of stress on eyewitness memory has produced inconsistent findings, with some studies finding that stress can enhance memory and others finding that it produces a debilitating effect on witness memory (see Deffenbacher, Bornstein, Penrod, & McGorty, 2004, for a review). The different influences of stress on memory appear to be related to which neural control system is activated by the stressful situation (Deffenbacher, 1994). In less threatening situations, stress is more likely to elicit an *arousal* mode resulting in decreased physiological arousal and attention focused on the most important aspects of the criminal event. The arousal mode tends to increase witness memory for central but not peripheral details. However, in more threatening situations, the stress is more likely to elicit an *activation* mode leading to increased cognitive anxiety (worry) and physiological arousal. As anxiety and physiological arousal increase to very high levels, witnesses may experience a catastrophic decrease in memory performance. A meta-analysis of the research literature shows that situations that produce an activation response are more likely to produce a deficit in eyewitness memory that results in fewer correct identifications in target-present lineups and decreased performance in recall (Deffenbacher et al., 2004).

When the eyewitness observes a crime in which the perpetrator brandishes a weapon, it is often assumed that recognition and recall accuracy will be adversely affected, partly due to the highly stressful nature of the experience. A reasonable assumption, supported by experimental evidence, is that witnesses to crimes committed by armed perpetrators tend to focus their attention on the weapon,

thereby producing poorer memory for the perpetrator than witnesses faced with unarmed individuals. Although it is reasonable to assume that this weapon focus effect (WFE) is a product of the stress caused by the threat associated with a weapon, a weapon's presence can be detrimental to witness memory due to the "attentional capture" that is caused by the unexpectedness of the weapon in the environment (Pickel, Ross, & Truelove, 2006). Attentional capture causes witnesses to focus their attention on the central details of the scene (i.e., the weapon), resulting in decreased accuracy for the peripheral details (i.e., the offender). Laboratory studies on the WFE indicate that it impairs a witness's ability to describe the perpetrator; however, its effect on identification accuracy is much smaller (Fawcett, Russell, Peace, & Christie, 2013).

POSTEVENT FACTORS, INCLUDING THE ACTIONS OF LAW ENFORCEMENT AND THE COURTS

To this point, the relevant factors that contribute to eyewitness memory have all been beyond the control of the criminal justice system. This is because they have operated prior to the involvement of law enforcement. They are for that reason also less easy to assess and evaluate. However, from the point at which law enforcement becomes involved, there is generally a better record of events, and this affords a better opportunity to assess the reliability and validity of the eyewitness's evidence. For this reason, the expert should be especially familiar with these postevent issues, as they will likely be among the most common topics for which he or she has information. In addition, the expert should be sensitive to these issues under the control of law enforcement, because, although many aspects of the witness's experience may affect his or her memory, the actions of officials can affect witness memory so significantly as to render subsequent witness memory and reporting uninterpretable. Some postevent factors exert their influence on informational processes; and others affect decisional processes.

Delay. The time delay between the criminal event and the eyewitness's identification or testimony can vary considerably. It may be several days or longer between the event and the initial police interview and many months between the event and the trial. The danger is that the witness's memory may deteriorate in this period, in line with Ebbinghaus's (1886/1962) empirical observations that the probability of successful memory retrieval declines exponentially over time. In order to circumvent this, many police precincts have practice rules about the length of the acceptable delay between a crime and the elicitation of information from an eyewitness. The question of delay has been addressed frequently in the face recognition literature but less frequently or rigorously in the eyewitness literature. Although research indicates that correct rejections of lineups containing innocent suspects are made more frequently when delays are shorter, it is not clear that correct identification of guilty suspects is assisted by shorter delays (e.g., see Dysart & Lindsay, 2007).

Interviewing Techniques. Eyewitnesses typically are asked to provide a narrative report about the crime early in the investigation. Many times these reports are quite incomplete, and witnesses often fail to report much detail. This failure can occur because information is not available in memory, due to it not having been encoded in the first place or because it has been “lost” from memory (e.g., through decay or distortion). Alternatively, the information may be available in memory but not accessible (Tulving & Pearlstone, 1966), or the information may have been available and was accessed (i.e., recalled) but not reported (e.g., the witness did not think it relevant to the case). These last two aspects (accessibility and reporting) are influenced, in part, by the investigator’s interview techniques.

The typical investigative interview is relatively inefficient at eliciting information from witnesses (for a review, see Fisher & Schreiber, 2007). This inefficiency stems from a reliance on techniques that do not encourage memory retrieval. Although law enforcement agencies employ different standards and training in regard to interviewing, the interview process can be described as consisting of (a) the presence of predetermined topics; (b) the asking of questions about each of these topics; and (c) placing the interviewer in overt control of the interview, thereby relegating the witness to a passive role. In addition, investigators tend to rely on closed-ended questions (e.g., How tall was he? What color hair did he have? What did you see him do when he entered the store?) rather than encouraging witnesses to provide an open-ended, free narrative of the event. Not only do closed-ended questions discourage detailed witness reports, but they are more likely to be suggestive in nature (e.g., was the car blue?) than open-ended questions. As a result, police interviews tend to elicit less information from witnesses and increase the potential for inaccurate information.

Given the relative inefficiency of typical police interviews, psychologists have developed techniques that enhance the amount of information witnesses can provide in their narrative reports (Technical Working Group on Eyewitness Evidence, 1999). One of the more extensively researched approaches for interviewing adult eyewitnesses is the Cognitive Interview (for a brief overview, see Fisher, Ross, & Cahill, 2010), a package of techniques that increases the amount of information eyewitnesses report about a criminal event without decreasing the overall accuracy of that information (Memon, Meissner, & Fraser, 2010).

Consistency Across Witness Reports. Witnesses often are asked to provide multiple narrative reports over time, and these reports may be inconsistent. Although witnesses who provide inconsistent statements typically are viewed as unreliable (Fisher, Brewer, & Mitchell, 2009; Oeberst, 2012), inconsistency across reports does not necessarily mean the statements are inaccurate (e.g., Oeberst, 2012; see Fisher et al., 2009, for a review).

There are at least three types of inconsistency. The first, and the one most likely to signal inaccuracy in the report, occurs when a witness provides a specific detail

during an early interview (e.g., the perpetrator wore a red shirt) and subsequently reports contradictory information (e.g., he wore a blue shirt). Other inconsistencies include reporting a detail during an earlier interview but failing to report it subsequently (i.e., forgetting) or failing to report a detail in an earlier interview and first reporting it in subsequent interviews (i.e., reminiscence). Although laypersons are skeptical of the accuracy of forgotten or reminiscence details (Oeberst, 2012), research on basic memory processes indicates that forgotten and reminiscence memories can be quite accurate (Erdelyi, 2010), and research specifically addressing eyewitness memory indicates that, absent poor interviewing techniques, these details are just as likely to be accurate as consistent details (i.e., those details reported across all interviews; e.g., Oeberst, 2012).

Memory-Refreshing Techniques (i.e., Hypnosis). Studies of hypnosis as a memory-refreshing technique consistently show that it produces errors in eyewitness memory (see Mazzoni & Lynn, 2007, for a review). This deleterious effect is a product of increased suggestibility and the resultant increased likelihood of false memories. As a result, many states have a per se rule against admitting testimony from a witness who has previously been hypnotized in relation to the case at hand (Webert, 2003).

Composites. Witnesses frequently are required to produce visual likenesses of the perpetrator's face, in interaction with police officers, using specialized software programs, such as Identikit or FACES. Composite portraits typically are published widely on television, in the print media, and in public places. In some highly publicized cases, the release of these portraits has led to a deluge of information from the public about the supposed identity of the person depicted, most of which is unhelpful. However, the composite portrait may well be the basis for choosing a suspect for further investigation or may even be the sole basis for placing the suspect in a lineup. The expert should be alert to the conditions under which the composite was constructed and the resulting portrait was displayed to witnesses and the public. Composite portraits are constructed in more cases than defense attorneys may be aware of and these composites may not be included in the final documents provided to prosecutors for many reasons, including that they were considered poor representations or were not considered evidence of the suspect's guilt in their own right.

Over 40 years of research has shown that, with very few exceptions, witnesses generally are unable to produce composite portraits that resemble the perpetrator sufficiently for someone else to identify him or her—indeed, in laboratory studies, performance of evaluator-witnesses is usually at chance level (for an overview of this research, see Davies & Valentine, 2007). Recent technological developments that allow witnesses to construct faces holistically (rather than in the piecemeal fashion of the older systems) with so-called eigenface models (cf. Frowd, Hancock,

& Carson, 2004; Tredoux, Nunez, Oxtoby, & Prag, 2007) report slightly better results, but the representations elicited from these systems are still too poor to challenge the conclusion that composite portraits very rarely resemble the perpetrators.

More controversial is whether producing or viewing composite portraits contaminates witness memory for the perpetrator. For instance, if the composite portrait shows a face with a mustache when the perpetrator had none, the witness may come to believe that the perpetrator indeed had a mustache (see Jenkins & Davies, 1985, for an experimental demonstration). The evidence for this claim is mixed. It seems clear that when the composite contains a misleading item of information, this can be incorporated into the witness's later reports. Although some authors have gone further and argued that merely creating a composite portrait can damage memory (e.g., Wells, Charman, & Olson, 2005), the balance of evidence suggests that this is not the case (cf. Meissner & Brigham, 2001a).

There is another problem with the use of composites. The police often claim that composite images displayed to the public are used only for investigative leads. However, when a member of the public claims that the composite looks like someone who is not yet a suspect, the investigation may come to focus on that person, with other leads getting less scrutiny (as happened in the Northrop/Davis case).

Lineup Composition. The notion of a biased lineup is so well known that it is urban lore. A notorious example is cited in Ellison and Buckhout (1981). Minneapolis police arrested an African American man as the prime suspect in a case they were investigating and placed him in a parade otherwise constituted only by White men. Similar cases have occurred elsewhere in the world; in South Africa, in the case of *Pelwani v. S.* (1963), a lineup was composed of three Indian men and three White men, the suspects being three Indian men. In these cases, the lineups were clearly (perhaps self-evidently) biased—and easily identifiable as such—but more usually the bias is difficult to identify and describe. In addition, this assessment of bias will always depend on inherently subjective observations, especially the case for infractions of the requirement of physical similarity.

Despite these difficulties, psychologists have developed a set of methods and measures for assessing the fairness of police lineups, and these methods are an invaluable part of the toolbox for any expert witness in this area. We give more attention to this section of the chapter than to many others, because we consider the work to be of great importance to expert witnesses.

Work on measures of lineup fairness dates from an article published by Doob and Kirshenbaum (1973) in which they reported a test of the fairness of a police lineup in a Canadian case, *R. v. Shatford*. The authors questioned the eyewitness identification in this case, given that the sole description she was able to give the police was that the perpetrator was "attractive." Doob and Kirshenbaum showed a photograph of the identification parade to 23 participants along with the witness's "description"

of the suspect. They reasoned that subjects who had not been present at the crime should not be able to identify the suspect with a probability exceeding random selection ($1/\text{number of lineup members} = 1/12$ in this case). However, 64% of the witnesses correctly identified the suspect. Doob and Kirshenbaum computed the probability of this occurring to be less than 0.001 and concluded that the lineup was biased. The authors suggested that this method—which has come to be known as the method of the mock witness—could be used in general to assess the fairness of identification parades. The method posits that the lineup is fair when the proportion of witnesses choosing the suspect does not exceed that expected on the basis of mere random choice.

The purpose of mock witness evaluation is to assess the structural fairness of the lineup—whether the fillers are adequate alternatives to the suspect and whether the suspect stands out from the fillers. Its purpose is not to predict what witnesses would do. The central assumption is that, if persons who have had no exposure to the suspect prior to viewing the lineup select him or her at a rate greater than chance, then the lineup is biased toward identification of the suspect. The intention of Doob and Kirshenbaum's (1973) study was to provide a measure of lineup fairness. Information regarding lineup fairness is provided to some extent by the proportion of accurate mock witness choices, but it is certainly not a complete assessment of the lineup's quality. An important additional feature of a lineup is the number of plausible fillers that it contains.

Malpass (1981; Malpass & Devine, 1983) argued for a distinction between lineup size and lineup bias. *Lineup size* refers to the number of plausible members that the lineup contains, and it contributes directly to the fairness of the lineup by decreasing the probability that the suspect is identified by a witness who willfully chooses at random. *Lineup bias* is reflected by the extent to which mock witnesses choose the defendant at rates greater (or smaller) than chance expectation. When the proportion of mock witnesses choosing the suspect equals that expected by chance (i.e., $1/k$, where k is the number of lineup members), the lineup is unbiased. When it deviates from the expected value, it is biased. The measure has interpretable limits at both the upper and lower ends. As the proportion approaches 1.0, mock witnesses are selecting the suspect to the exclusion of the fillers, and, when the proportion approaches zero, witnesses are failing to choose the suspect at all.

However, one can expect the proportion to show random sampling variation, and an important question thus concerns how to interpret the observed proportion. For example, if 7 of 20 mock witnesses (35%) choose the suspect from a 5-person lineup, we need to know whether this could reasonably be explained as chance variation from the expected value of 4 of 20 (20%). Doob and Kirshenbaum (1973) used a z -test to make this decision, but Tredoux (1998) recommended the direct calculation of binomial probabilities, because that method does not make the assumption of an approximating distribution. Wells, Leippe, and Ostrom (1979) and Tredoux (1998) recommended reporting the proportion as a confidence interval (CI) rather than as a point estimate.

Usually experts are interested in situations where the lineup is biased against the suspect (i.e., where the suspect is chosen by a higher proportion of mock witnesses than is expected by chance), but it is also possible that the lineup could be biased in favor of the suspect. Imagine a lineup where 1 of 30 mock witnesses (3%) chooses the suspect from a 5-person lineup. This is significantly less than expected by chance (binomial $p < 0.01$). Although this might appear to be of no consequence to the police or to eyewitness researchers, as the suspect's liberty has not been jeopardized unfairly, this perspective only recognizes the problem of false identifications and fails to recognize that a second kind of error can be committed when using a lineup, namely the failure to identify a guilty perpetrator. Lineups in which suspects are chosen by mock witnesses at levels significantly below chance are poorly constructed and run the risk of committing the second kind of error. Such lineups may be rare, but it is useful to extend the reasoning behind the measure of lineup bias so that it can be used as a warning indicator for both types of fundamental error.

Malpass (1981) suggested "effective size" as a measure of the number of plausible lineup members. Effective size has a maximum value of k —the number of lineup members—and a minimum of 1 (assuming that mock witnesses must choose one member of the lineup). The assumption underlying the notion of effective size is appealing: One or more of the fillers in a lineup may present an inadequate test of a witness who has little more than general knowledge of the appearance of the offender, and we should not take the ability of a witness to reject such fillers very seriously. The calculation of effective size acts on this assumption by reducing the nominal size of the lineup according to departures of proportionate identification of individual fillers from that expected by an equi-probability model (every lineup member drawing the same number of mock witness choices; see Malpass, 1981, for details on the calculation).

However, there are a number of weaknesses with the effective size measure (see Tredoux, 1998). Most important, there is no known sampling distribution for effective size, which weakens the kinds of conclusions researchers or practitioners can draw about particular lineups. Tredoux (1998) suggested an alternative computational formula for effective size that retains most of the desirable properties, with the important added benefit of a known sampling distribution. The statistic produced by this formula is known as E' . Specifically, E' has a maximum value of k —the number of lineup fillers—and a minimum value of 1 (assuming that mock witnesses are required to choose a lineup member). If some lineup members attract more choices than others, this will result in a reduction of the value of E' from k toward 1. Methods for using E' inferentially can be found in Tredoux (1998). These methods are relatively uncomplicated and can be incorporated into a spreadsheet for easy computation (for an example, download <http://www.eyewitness.utep.edu/images/size-calc.xls>).

Each of the measures discussed can be used in a practical sense to evaluate lineup fairness. The method in each case is similar to that employed in mock witness research—participants are provided a description of the suspect and must select

the parade member who best fits this description. The measures are then calculated on the basis of the mock identifications.

There are many audiences for the concepts and empirical results from actual mock witness assessments—researchers, law enforcement officers, judges, defense and prosecuting attorneys, and jurors—all of whom may draw conclusions about lineup fairness. It would be useful to make these ideas and their quantification available and conceptually accessible to these constituencies. The problem is to represent the quantifications of size and bias in a way that is understandable. This is less a problem for professional participants in the criminal justice system and much more a problem for novices (i.e., jurors). In our experience, simple graphic presentations of choice rates for lineup members against a background of chance expectation are easily understood.

The techniques and procedures needed to implement mock witness evaluation of a lineup can be assembled from information referenced in this chapter. For practitioners in the criminal justice system, however, constructing a complete procedure in this way is time consuming and contains many uncertainties and disincentives. Detailed instructions do not exist in the law enforcement literature. Malpass provides detailed instructions and access to computer-based calculations of the relevant statistics on his Web site (<http://eyewitness.utep.edu/consult02a.html>). As training in lineup construction and evaluation techniques becomes more common for law enforcement personnel, access to these techniques should become more widely available.

In the current legal environment, bias against the suspect in a lineup is the most useful argument for suppression of identification evidence, because it is directly related to the “second prong” of the *Manson v. Braithwaite* (1977) “totality of the evidence” reasoning. In *Manson*, the United States Supreme Court identified five “independent” indicators of the accuracy of an identification: (1) opportunity to view, (2) attention to the offender’s face, (3) accuracy of the witness’s description, (4) time to identification, and (5) witness certainty. In fact, however, scientific study has shown the relation of each of these indicators to witness identification accuracy to be flawed, complex, subject to outside manipulation, inaccessible to witness memory, insufficiently strong to serve as an indicator of identification accuracy, or blatantly incorrect. The *Manson* case is a classic example of applying common knowledge in lieu of that developed from experimental psychology and psychometrics. When examined through the lens of experimental psychology, it was a poisonous failure because the court’s screens for suggestion fail to catch eyewitness identifications based on suggestive lineups and administration procedures and allows them to go before fact finders who are without both the tools to evaluate the evidence and the assistance of qualified experts (see Wells & Quinlivan, 2009, for a detailed analysis).

Postevent Information. Witnesses may be exposed to information from a variety of sources after the event has taken place. They may engage in conversation with co-witnesses; read newspaper accounts; view television reports; and be interviewed

by police officers, lawyers, and other officials. The important question here concerns what happens to their recollections or identification abilities when this information is contrary to what they witnessed and recollect. This has been one of the central questions for empirical research into eyewitnesses, and it is particularly identified with the work of Elizabeth Loftus and her colleagues (for a review of 30 years of research, see Loftus, 2005).

Early work established that subtle questioning can influence the recollection of speed estimates, the perceived severity of an automobile accident, and specific details about an event (Loftus, Miller, & Burns, 1978; Loftus & Palmer, 1974). For instance, witnesses can easily be misled into reporting the presence of a yield instead of a stop traffic sign or that the perpetrator had a mustache when he in fact did not (Jenkins & Davies, 1985). Later work established that entire events can be introduced into a witness's memory, even false events that are relatively traumatic in nature. Loftus and Pickrell (1995) were able to introduce false memories in students for a traumatic event alleged to have occurred in their childhood, namely separation from their mother in a shopping mall, and Wade, Garry, Read, and Lindsay (2002) implanted a false memory for a childhood trip in a hot air balloon in a sample of adults.

The potential for postevent information to contaminate memory reports increases with long delays between the event and the recollection and long delays between the misinformation and the recollection in addition to a variety of factors that make the discrepancy between the original information and the postevent information difficult to discern (e.g., if the misinformation is introduced by an authoritative or trusted source; Loftus, 2005).

Repeated Identification Procedures. In some cases, witnesses may be asked to participate in multiple activities assessing their recognition memory for the perpetrator. For example, they may view mug books in an attempt to identify a perpetrator, they may assist in the construction of a composite of the perpetrator, or they may be shown one or more photo montages prior to participating in a live lineup. Research on the effects of these repeated identification procedures has found that exposure to a suspect in prior identification procedures increases the likelihood that the suspect will be identified in subsequent procedures, regardless of whether he or she is the perpetrator. This tendency for carryover effects from one procedure to the next is even greater if the witness has a weak memory for the event (Pezdek & Blandon-Gitlin, 2005) or identified the suspect in a prior identification procedure (Deffenbacher et al., 2006; Haw, Dickinson, & Meissner, 2007). A by-product of this repeated exposure to a suspect across identification procedures is the artificial inflation of witness confidence produced by the increased familiarity the witness experiences as a result of multiple exposures. The implications of confidence in assessing witness accuracy are discussed in our later section on witness confidence and response latency.

Instructions Provided to Witnesses Prior to Identification Procedures. Legal scholars and psychologists have pointed to the inherent suggestiveness of many identification procedures (Wells, 1984; Williams & Hammelman, 1963). The police lineup is one of these inherently suggestive procedures, even though one of its explicit purposes is to reduce suggestiveness. Williams and Hammelman (1963, p. 487) argued that the social situation in which a lineup identification is attempted invites the witness to draw the inference that the police “got their man” and put him in the lineup for the witness to select—it is more like a multiple-choice quiz than a recognition task. Eyewitness researchers point out how multiple aspects of lineup identification tasks are prone to suggestive influences (Malpass & Devine, 1981; Wells, 1984). The lineup members may be dressed differently, which may suggest the identity of the suspect to the witness, or the lineup may be administered in such a manner that the identity of the suspect is revealed—indeed, there are myriad potential sources of bias. It is important to note that bias might not be intended by the law enforcement officer managing the lineup, as the power of the situation can overwhelm the best efforts to make the process as fair as possible. The inadvertent influence of expectations is now well demonstrated in a variety of contexts, including classroom settings (Rosenthal & Jacobson, 1968) and situations where obedience is required of subordinates (Milgram, 1963).

Some courts have recognized the suggestiveness of police lineups and attempted to remedy it by regulating how lineups are conducted. For instance, the police may be urged to ensure that (a) the witness does not see the lineup being assembled (as this may reveal who the suspect is); (b) the lineup members are dressed similarly (even if this means instructing the suspect to change his or her clothes); and (c) the officer conducting the lineup is not the investigating officer in charge of the case (as he or she will have a vested interest that the witness will choose the suspect), or—more dramatically—that the officer conducting the lineup does not know who the suspect is.

The instructions issued to the eyewitness at the lineup may strongly influence the task expectations the witness has and the provision of these instructions is an important micromoment in the process. Malpass and Devine (1981) demonstrated that an instruction to witnesses that the lineup administrators assumed the perpetrator to be present resulted in significantly more participants identifying one lineup member when compared with an instruction that the perpetrator may or may not be present. A number of studies have replicated this finding, and, although there is some dispute about whether unbiased instructions simply affect rates of choosing or whether they result in more discriminating decisions (cf. Clark, 2005), the National Institute of Justice guidelines for conducting police lineups now require warning witnesses that the perpetrator may not be present in the lineup (“Advise the witness that the person who committed the crime may or may not be present in the group of individuals,” Technical Working Group in Eyewitness Evidence, 1999, p. 32).

It is important to recognize that the influence on an eyewitness may begin well before the formal identification process. Consider the likely telephonic interaction

between witness and police officer when the witness is asked to attend the identification procedure: Quinlivan et al. (2012) show that a suggestion to the witness prior to the formal admonition may in effect nullify the admonition.

Blind Administration. A number of eyewitness researchers have urged that lineups be blind administered, meaning that the police official administering the lineup should not know who the suspect is (e.g., Wells et al., 1998). In addition, when the lineup is corporeal (i.e., a live lineup), the fillers should not know who the suspect is. This recommendation places a considerable burden on the police, potentially committing additional police officers to the task and reassigning resources from where they might be needed. It might also be a difficult thing to achieve in practice, particularly in small police precincts or geographical areas where the suspect is well known. For that reason, empirical evidence on the benefits of blind administration ought to be convincing and clear, but this is not so. Although some research indicates that properly blinded administration of lineups results in identifications that are more diagnostic of guilt than less well blinded administration (Greathouse & Kovera, 2009), other research has produced different results (see Russano, Dickinson, Greathouse, & Kovera, 2006, for a review), and a recent synthesis of the evidence suggests that gains achieved by blind administration regarding reducing false identifications of innocent suspects is directly offset by a reduction in identifications of perpetrators (Clark, 2012).

Identification Procedure Presentation Format. The classic, time-tested way of conducting a police lineup comes to us from 19th-century England. A suspect is asked to stand alongside a suitable number of people, “‘who are as far as possible of the same age, height, general appearance (including standard of dress and grooming) and position in life’ as the suspect” (Devlin, 1976, p. 113), and the witness is asked to indicate if the perpetrator is standing in the line and, if so, who he (or she) is. The police lineup was designed in part to deal with the problem of in-court identifications, which are still required as part of court formality in criminal trials.

The formation of lineup members alongside each other is known as the simultaneous lineup, but it is possible to imagine many other ways of conducting an identification. An alternate method that police often use when the identification takes place very soon after the event is to stage or order a “showup,” where the witness is confronted directly with the suspect and asked if this is the perpetrator. Yet another method, devised by Rod Lindsay and Gary Wells (Lindsay & Wells, 1985), is to present an undeclared number of lineup members to the witness one at a time and require the witness to indicate for each one whether he or she is the perpetrator. Each method has its drawbacks.

The most pernicious form of eyewitness identification evidence is in-court identification. It deprives the defendant of all of the protections against false identification

conceived by psychological scientists and others concerned with reducing mistaken identification:

- All protections related to the construction of a fair lineup are absent.
- All of the protections afforded a suspect at a confrontation or showup are negated.
- The situation is suggestive because of the presence of the defendant.
- The prosecutor is free to preface the request for identification with a summation of the many reasons why the witness should be able to make the identification and why it is the right thing to do.
- The admonition that the actual offender may not be present is absent or undermined.
- The admonition that the person administering the identification does not know who is the suspect is obviously false.
- The admonition that the witness is not required to make an identification is undermined.

The independent source doctrine—the idea that a person can make an in-court identification based on a pristine, “original” memory for the perpetrator, untouched by and independent from previous identification attempts with the suspect or a photo of the suspect—is largely illusory. Yet pretrial corporeal identifications and in-court identifications that are preceded by photospread identifications are often accompanied by a request for the witness to recall the original event and make his or her in-court identification based only on this “independent” source of memory and without any effect of any occasions on which the witness may have observed the defendant or his or her photo. Even the idea that the witness can know whether he or she can accomplish this independent-source maneuver is itself illusory. These ideas come from common knowledge and not from the science and empirical research relating to eyewitness identification. This belief that in-court identifications are based upon an independent source of memory developed from the criminal event is a product of legal tradition for which there is little or no basis in science. It is mainly used as an end run around findings of suggestiveness in pretrial identification procedures (Sobel, 2007, p. 154).

Evaluation of the validity of an in-court identification in any given case is nearly impossible. From the perspective of the science of memory and identification, this procedure serves only as a dramatic performance for its influence on the jury’s decision making. The role of a psychological scientist in a consulting/evaluator role is strong here, because the in-court identification contains so many elements that the science knows elevates the likelihood of identification and ignores memory changes.

The showup is preferred to in-court identification, but it is nevertheless widely recognized as highly suggestive and especially sensitive to a range of decisional

factors. A witness who thinks that the police generally get their man may be inclined to identify the suspect as the perpetrator, even though his or her memory is lacking. This would be less problematic if the witness were shown a properly selected group of people (typically six in most U.S. jurisdictions) even a witness with virtually no information about the identity of the perpetrator could not randomly guess the identity of the suspect more than one-sixth of the time. In a lineup, this protection is no longer present, and legal texts have advised against it for a long time. However, studies that have formally tested whether lineups and showups differ in terms of their promotion of (a) correct identifications of guilty suspects and (b) protection of innocent suspects have not shown a general advantage for lineups, except in cases where the suspect and perpetrator strongly resemble each other (see the meta-analysis by Steblay, Dysart, Fulero, & Lindsay, 2003).

Because the sequential lineup has been promoted so aggressively, several jurisdictions require sequential lineups rather than simultaneous ones. The sequential lineup is said to elicit nearly as many correct identifications of guilty perpetrators as the traditional simultaneous lineup but only 40% as many incorrect identifications of innocent suspects (e.g., Steblay, Dysart, Fulero, & Lindsay, 2001; Steblay, Dysart, & Wells, 2011). However, a growing list of original and synthesis studies have suggested that the trade-off between correct identifications and false alarms is equivalent across the two kinds of lineup (e.g., Clark, 2012; Gronlund, Carlson, Dailey, & Goodsell, 2009; McQuiston-Surret, Malpass, & Tredoux, 2006; Meissner et al., 2005) and that sequential lineups simply make witnesses more conservative (i.e., less likely to choose anybody). In other words, sequential lineups lose correct identifications of guilty perpetrators in direct proportion to the incorrect identifications of innocent suspects they avoid. In terms of the informational and decisional aspects of eyewitness testimony, the sequential lineup does not differ from traditional police lineups in any way that has to do with the quality or kind of information that the witness is able to retrieve; instead, it affects the willingness of a witness to choose someone from the lineup.

Multiple "Laps" in Sequential Lineups. In their original investigations, researchers were clear that sequential lineups should be discontinued once the witness makes a choice, whether it is the suspect or a filler. Indeed, part of the rationale for the procedure was that an identification of a filler should be considered indicative of a poor memory. In the implementation of sequential lineups into legal and police practice in the United States, however, witnesses are usually allowed to view the remaining lineup members even after making an identification and are afforded multiple laps (i.e., viewings of the entire sequence of lineup members). Police officials have been reluctant to conduct a lineup procedure in which the witness might never get to view the suspect and wish to accommodate witnesses that request a second viewing of the lineup members. Recent empirical data suggest that the use of multiple laps reduces the value of witness identifications from

sequential lineups. Steblay, Dietrich, Ryan, Raczynski, and James (2011) showed in two experiments that witnesses who viewed two laps made more errors than witnesses who viewed only one lap. It is not clear whether multiple laps affect performance on simultaneous lineups. It would be very unusual for a simultaneous lineup to be administered in this way.

OTHER INFORMATION AVAILABLE TO THE EXPERT

Throughout this chapter, we have advanced a general framework for considering the factors that affect the reliability of eyewitness testimony. There are, however, some factors that do not fit clearly into this framework. In this section we identify these other factors and discuss how they relate to the evaluation of eyewitness testimony.

Witness Confidence and Response Latency. The U.S. Supreme Court, in *Manson v. Braithwaite* (1977), held that witness confidence in the accuracy of the identification is a valid indicator of accuracy. Psychological scientists were skeptical, and confidence has since been intensively studied in the research literature. Although confidence–accuracy correlations are often statistically significant, overall, they are modest (average $r_{\text{unweighted}} = .29$, 95% CI [.25–.33]; Sporer, Penrod, Read & Cutler, 1995). Further analyses have revealed that the magnitude of this confidence–accuracy relationship is greater for those who identify an individual in a lineup ($r_{\text{unweighted}} = .41$, 95% CI [.33–.46]); however, these results indicate that accuracy accounts for only approximately 16% of the variance in confidence ratings made by choosers (Sporer et al., 1995). This finding stems from the fact that although choosing and accuracy share certain precursors (e.g., memory strength), they do not have many other precursors in common. Among those not shared is the effect of feedback given to the witness after the identification. Confidence statements are notoriously easy to manipulate through feedback (Luus & Wells, 1994). That there are nonshared precursors should raise serious caution. Recent research suggests that confidence ratings may be more useful when made under time pressure (Brewer, Weber, Wootton, & Lindsay, 2012) and when the witness is extremely confident in his or her decision (Brewer & Weber, 2008). The question of course is whether the science can reach an acceptable low and stable level of error in predicting accuracy from confidence ratings, given the instability of confidence and the unshared precursors.

One of the significant limitations of relying on confidence judgments as an indicator of the witness's decision accuracy is the inconsistency regarding when (and how) the confidence judgment is elicited. Although research by Brewer and his colleagues has suggested that confidence may be a greater indicator of accuracy than previously believed (see Brewer & Weber, 2008, for a brief overview), it is important to note that this research relies on collection of confidence judgments

immediately after the identification decision and without providing any feedback to witnesses. In police investigations, witness confidence often is collected at many points subsequent to the identification decision, including during testimony at trial. As previously discussed, feedback and other forms of postevent information can manipulate witness confidence, thereby reducing the usefulness of confidence judgments collected later in the investigation.

Other difficulties include incommensurate forms of eliciting confidence judgments, which may not be equivalent or call on the same set of cognitive processes or source monitoring tags. Examples include the form through which the witness chooses to provide his or her confidence judgment (e.g., verbal expressions such as “very sure” or “pretty sure”; numerical ratings on a scale from 1 to 10); and variations in the question asked [“How sure (certain) are you that (i) the person you identified is the person who (describe crime), or (ii) your identification is correct, or (iii) this is the person who you saw on (date) at (place) who (describe specific act)?”]. Empirical research evaluating the influence of the question and response form on the confidence–accuracy association is not available.

Another difficulty surrounds the familiarity witnesses have with making quantitative judgments reflecting their internal states. Readily identifiable segments of the population have limited education and fluency such that the language used in instructions and information requests and linguistic forms made available for expression of confidence distort whatever relationship confidence and accuracy might have in a given instance.

Measuring subjective states, such as confidence, and finding the most appropriate way to ask the seemingly simple identification question itself present problems. Should one ask whether the witness finds anyone in the lineup to be familiar, and then, subsequent to an affirmative answer, ask the witness the circumstances under which the familiarity was achieved? Or should one ask the witness to identify the offender (criminal?) or the person who (description of crime)? Should witnesses simply be given a yes/no choice for whether the perpetrator is in the group of persons (photos) shown, or should they be given the option to say whether any member of the lineup *is* the offender, *resembles* the offender, or *is not* the offender? Should they be asked these questions about each member of the lineup, whether it is presented simultaneously or sequentially? It is disconcerting that the research literature does not provide clear and certain guidance.

Although most research suggests that the utility of confidence judgments as indicators of witness decision accuracy is limited, a small body of literature consistently has found that the speed with which a witness makes an identification decision (i.e., response latency) is reliably associated with the accuracy of that decision (e.g., Brewer, Caon, Todd, & Weber, 2006; Dunning & Perretta, 2002; Sauerland & Sporer, 2009). This research has consistently shown that immediate identification decisions are much more likely to be accurate than slower decisions; however, these studies have been unable to identify a specific latency window that discriminates between accurate and inaccurate responses. As a result, outside of immediate identification

decisions, the speed with which an identification decision is made becomes a less useful indicator for estimating the decision's accuracy.

PRACTICAL CONSIDERATIONS FOR CONDUCTING AN EVALUATION OF EYEWITNESS TESTIMONY

So far in this chapter we have provided a framework for evaluating how and why some factors may influence the reliability of eyewitness testimony. We now turn our attention to addressing some of the important practical considerations when evaluating an eyewitness's identification or testimony.

Expert Training and Role. Anyone attempting to evaluate eyewitness identification from a psychological perspective should have training in the methods and theory of experimental psychology and should be intimately familiar with the empirical literature on eyewitnesses. Given that the scientific account of what happens during the perception of objects and events, the formation of memory, and the processes through which response decisions are made has become very detailed, this knowledge has become increasingly technical and unavailable to those without the requisite background. However, an expert with a current understanding of the basic theories of perception, memory, and decision making will likely have the foundational knowledge necessary to undertake this task. This understanding, coupled with comprehensive knowledge of the research on eyewitness testimony and the specific tools developed to assess the quality of eyewitness memory tests, provides access to the tools required to assess the eyewitness evidence and identify factors that may influence the accuracy of witness testimony. In addition, the expert should have published empirical research on cognitive processes that are relevant to the issues influencing eyewitness memory and decision making, identification procedures, and/or testimony.

Provided the expert has the requisite training, knowledge, and skills to evaluate eyewitness testimony adequately, the question then arises as to how this knowledge should be used to inform the judgment of the legal system. There are at least three models of how experts may influence the legal system's evaluation of eyewitness testimony.

In one model, the evaluator constructs judgments about the likely accuracy of the witness's testimony, perhaps even attempting to assess the witness's memory. There are several problems with such an approach—not the least of which is that it invades the province of the finder of fact. That is, the expert would be giving ultimate opinion testimony, providing a conclusion that is actually the task of the jury. Such testimony by a psychological expert would readily be seen as inappropriate and potentially prejudicial.

The testimony of an eyewitness to a criminal event is conditioned by many influences associated with the environment of a unique event and by the actions of many people, both as event actors and as actors representing the criminal

justice system and process. Although the accuracy of the witness's statements and identification decision(s) are the ultimate questions for the fact finder, the interpretation of these statements is influenced by many limiting factors for which tests of memory are not pertinent. Often no other witnesses were present during the criminal event. In addition, after the criminal event, trained personnel typically are not present to collect observations and technical measurements. First responders, who arrive after some delay, may not have requisite technical training or experience in what to look for and what to record and preserve. Often their reports are written from memory, perhaps supplemented by notes, after a delay of hours or more. In addition, first responders often have other competing investigative roles and responsibilities that may interfere with their ability to adequately, and accurately, collect eyewitness evidence (e.g., securing the crime scene, providing support to medical personnel, etc.). However, the information needed to arrive at a useful estimate of the accuracy of the witness's statements necessarily would be based on information that contains significant detail and complexity, is based on technical expertise rather than common knowledge, and reaches a plausible level of internal integration. Given the limitations described, such an information base seems unlikely to be readily available.

In another model, there is an important differentiation of tasks between the finders of fact in the legal system (judges, juries) and the eyewitness expert. It is the jury's task to come to a judgment of the likelihood of guilt, given the evidence, and then evaluate whether that judgment surpasses or sets aside reasonable doubt. As previously discussed, this is not the role of the expert. Rather, the expert's role is dramatically different. It is to explain to the fact finder the science of eyewitness testimony as it relates to the case at bar. The scientist's task, then, is to set before the jury the scientific knowledge on the relevant aspects of the case so that the jury may use this information as it sees fit.

An alternative approach is to consider the relevant factors found in the science, such as the matters reviewed earlier, but to ask an unusual question: Could these events and factors produce an identification of the defendant given that the defendant is innocent? Put another way, is guilt of the defendant required to achieve an identification under the conditions of the present case, or could an identification under these same conditions equally well result in identification of an innocent person? This is essentially asking the expert to give an opinion as to whether the science finds that the presence or absence of each factor in the case at bar contributes to an identification decision that would render identification of an innocent person as likely as that of the guilty person?

This approach may be characterized as an assessment of the comparative likelihood of an identification given innocence or guilt. Of course, one must to some degree be interested in its accuracy. But the absence of a criterion against which to compare and the absence of information about aspects of the witness, the event, and the experiences of the witness in his or her interactions with law enforcement and the court system make the task difficult. And in any case, that particular

question of accuracy is what the trial process is about. Evaluation of testimony for psychological experts takes on a different form, as described by Wagenaar (1988). Expert testimony of the form “the defendant is more likely to be guilty because s/he was identified in a lineup,” focusing as it does on the defendant, borders on ultimate opinion testimony and can be viewed as an invasion of the province of the jury. However, expert testimony in the form of “the defendant is as likely to have been identified in this lineup given innocence as given guilt” reflects a focus on the factors that influence identification, the presence of those factors that support low or high identification decision criterion, memory contamination, and other false identification risks.

What is the goal of evaluating eyewitness testimony? We agree with Wagenaar (1988) that it is to use the best scientific knowledge available to assist the jury in deciding the facts of the case and the weight jury members will give to the evidence provided by eyewitnesses.

What Is Evidence and What Is Not? We know that some things offered as “evidence” of identity are so likely to yield errors of identification that they should be considered high risk and identified as such to the jury. If offered at trial, this evidence certainly should be countered with expert testimony. Contained in this category is evidence based on:

- the five *Manson* criteria for deciding the presence of suggestion in an identification procedure (opportunity to view, attention, description, time to identification, and certainty; see Wells & Quinlivan, 2009);
- composite images (either computer based or by sketch artist);
- hypnosis or memory refreshing;
- witness confidence (even if assessed at the time of an otherwise excellent identification procedure);
- an in-court identification and the related claim of “independent source”; and
- identifications made of a suspect when the offender wore a disguise.

In addition, identifications made based on viewing a surveillance video are also highly likely to be made in error (Davis & Valentine, 2009; Davis, Valentine, & Davis, 2010). Some of these identifications are offered based solely on the basis of folk beliefs and have essentially no support in the science for their use in an eyewitness identification context.

Other matters are of interest from both psychological and procedural justice perspectives. For example, often the only evidence available comes from cryptic notes made by police personnel a long time after the events being reported occurred. These notes fail to detail how the witness was instructed, whether and how witnesses were separated when they made their identifications, what questions were asked of witnesses, and any qualification the witnesses may have made in response (“Well, I don’t really know. It might be number 3, but I’m not sure”). Department

policies often require recording. What are the sanctions when department policies are ignored? The statements psychologists make when evaluating testimony may be equivocal (e.g., “leaves many important questions unanswered”). But an expert in a consulting role can explain the history of the problem of poor documentation, reference the rules that have been put into place by various law enforcement and nongovernmental organizations and their basis in good practice, and assist with the formulation of the psychological science portion of motions to suppress.

Sources of Information for the Expert. It is crucial that the expert have access to all materials that are relevant to the collection of eyewitness evidence during the police investigation. Without this information, any opinions regarding the validity of the eyewitness evidence will be incomplete and potentially unreliable. The material experts should seek includes description(s) of the perpetrators given by witnesses and statements pertaining to other aspects of the witnessed event, from the time of the criminal event through the pretrial processes (e.g., 911 recordings, police reports, testimony during depositions and other pretrial hearings). These statements typically will be available in police reports; however, they may also come from interactions with the press (e.g., in the case of a media interview) and through postings on social media sites (e.g., Facebook, blogs, discussion boards, etc.). It is of vital importance to acquire documentation of any identification tests that were used during the investigation, including any documents provided to the witnesses and information about the specific procedures employed by the investigators (e.g., mug shot searches, showups, composites produced, pre-lineup instructions, photoarrays/lineups used, hypnosis/memory refreshing techniques, etc.) as well as any potential exposure the witness may have had to additional case information in the media (e.g., reports about potential suspects, photographs, composite sketches, etc.).

An important aspect of case materials often overlooked by attorneys is whether witnesses participated in identification procedures that did not result in an identification of the suspect. Law enforcement officers often view nonidentifications as uninformative and therefore exclude this information from the case materials provided to the prosecuting attorney’s office and defense counsel. However, for the consulting expert, evidence of nonidentifications is important to consider: Non-identifications decrease the extent to which identifications by other witnesses are diagnostic of the suspect’s guilt (see Clark & Wells, 2008, for a detailed discussion). Therefore, the evaluator should have access to all case material related to every identification procedure, regardless of its outcome.

Unfortunately, comprehensive records are often elusive. In some cases, this information is not available, due to an absence of records or a lack of verbatim recording of the witness’s statements. However, the expert should never assume that the information provided by the attorney is all that there is to be had. Rather, attorneys are often unaware of what information is relevant. Thus, experts should be

explicit with the request for this information and encourage the retaining attorney to pursue any and all documentation of the eyewitness evidence from all witnesses and all suspects in the case. This is important, as the procedures used on one witness may influence another witness in the case, especially if the witnesses came into contact with one another.

Once comprehensive records of the witness's statements and the case investigative methods are gathered, the expert can begin examining this material to identify factors that may influence the reliability of the witness's testimony. This examination should rely on the scientific literature on the study of human perception, memory, and decision making that we discussed earlier, in addition to the particular techniques and concepts developed in that literature.

APPLYING WHAT WE KNOW (USING THE NORTHROP/DAVIS CASE)

The wrongful convictions of Alan Northrop and Larry Davis provide us with the opportunity to present how an expert could go about assessing eyewitness testimony in a criminal case. However, two caveats are in order. First, hindsight bias affords us the ability to look back on this case and see the "obvious" signs of eyewitness error, but this is a somewhat immodest suggestion about the general perspicacity of experts. We use this case as an example of what types of information an evaluator should seek and where it may come from, but it is, of course, not always easy to identify potential sources of error. Second, we evaluated these materials only recently—well after Northrop and Davis were exonerated. None of the authors was involved with this case at any stage of the legal process. Rather, these case materials were provided by attorneys who represented Northrop and Davis in postconviction processes. We do not have all the relevant information on these cases. Our evaluation is based on two legal documents—the *Motion and Memorandum for a New Trial* (Innocence Project Northwest, 2010) and the subsequent *Findings of Facts and Conclusions of Law and Order* (i.e., the court decision to overturn conviction and grant the motion for new trial; *Washington v. Davis/Northrop*, 2010), images used during the criminal investigation (i.e., a composite sketch, photo montages, and a photograph of the Northrop live lineup), and conversations with the postconviction attorneys and exonerees.

We began our evaluation by thoroughly examining the records that were made available to us for any information that may be relevant to the eyewitness's memory of the event. Rather than creating a list of factors that might or might not have been present, we developed an "eyewitness timeline" that set out a chronological sequence of events and reports. Constructing a detailed timeline is imperative, as this information generally serves as the data for subsequent evaluation. It is with the assistance of this timeline that the evaluator can identify the sequence of events and when information was reported.

Using this timeline as a starting point, we searched for information germane to the many issues we have discussed in this chapter. As with any case, many of the issues

that can influence eyewitness memory were not present in this case (e.g., there were no cross-race, age, disguise, or alcohol intoxication issues), and we were unable to determine whether many other issues were present due to the limited information in the reports we were provided (e.g., witness motives/beliefs, stress, interviewing techniques, and postevent information). However, we identified many aspects of this case that could have contributed to the errors made by the eyewitness and about which we would have proffered expert testimony during the trial.

FACTORS RELATED TO INFORMATIONAL PROCESSES

As we discussed earlier, a useful framework for evaluating eyewitness testimony is to assess the extent to which the factors influenced the informational and/or decisional processes associated with eyewitness identification and testimony. In this section, we consider some of the factors that could have influenced the informational processes of the witness in the Davis/Northrop case.

Composites. Within a few days of the crime, the witness was asked to assist with developing a composite portrait of the perpetrators. Although the victim stated that she did not get a good enough look at the blond assailant to assist with the composite, she was willing to help the sketch artist produce an image of her second assailant. This composite set the ball rolling toward the wrongful convictions of Northrop and Davis, as Alan Northrop became a suspect primarily because he resembled the sketch (see Figure 17.6). As discussed earlier, composite portraits rarely look much like the perpetrator. Witnesses have significant difficulty constructing composites of familiar individuals, let alone complete strangers observed for a short period

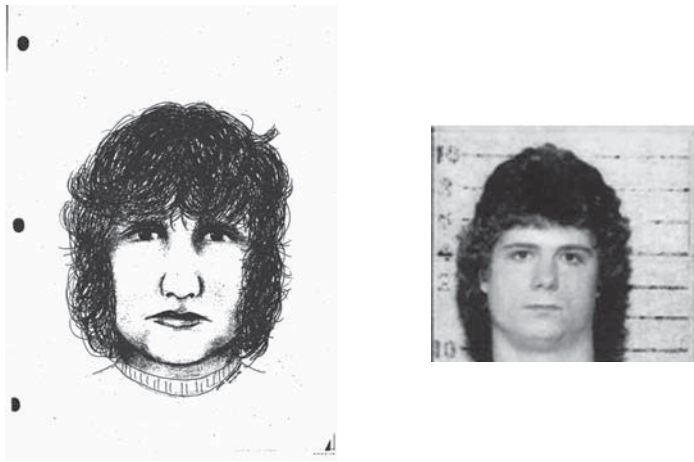


Figure 17.6 Composite Sketch Created by the Victim in the Davis/Northrop Case (left) and Police Mug Shot of Alan Northrop (right)

of time (Davies & Valentine, 2007). Considering also that the victim could not construct a nose, resulting in the sketch artist including “a nose that fits with [the] face,” this potentially misleading piece of information may have influenced the witness’s memory for the perpetrator (cf. Jenkins & Davies, 1985). The evidence suggests that investigators developed a composite that probably did not look much like the perpetrator but was used as an investigative tool for identifying suspects, nevertheless. Indeed, Alan Northrop was identified as a suspect because community members thought he looked similar to the composite. In addition, and more problematically, the process of constructing the composite and the incorporation of a “nose that fits” may have altered the victim’s memory for the perpetrator.

Nonidentifications. The victim viewed two photo montages, one containing Alan Northrop and one containing Larry Davis. The witness did not identify anyone from the Northrop montage. Rather than considering nonidentification as evidence of innocence, the investigators continued to search into Northrop’s potential involvement. In addition, a discrepancy exists regarding whether the victim positively identified Davis from his photo montage. Although the investigator’s report states that the witness evaluated the montage and stated “that’s not him, but that’s the one” while pointing to the photograph of Larry Davis (Innocence Project Northwest, 2010, p. 10), the witness later testified that she thought Davis may have been the man who assaulted her, but she could not identify his face, rather saying that his neck looked familiar and he seemed familiar to her. This tentative response was interpreted by law enforcement as a positive identification.

Lineup Composition. The narrative reports did not provide any information regarding how the lineups were constructed. Investigators are encouraged to evaluate a lineup prior to administering it and to determine whether the suspect stands out for any reason (Technical Working Group on Eyewitness Evidence, 1999). Often this is done with an eyeball test—the investigator relying on his or her subjective evaluation. However, this subjective assessment is problematic due to the unreliability of evaluations across individuals. For example, one could examine the Northrop montage and conclude that there is no evidence of bias, since all individuals are young males with brown hair, whereas another may conclude that individual #6 stands out because he is the only one who has an open mouth. In addition, another expert may examine the Davis montage and perceive this lineup as being potentially biased, because Larry Davis is the only individual without a full mustache. Rather than rely on these subjective assessments, we conducted a mock witness evaluation of the photo montages used in this case similar to the approach described earlier.

A total of 34 individuals participated in the evaluation. Each participant evaluated both the Davis and Northrop photo montages; however, the order of evaluation was counterbalanced. Following standard mock witness procedures, participants

were given the description of each perpetrator that was provided by the victim. In addition, participants were shown the composite sketch of the perpetrator prior to viewing the Northrop montage.

As discussed earlier, the mock witness procedure affords the ability to assess two components related to lineup fairness: lineup size (the number of plausible fillers in the lineup) and lineup bias (the extent to which the suspect stands out amongst those plausible fillers). Figure 17.7 shows the identification frequencies for each member of the photo montages. Evaluation of the adequacy of the fillers indicated that not all members fully served their protective purpose. Although both montages contain six individuals, the effective sizes are significantly smaller than this nominal size (*Tredoux's* $E'_{Northrop} = 2.18$, 95% CI: 1.59–3.50; *Tredoux's* $E'_{Davis} = 2.86$, 95% CI: 2.15–4.26), indicating that both montages had low effective sizes. We then assessed the extent to which the montages were biased toward Northrop and Davis. These analyses indicated that the Northrop photo montage was biased toward Northrop, with 65% of the mock witnesses choosing him from the montage (95% CI: .49–.81; chance expectation = .166). In the second lineup, half of the mock witnesses selected Davis from the Davis photo montage (95% CI: .33–.67; chance expectation = .166). Based on these analyses, we concluded that both the Northrop and Davis photo montages demonstrate low effective lineup size *and* high bias against the defendants.

Repeated Procedures. The victim in this case participated in multiple identification procedures for both Alan Northrop and Larry Davis. The first identification procedures for both men were separate photo montages during which the victim failed to clearly and positively identify either man as a perpetrator. However, the victim was also shown two separate live lineups (one containing Alan Northrop and one containing Larry Davis) in which the only individuals who were the same in both the photo montages and live lineups were the suspects. These live lineups resulted in positive identifications of both Davis and Northrop. As discussed earlier, repeated identification procedures are problematic, because it becomes difficult to determine whether the witness's subsequent identification is based on memory for the perpetrators or because the witness remembers these individuals from prior identification procedures.

FACTORS RELATED TO DECISIONAL PROCESSES

In the previous section we considered factors that may have affected the quality of the information on which the witness in the Davis/Northrop case based her decision. However, as we made clear earlier in the chapter, when a witness makes a decision it is not only the quality of information at his or her disposal, but also a set of factors that affect the willingness to make a decision. We discuss these "decisional factors" that were present in the Davis/Northrop case in this section.

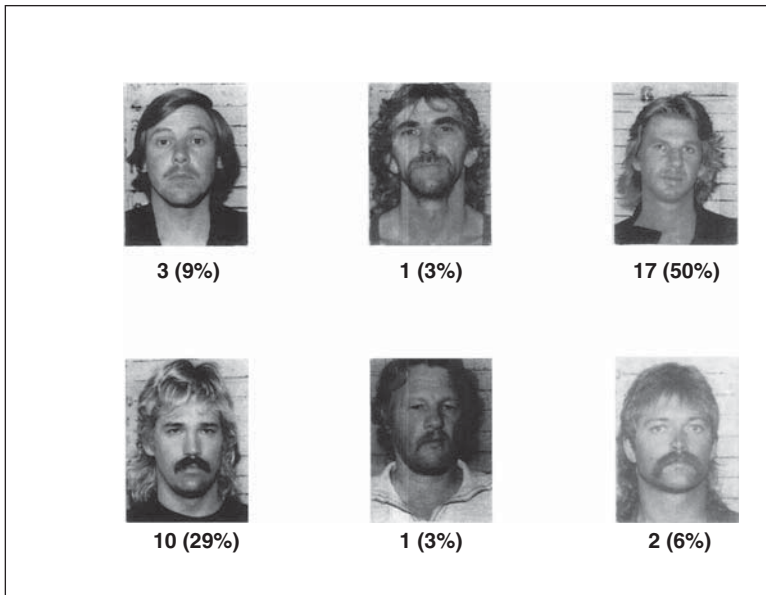
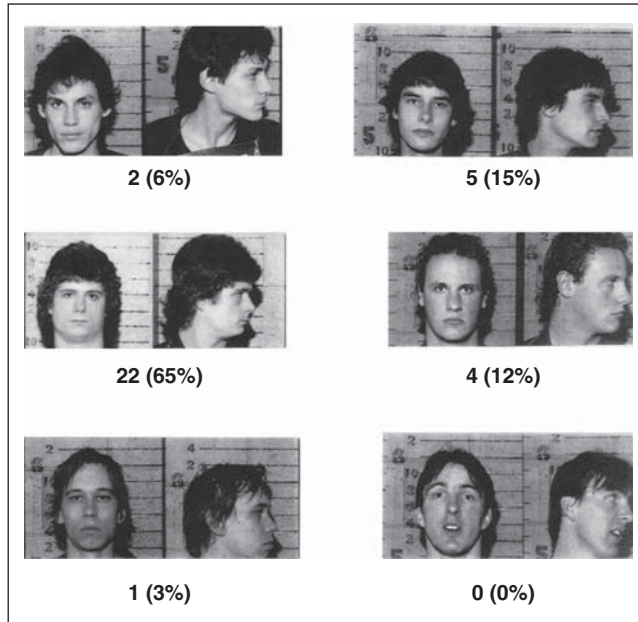


Figure 17.7 Results of the Mock Witness Evaluations for the Northrop and Davis Photo Montages Choice frequencies (percentages in parentheses) are provided for each lineup member.

Witness Instructions/Expectations. Prior to the live lineups, the victim was contacted by a friend who informed her that there were suspects in the case, that one of the suspects had been arrested, and provided her the names of the suspects. The victim then called the jail, confirmed the arrest of Larry Davis, and deduced that he must have been the blond perpetrator since he was arrested for burglary and not sexual assault. A short time later, investigators called the victim and informed her that they had arrested someone and may be doing a lineup. After the Davis live lineup, the victim was informed by investigators that there was another suspect in the case, that the suspect might agree to participate in a live lineup, and if he did agree, she would be contacted to come down to the station. Later that day, the victim was called down to the station for the Northrop live lineup. As a result of this information, the victim went into both of the live lineups with the expectation that someone in the lineup was going to be the suspect. In addition, the victim was also informed prior to the Northrop live lineup that the reason the suspect had not already been arrested was that she had failed to pick him out in the photo lineup she saw earlier. Therefore, the victim likely went into this live lineup procedure with the mind-set (a) that a suspect is present and (b) to look for someone whom she had seen in the photo lineup. As discussed earlier, this type of pre-lineup information is extremely suggestive and increases the likelihood of a witness choosing someone, regardless of guilt or innocence, simply due to the expectation that the police have their man.

SUMMARY AND CONCLUSIONS

Forensic psychology is first and foremost an adventure in applied psychology. However, as Belbin (1979) pointed out, very often what goes by the name applied psychology is at best only applicable psychology. Findings made under the name of applied psychology are generally not applied—the label is assumed because of a concern with issues or problems in society, but to do anything that matters is extraordinarily difficult. There are elevated evidentiary standards (the research must pass particularly stringent tests of internal and external validity, since the welfare of the public may be at stake), and it is contentious whether researchers have any standing in the public policy domain.

In the case of research on eyewitnesses, however, psychologists have had a measure of success, and we have reported various aspects of this work here. The contributions we can make to combating mistaken eyewitness identification and its grievous misdirections of criminal justice are at once both practical and theoretical. We have shown in this chapter how an expert can assess the reliability of an eyewitness's identification based on knowledge of the expanding research literature and some of the tools that have emerged from that literature, including those for assessing the fairness of police lineups. We have also argued throughout the chapter that experts and law enforcement officials can all profit considerably by thinking about an eyewitness's testimony as being the product of two independent

processes. On one hand, the eyewitness might have a sound memory of an event or of a perpetrator's physical features. The memory may be detailed, and the witness might be able to correctly recall many aspects of the event and be able to recognize the perpetrator in a lineup containing fillers who closely resemble him. In other words, the *informational processes* may have been sound and have resulted in a stable and well-elaborated memory.

However, the witness might fail to deliver testimony to the police or the court that does justice to the quality of the information that he or she has about the event and the perpetrator(s). We have discussed research, for instance, that shows how important the social nature of the identification task is to the reliability of the identification and how suggestiveness can enter the investigative process at multiple points (the composite construction task, repeat viewings of mug shots and photospreads, the lineup itself, the interview process, and many others). These *decisional processes* influence a witness's testimony just as surely as the experience of the original event does. In particular, the interaction between *informational* and *decisional* processes is best understood through signal detection theory. The implications of using SDT are multifold, and researchers in the field are making important discoveries, starkly illuminated recently by Clark's (2012) evidence that alterations to the way police collect eyewitness evidence and identifications involve a trade-off between fewer identifications of innocent suspects and fewer identifications of guilty suspects. There is little doubt in our minds that the conceptual framework afforded by SDT offers an opportunity to make eyewitness research considerably more rigorous and to augment its considerable extant achievements.

REFERENCES

- Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin and Review*, *12*, 1043–1047.
- Areh, I. (2011). Gender-related differences in eyewitness testimony. *Personality and Individual Differences*, *50*, 559–563.
- Belbin, E. (1979). Applicable psychology and some national problems. *British Journal of Psychology*, *70*, 187–197.
- Brewer, N., Caon, A., Todd, C., & Weber, N. (2006). Eyewitness identification accuracy and response latency. *Law and Human Behavior*, *30*, 31–50. doi:10.1007/s10979-006-9002-7
- Brewer, N., & Weber, N. (2008). Eyewitness confidence and latency: Indices of memory processes not just markers of accuracy. *Applied Cognitive Psychology*, *22*, 827–840. doi:10.1002/acp.1486
- Brewer, N., Weber, N., Wootton, D., & Lindsay, D. S. (2012). Identifying the bad guy in a lineup: Using confidence judgments under deadline pressure. *Psychological Science*, *23*, 1208–1214. doi:10.1177/0956797612441217
- Chiroro, P. M., Tredoux, C. G., Radaelli, S., & Meissner, C. A. (2008). Recognizing faces across continents: The effect of within-race variations on the own-race bias in face recognition. *Psychonomic Bulletin and Review*, *15*, 1089–1092. doi:10.3758/PBR.15.6.1089
- Clark, S. E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior*, *29*, 395–424.

- Clark, S. E. (2012). Costs and benefits of eyewitness identification reform: Psychological science and public policy. *Perspectives on Psychological Science, 7*, 238–259.
- Clark, S. E., & Wells, G. L. (2008). On the diagnosticity of multiple-witness identifications. *Law and Human Behavior, 32*, 406–422. doi:10.1007/s10979-007-0115-7
- Davies, G., & Valentine, T. (2007). Facial composites: Forensic utility and psychological research. In R.C.L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *The handbook of eyewitness psychology, Vol. 2, Memory for people* (pp. 59–83). Mahwah, NJ: Erlbaum.
- Davis, J. P., & Valentine, T. (2009). CCTV on trial: Matching video images with the defendant in the dock. *Applied Cognitive Psychology, 23*, 482–505. doi:10.1002/acp.1490
- Davis, J. P., Valentine, T. Y., & Davis, R. E. (2010). Computer assisted photo-anthropometric analyses of full-face and profile facial images. *Forensic Science International, 200*, 165–176. doi:10.1016/j.forsciint.2010.04.012
- deJong, M., Wagenaar, W. A., Wolters, G., & Verstijnen, I. M. (2005). Familiar face recognition as a function of distance and illumination: A practical tool for use in the courtroom. *Psychology, Crime and Law, 11*, 87–97.
- Deffenbacher, K. A. (1994). Effects of arousal on everyday memory. *Human Performance, 7*, 141–161.
- Deffenbacher, K. A., Bornstein, B. H., & Penrod, S. D. (2006). Mugshot exposure effects: Retroactive interference, mugshot commitment, source confusion, and unconscious transference. *Law and Human Behavior, 30*, 287–307. doi:10.1007/s10979-006-9008-1
- Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior, 28*, 687–706.
- Devlin, Patrick. (1976). *Report to the Secretary of State for the Home Department of the Departmental Committee on Evidence of Identification in Criminal Cases*. London, England: Her Majesty's Stationery Office.
- Doob, A. N., & Kirshenbaum, H. M. (1973). Bias in police lineups—partial remembering. *Journal of Police Science and Administration, 18*, 287–293.
- Dunning, D., & Perretta, S. (2002). Automaticity and eyewitness accuracy—A 10- to 12-second rule for distinguishing accurate from inaccurate positive identifications. *Journal of Applied Psychology, 87*, 951–962.
- Dysart, J. E., & Lindsay, R. C. L. (2007). The effects of delay on eyewitness identification accuracy: Should we be concerned? In R. C. L. Lindsay, D. R. Ross, J. D. Read, & M. P. Toglia (Eds.), *The handbook of eyewitness psychology, Vol. 2, Memory for people* (pp. 361–376). Mahwah, NJ: Erlbaum.
- Ebbinghaus, H. (1962). *Memory: A contribution to experimental psychology*. New York, NY: Dover. (Originally published 1885).
- Ellison, K. W., & Buckhout, R. (1981). *Psychology and criminal justice*. New York, NY: Harper & Row.
- Erdelyi, M. H. (2010). The ups and downs of memory. *American Psychologist, 65*, 623–633. doi:10.1037/a0020440
- Evans, J. R., Schreiber Compo, N., & Russano, M. (2009). Intoxicated witnesses and suspects: Procedures and prevalence according to law enforcement. *Psychology, Public Policy, and Law, 15*, 194–221. doi:10.1037/a0016837

- Fawcett, J. M., Russell, E. J., Peace, K. A., & Christie, J. (2013). Of guns and geese: A meta-analytic review of the "weapon focus" literature. *Psychology, Crime and Law, 19*, 35–66. doi:10.1080/1068315X.2011.599325
- Fisher, R. P., Brewer, N., & Mitchell, G. (2009). The relation between consistency and accuracy of eyewitness testimony: Legal versus cognitive explanations. In R. Bull, T. Valentine, & T. Williamson (Eds.), *Handbook of psychology of investigative interviewing: Current developments and future directions* (pp. 121–136). Chichester, England: Wiley-Blackwell.
- Fisher, R. P., Ross, S. J., & Cahill, B. S. (2010). Interviewing witnesses and victims. In P. A. Granhag (Ed.), *Forensic psychology in context: Nordic and international approaches* (pp. 56–74). Devon, England: Willan.
- Fisher, R. P., & Schreiber, N. (2007). Interviewing protocols for improving eyewitness memory. In M.P. Toglia, J. D. Read, D. F. Ross, & R.C.L. Lindsay (Eds.), *The handbook of eyewitness psychology, Vol. 1, Memory for events* (pp. 53–80). Mahwah, NJ: Erlbaum.
- Frowd, C. D., Hancock, P. J., & Carson, D. (2004). Evo-fit: A holistic, evolutionary facial imaging technique for creating composites. *ACM Transactions on Applied Perception, 1*, 219–239.
- Garrett, B. L. (2012). Eyewitnesses and exclusions. *Vanderbilt Law Review, 65*, 451–506.
- Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Krüger, L. (1989). *The empire of chance: How probability changed science and everyday life*. Cambridge, England: Cambridge University Press.
- Greene, E., & Fraser, S. (2002). Observation distance and recognition of photographs of celebrities' faces. *Perceptual and Motor Skills, 95*, 637–651.
- Greathouse, S. M., & Kovera, M. B. (2009). Instruction bias and lineup presentation moderate the effects of administrator knowledge on eyewitness identification. *Law and Human Behavior, 33*, 70–82.
- Gronlund, S. D., Carlson, C. A., Dailey, S. B., & Goodsell, C. A. (2009). Robustness of the sequential lineup advantage. *Journal of Experimental Psychology: Applied, 15*, 140–152. doi:10.1037/a0015082
- Gronlund, S. D., Carlson, C. A., Neuschatz, J. S., Goodsell, C. A., Wetmore, S. A., Wooten, A., & Graham, M. (2012). Showups versus lineups—An evaluation using ROC analysis. *Journal of Applied Research in Memory and Cognition, 1*, 221–228. doi:10.1016/j.jarmac.2012.09.003
- Gross, S. R., & Shaffer, M. (2012). Exonerations in the United States, 1989–2012. *Public Law and Legal Theory Working Paper Series*, University of Michigan Law School. Retrieved from <http://ssrn.com/abstract=2092195>
- Hassin, R., & Trope, Y. (2000). Facing faces: Studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology, 78*, 837–852.
- Haw, R. M., Dickinson, J. J., & Meissner, C. A. (2007). The phenomenology of carryover effects between show-up and line-up identification. *Memory, 15*, 117–127. doi:10.1080/09658210601171672
- Hilliar, K. F., Kemp, R. I., & Denson, T. F. (2010). Now everyone looks the same: Alcohol intoxication reduces the own-race bias in face recognition. *Law and Human Behavior, 34*, 367–368. doi:10.1007/s10979-009-9204-x
- Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review, 117*, 1168–1187. doi:10.1037/a0020463

- Innocence Project. (n.d.). The Innocence Project. Retrieved from <http://www.innocenceproject.org>
- Innocence Project Northwest. (2010). Unpublished motion and memorandum for a new trial in case of Washington v. Davis/Northrop.
- Jenkins, F., & Davies, G. (1985). Contamination of facial memory through exposure to misleading composite pictures. *Journal of Applied Psychology, 70*, 164–176.
- Lindsay, R., Semmler, C., Weber, N., Brewer, N., & Lindsay, M. (2008). How variations in distance affect eyewitness reports and identification accuracy. *Law and Human Behavior, 32*, 526–535.
- Lindsay, R. C. L., & Wells, G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentations. *Journal of Applied Psychology, 70*, 556–564.
- Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning and Memory, 12*, 361–366.
- Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 19–31.
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13*, 585–589.
- Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals, 25*, 720–725.
- Luus, C.A.E., & Wells, G. L. (1994). The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology, 79*, 714–723.
- MacLin, O. H., & MacLin, M. K. (2004). The effect of criminality on face attractiveness, typicality, memorability and recognition. *North American Journal of Psychology, 6*, 145–154.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.
- Malpass, R. S. (1981). Effective size and defendant bias in eyewitness identification lineups. *Law and Human Behavior, 5*, 299–309.
- Malpass, R. S., & Devine, P. G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology, 66*, 482–489.
- Malpass, R. S., & Devine, P. G. (1983). Measuring the fairness of eyewitness identification lineups. In S. M. Lloyd-Bostock & B. R. Clifford (Eds.), *Evaluating witness evidence* (pp. 81–102). London, England: Wiley.
- Malpass, R. S., & Devine, P. G. (1984). Research on suggestion in lineups and photospreads. In G. L. Wells & E. F. Loftus (Eds.), *Eyewitness testimony: Psychological perspectives* (pp. 64–91). New York, NY: Cambridge University Press.
- Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other "race." *Journal of Personality and Social Psychology, 13*, 330–334.
- Malpass, R. S., Ross, S. J., Meissner, C. A. & Marcon, J. L. (2009). The need for expert psychological testimony on eyewitness identification. In B. L. Cutler (Ed.) *Expert testimony on the psychology of eyewitness identification* (pp. 3–28). New York, NY: Oxford University Press.
- Malpass, R. S., Susa, K. J., & Meissner, C. A. (2008). Training of eyewitnesses. In B. L. Cutler (Ed.), *Encyclopedia of psychology and law* (pp. 807–808). Thousand Oaks, CA: Sage.

- Manson v. Braithwaite. 432 U.S. 98 (1977).
- Maylor, E. A., Rabbitt, P. M., & Kingstone, A. (1987). Effects of alcohol on word categorization and recognition memory. *British Journal of Psychology*, *78*, 233–239.
- Mazzoni, G., & Lynn, S. J. (2007). Using hypnosis in eyewitness–memory: Past and current issues. In M. P. Toglia, J. D. Read, D. F. Ross, & R.C.L. Lindsay (Eds.) *The handbook of eyewitness psychology, Vol. 1, Memory for events* (pp. 321–338). Mahwah, NJ: Erlbaum.
- McKelvie, S. J. (1976). The role of eyes and mouth in the memory of a face. *American Journal of Psychology*, *89*, 311–323.
- McQuiston-Surrett, D., Malpass, R. S., & Tredoux, C. G. (2006). Sequential vs. simultaneous lineups: A review of methods, data, and theory. *Psychology, Public Policy, and Law*, *12*, 137–169.
- Meissner, C. A., & Brigham, J. C. (2001a). A meta-analysis of the verbal overshadowing effect in face identification. *Applied Cognitive Psychology*, *15*, 603–616.
- Meissner, C. A., & Brigham, J. C. (2001b). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, *7*, 3–35.
- Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. *Memory and Cognition*, *33*, 783–792.
- Memon, A., Meissner, C. A., & Fraser, J. (2010). The cognitive interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law*, *16*, 340–372. doi:10.1037/a0020518
- Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, *67*, 371.
- Münsterberg, H. (1908). *On the witness stand: Essays on psychology and crime*. New York, NY: Doubleday.
- Noor, F., & Evans, D. C. (2003). The effect of facial symmetry on perceptions of personality and attractiveness. *Journal of Research in Personality*, *37*, 339–347.
- Oeberst, A. (2012). If anything else comes to mind...better keep it to yourself?: Delayed recall is discrediting—unjustifiably. *Law and Human Behavior*, *36*, 266–274. doi:10.1037/h0093966
- Parker, J. F., & Ryan, V. (1993). An attempt to reduce guessing behavior in children’s and adults’ eyewitness identifications. *Law and Human Behavior*, *17*, 11–26.
- Pezdek, K., & Blandon-Gitlin, I. (2005). When is an intervening line-up most likely to affect eyewitness identification accuracy? *Legal and Criminological Psychology*, *10*, 247–263. doi:10.1348/135532505X49846
- Pelwani v. S. 1963 (2) (PH) H237 (T). (South Africa)
- Pickel, K. L., Ross, S. J., & Truelove, R. S. (2006). Do weapons automatically capture attention? *Applied Cognitive Psychology*, *20*, 871–893.
- Quinlivan, D. S., Neuschatz, J. S., Cutler, B. L., Wells, G. L., McClung, J., & Harker, D. L. (2012). Do pre-admonition suggestions moderate the effect of unbiased lineup instructions? *Legal and Criminological Psychology*, *17*, 165–176.
- Ranganathan, M., & D’Souza, D. C. (2006). The acute effects of cannabinoids on memory in humans: A review. *Psychopharmacology*, *188*, 425–444.
- Rehman, J., & Herlitz, A. (2007). Women remember more faces than men do. *Acta Psychologica*, *124*, 344–355.

- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom*. New York, NY: Holt, Rinehart & Winston.
- Ross, D. F., Ceci, S. J., Dunning, D., & Tolia, M. P. (1994). Unconscious transference and lineup identification: Toward a memory blending approach. In D. F. Ross, J. D. Read, & M. P. Tolia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 80–100). New York, NY: Cambridge University Press.
- Russano, M. B., Dickinson, J. J., Greathouse, S. M., & Kovera, M. B. (2006). Why don't you take another look at number three? Investigator knowledge and its effects on eyewitness confidence and identification decisions. *Cardozo Public Law, Policy, and Ethics Journal*, 4, 355–379.
- Sauerland, M., & Sporer, S. L. (2009). Fast and confident: Postdicting eyewitness identification accuracy in a field study. *Journal of Experimental Psychology: Applied*, 15, 46–62. doi:10.1037/a0014560
- Schreiber Compo, N., Evans, J. R., Carol, R. N., Villalba, D., Ham, L. S., Garcia, T., & Rose, S. (2011). Intoxicated eyewitnesses: Better than their reputation? *Law and Human Behavior*, 36, 77–86. doi:10.1037/h0093951
- Shepherd, J., Gibling, F., & Ellis, H. D. (1991). The effects of distinctiveness, presentation time and delay on face recognition. *European Journal of Cognitive Psychology*, 3, 137–145.
- Sobel, N. R. (2007). *Eyewitness identification: Legal and practical problems*. Eagan, MN: Thompson/West.
- Sporer, S. L. (1992, March). *An archival analysis of person descriptions*. Paper presented at the biennial meeting of the American Psychology–Law Society. San Diego, CA.
- Sporer, S. L. (2001). Eyewitness identification: Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law*, 7, 1–45.
- Sporer, S. L., Penrod, S., Read, R., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin*, 118, 315–327.
- Stebly, N., Dysart, J., Fulero, S., & Lindsay, R.C.L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentation: A meta-analytic comparison. *Law and Human Behavior*, 25, 459–473.
- Stebly, N., Dysart, J., Fulero, S., & Lindsay, R.C.L. (2003). Eyewitness accuracy rates in police showup and lineup presentations: A meta-analytic comparison. *Law and Human Behavior*, 27, 523–540.
- Stebly, N., Dysart, J., & Wells, G. L. (2011). Seventy-two tests of the sequential lineup superiority effect: A meta-analysis and policy discussion. *Psychology, Public Policy, and Law*, 17, 99–139.
- Stebly, N. K., Dietrich, H. L., Ryan, S. L., Raczynski, J. L., & James, K. A. (2011). Sequential lineup laps and eyewitness accuracy. *Law and Human Behavior*, 35, 262–274. doi:10.1007/s10979-010-9236-2
- Steele, C. M., & Josephs, R. A. (1990). Alcohol myopia: Its prized and dangerous effects. *American Psychologist*, 45, 921–933. doi:10.1037/0003-066X.45.8.921
- Technical Working Group on Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement*. Washington, DC: National Institute of Justice.
- Tredoux, C. G. (1998). Statistical inference on measures of lineup fairness. *Law and Human Behavior*, 22, 217–237.

- Tredoux, C. G., Nunez, D. T., Oxtoby, O., & Prag, B. (2007). An evaluation of ID: An eigenface based construction system. *South African Computer Journal*, 37, 90–97.
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5, 381–391.
- Valentine, T., & Mesout, J. (2009). Eyewitness identification under stress in the London dungeon. *Applied Cognitive Psychology*, 23, 151–161. doi:10.1002/acp.1463
- Valentine, T., Pickering, A., & Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology*, 17, 969–993.
- Wade, K. A., Garry, M., Read, J. D., & Lindsay, D. S. (2002). A picture is worth a thousand lies: Using false photographs to create false childhood memories. *Psychonomic Bulletin and Review*, 9, 597–603.
- Wagenaar, W. A. (1988). *Identifying Ivan: A case study in legal psychology*. Cambridge, MA: Harvard University Press.
- Wagenaar, W. A., & van der Schrier, J. H. (1996). Face recognition as a function of distance and illumination: A practical tool for use in the courtroom. *Psychology, Crime and Law*, 2, 321–332. doi:10.1080/10683169608409787
- Washington v. Davis/Northrop (2010). Findings of fact and conclusions of law and order.
- Webert, D. R. (2003). Are the courts in a trance? Approaches to the admissibility of hypnotically enhanced witness testimony in the light of empirical evidence. *American Criminal Law Review*, 40, 1301–1327.
- Wells, G. L. (1984). The psychology of lineup identification. *Journal of Applied Social Psychology*, 14, 89–103.
- Wells, G. L., Charman, S. D., & Olson, E. A. (2005). Building face composites can harm lineup identification performance. *Journal of Experimental Psychology: Applied*, 11, 147–156.
- Wells, G. L., Leippe, M. R., & Ostrom, T. M. (1979). Guidelines for empirically assessing the fairness of a lineup. *Law and Human Behavior*, 3, 285–293.
- Wells, G. L., & Quinlivan, D. S. (2009). Suggestive eyewitness identification procedures and the Supreme Court's reliability test in light of eyewitness science—30 years later. *Law & Human Behavior*, 33, 1–24.
- Wells, G. L., Small, M., Penrod, S. D., Malpass, R. S., Fulero, S. M., & Brimacombe, C.A.E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior*, 22, 603–647.
- Wickens, T. D. (2002). *Elementary signal detection theory*. New York, NY: Oxford University Press.
- Williams, G., & Hammelmann, H. A. (1963). Identification parades. *Criminal Law Review*, 62, 479–491, 545–555.
- Wright, D. B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta Psychologica*, 114, 101–114. doi:10.1016/S0001-6918(03)00052-0
- Yuille, J. C., & Tollestrup, P. (1990). Some effects of alcohol on eyewitness memory. *Journal of Applied Psychology*, 75, 268–273. doi:10.1037/0021-9010.75.3.268
- Yuille, J. C., Tollestrup, P. A., Marxsen, D., Porter, S., & Herve, H.F.M. (1998). An exploration on the effects of marijuana on eyewitness memory. *International Journal of Law and Psychiatry*, 21, 117–128.

CHAPTER 18

Evaluating Eyewitness Testimony of Children

SUE D. HOBBS, JONNI L. JOHNSON, GAIL S. GOODMAN, DANIEL BEDERIAN-GARDNER,
MICHAEL J. LAWLER, IVAN D. VARGAS, AND MACARIA MENDOZA

EACH year, hundreds of thousands of children suffer or witness crime; some of them then serve as witnesses in forensic investigations and legal proceedings. This is especially true in sexual abuse cases where the offense typically is committed in secrecy (Bala, Lee, & McNamara, 2001), and often there is no visible injury or physical evidence, so that the children's eyewitness memory accounts take center stage (Keeney, Amacher, & Kastanakis, 1992; Myers, 1993a). However, children witness and experience many other crimes as well, such as domestic violence, homicide, war atrocities, school shootings, and kidnappings. Children who are crime victims or witnesses may be questioned by police officers, child advocacy workers, social workers, and attorneys. Additionally, they may need to testify at depositions, preliminary hearings, and trials and, if a defendant is convicted, a sentencing hearing (Quas & Goodman, 2012; Troxel, Ogle, Cordon, Lawler, & Goodman, 2009).

When adults do not believe a child's accurate testimony, it can have devastating consequences. If child victims are not believed, a perpetrator is free to commit other crimes, and the victims may be placed in further danger due to retaliation against them by the perpetrator. For example, when a 14-year-old told her minister that she and her five sisters were being sexually abused by their parents, the minister warned her parents before alerting authorities. Because of this, her parents, Bruce and Glenda Dutro, subjected all six Dutro children to several days of confinement, beatings, rape, and starvation. When social workers finally visited the home, the girls were too terrified of their parents to disclose the abuse. The assaults continued, and it was not until several years later that authorities were alerted again and the parents were finally convicted. Bruce Dutro is currently serving a 300-year prison sentence, and his wife, Glenda, is serving a 15-year sentence (*People v. Dutro*, 2012).

However, when children's accounts are inaccurate, believing them can also lead to injustices that include conviction of the innocent. Take the case of David Wiggins. In July 1988, a 14-year-old girl opened her back door to let out her dog, and a man forced himself inside her home. He pushed her to the floor, put a towel over her face, and raped her. The child saw the man's face briefly. She called 911 shortly after her attacker left and was taken to the hospital. Two days after her assault, the child was shown a photo lineup that included Wiggins's photo. The child wrote next to Wiggins, "looks familiar." The following day, Wiggins, who had been arrested for riding in a stolen car, was placed in a live lineup. The child identified him as her attacker during the lineup and again at the trial. Wiggins was convicted and sentenced to life in prison. Later DNA testing indicated that Wiggins was not the perpetrator, and the prosecutor agreed that a wrongful conviction was likely. Wiggins was released after serving over 20 years in prison (Innocence Project, 2012).

Such real-world cases illustrate why children's eyewitness abilities are of paramount interest for legal professionals and researchers. Children's reports are the linchpins in many legal proceedings, especially when physical evidence is absent. Research on the abilities of child eyewitnesses may be particularly important in assisting investigators when children's reports are the only piece of evidence, as is often the circumstance in child sexual abuse cases. In this chapter, we discuss factors that may influence the accuracy and perception of children's reports. This review is not exhaustive, but we hope to draw attention to areas of consensus and foster dialogue about areas of controversy that will assist in building theoretical understanding and optimal legal application concerning children's eyewitness reports.

MEMORY DEVELOPMENT

Before delving into research and theory on children's eyewitness abilities, it is important to have a basic understanding of age trends in memory development. Children undergo marked changes in encoding, knowledge base, and retrieval with age (Howe, 2011). Although memory development continues into adolescence and adulthood, a qualitative jump occurs after the early preschool years. On eyewitness memory tasks, it is particularly challenging to obtain complete and accurate information from young preschoolers (e.g., Goodman & Reed, 1986). Compared to older children and adults, younger children recall less information in response to free recall questions and open-ended questions (e.g., "What happened?"), and they make more errors in response to direct questions, such as yes/no queries (e.g., "Was his shirt red?" "Did he shut the door?" "Did he kiss you?"), option-posing queries ("Did he have a knife or a gun?" "Was her hair straight, curly, or braided?"), and misleading questions (e.g., "He took your pants off first, didn't he?" when in fact, he did not; Dent & Stephenson, 1979; Goodman, Bottoms, Schwartz-Kenney, & Rudy, 1991; but see Ceci, Papierno, & Kulkofsky, 2007). Postevent misinformation that is stated as a presumption (e.g., "How fast was the car going when it passed the barn on the country road?" when in fact there was no barn) is also more likely

to contaminate young children's memory reports than those of older children and adults (Ceci, Ross, & Toglia, 1987; Schwartz-Kenney & Goodman, 1999), although adults can also be susceptible to such false information (Loftus, 1979). Although by about the age of 5 or 6 years, children often can identify a culprit as accurately as adults when presented with photo lineups that contain the perpetrator (Goodman, Hirschman, Hepps, & Rudy, 1991), younger children are also more likely than older children and adults to falsely identify an innocent person in photo lineups that do not include the offender (Pozzulo & Lindsay, 1998). There are, however, marked individual differences at any age; for example, some children as young as 2 or 3 years can be highly accurate and resistant to false suggestions (Harris, Goodman, Augusti, Chae, & Alley, 2009).

Children are likely to have weaker memory traces than adults and to have greater difficulty with source monitoring (Howe, 2011; Johnson & Foley, 1984). This then naturally leads to questions, such as whether children can maintain accuracy of their memory reports as time goes on and as memory traces become weaker or source monitoring becomes more difficult. Such questions have obvious legal relevance because some crimes are not readily reported; children may need to recall a forensically relevant event that occurred days, months, or even years earlier. Recently, Peterson (2011) suggested that children's reports about personally salient, stressful events remained accurate even with the passage of years. Yet how researchers assess the accuracy of these reports affects whether one concludes that accuracy is maintained or declines over time.

Overall, memory performance tends to improve across childhood and into adulthood, including on eyewitness memory tasks. However, the research base mainly concerns children's memory for unfamiliar people and briefly witnessed events. Situations about which children testify often involve familiar people and events that are traumatic or stressful.

TRAUMA, STRESS, AND MEMORY

Many criminal events are traumatic for children to witness or experience or, because of their potential for violence, cause child witnesses to experience considerable distress and anxiety. Thus, research investigating the impact of violence and stress on memory is of crucial importance to understanding children's eyewitness testimony. The extent to which children can remember and accurately report personally traumatic and stressful events is a topic of active research. Many children can, under a variety of circumstances, provide forensically relevant, accurate information about highly traumatic events they have witnessed or experienced (e.g., D. P. H. Jones & Krugman, 1986; McWilliams, Narr, Goodman, Ruiz, & Mendoza, 2013). In both children and adults, such events typically are recalled more accurately and for a longer period of time relative to benign or ordinary events (e.g., Peterson, 2012). Highly distressing events can also be recalled with error and are not immune to forgetting

and distortion, including false memory, in adults and children (e.g., Hirst et al., 2009; Neisser & Nicole, 1992; Terr, 1983).

In considering the research on children's memory for traumatic and stressful events, it is important to consider whether memory is assessed in field studies, in which witnesses of actual crimes are interviewed (Orbach, Lamb, La Rooy, & Pipe, 2012; see Paz-Alonso, Ogle, & Goodman, 2013, for review), or laboratory studies, where witnesses' emotional responses to and memories of more standardized incidents (e.g., slide-depicted or staged events) are examined. In particular, questions arise concerning the external validity of laboratory research (e.g., how well laboratory research sufficiently mimics the levels of distress induced by criminal events) and the internal validity of field research (e.g., how well field researchers can pin down cause-effect relations). Ideally, findings from laboratory and field research lead to the same conclusions, but this is not always so. For example, in a series of studies on children's memory for child sexual abuse, Leander and her colleagues (e.g., Leander, Christianson, & Granhag, 2007; Leander, Granhag, & Christianson, 2005) and Cederborg, Lamb, and Laurell (2007) found that children avoided detailed reports of child sexual abuse they had experienced (e.g., as documented by photographs of the assaults found in perpetrators' possession) and that some children completely denied that the abuse occurred, even when shown photographs of the molestations, likely due to embarrassment, feelings of guilt or complicity, or desire to protect the abuser. Such factors have rarely been studied in laboratory research (but see Saywitz, Goodman, Nicholas, & Moan, 1991) yet may have a profound effect on children's memory for actual crimes.

In any case, it is clear that many factors play a role in children's memory for traumatic and stressful events—too many to review in this chapter. Here we first consider some of the theoretical issues involved in memory for stressful and traumatic experiences. We then turn to a subset of the factors that affect children's memory for stressful events, such as age when events occurred, language and parental factors, centrality of the to-be-remembered information, whether the individual is a participant or a bystander witness, and whether events are repeated or single occurrences. Additionally, we review research on physiological stress responses—research that is furthering our knowledge about how stress affects children's memory of traumatic and stressful events. Clearly, a complex multivariate model of children's memory for stressful events is needed to integrate disparate findings.

THEORETICAL ISSUES

There has been considerable theoretical debate as to whether memory is diminished or enhanced for highly stressful experiences (e.g., Christianson, 1992; Deffenbacher, Bornstein, Penrod, & McGorty, 2004). Christianson (1992) argued that with heightened distress, attention becomes particularly focused on the central stressor as do elaborative processes later, resulting in particularly accurate memory for the main stressor. Deffenbacher et al. (2004) countered that in the studies reviewed by Christianson, sufficiently high levels of distress had not been achieved to uncover

decrements in memory that occur when people are confronted with life-threatening situations.

However, a growing body of research indicates that memory is particularly robust for highly stressful experiences and that neural circuits involving, for example, hippocampal and amygdala structures, support memory for highly emotional information (Phelps, 2004). When stimuli are high in arousal and negative in valence, memory is particularly likely to be robust (Kensinger, 2009; Mather & Sutherland, 2009). Perhaps relevant to memory for violent criminal events, Nairne, Thompson, and Pandeirada (2007) have proposed that, due to evolutionary forces, human memory is tuned to information that was relevant to survival in our ancestral past, a theoretical idea known as survival-based processing. Being the victim of a violent crime or observing others being assaulted would be particularly relevant to survival, both in our ancestral past and in the present, and thus one would expect especially accurate and robust memory for such acts. One possible reason for improved retention is that a specialized memory module exists for processing information important for survival.

Other possible mechanisms include that stressful, survival-relevant experiences are more emotionally arousing, more distinct, and more self-relevant, leading to better retention (Howe & Otgaar, 2013). Researchers still debate if robust memory for traumatic experiences reflects such basic memory processes (Howe, 2011) rather than emotional and neurobiological factors associated with trauma (Cordón, Pipe, Sayfan, Melinder, & Goodman, 2004). In any case, it is clear that, despite relatively strong retention, memories of highly stressful and traumatic events still may be subject to distortion and forgetting in children and adults (Otgaar & Smeets, 2010).

Chae, Goodman, and Edelstein (2011) argued that Christianson's and Deffenbacher et al.'s perspectives may have merit, depending on individual differences in the processing of emotional information, including for children. For example, individual differences in attachment-related insecurities influence the extent to which individuals encode, elaborate on, and later retrieve and report negative information. In particular, attachment-avoidant individuals (i.e., people who avoid intimacy and do not want to be emotionally needy) are likely to engage in "defensive exclusion" of information and emotions related to reminders of upsetting events (Bowlby, 1982). Defensive exclusion may in turn influence the processing of and memory for negative experiences: Compared to less attachment-avoidant individuals, more avoidant ones are less likely to have discussed their traumatic experiences with others and less likely to retain accurate memory over time (Edelstein et al., 2005). Thus, although memory in general is often particularly accurate and enduring for central details of events relevant to survival (Christianson, 1992), defensive processes may inhibit encoding, storage, and/or retrieval of memories of such experiences, leading to memory deficits or distortions in some individuals (Deffenbacher et al., 2004).

Several studies uncovered links between parents' attachment-related insecurities and children's memory for and suggestibility regarding stressful experiences.

For example, children of parents who score relatively high on measures of attachment avoidance provide less accurate memory reports and display heightened suggestibility regarding highly stressful medical procedures (e.g., Goodman, Quas, Batterman-Faunce, Riddlesberger, & Kuhn, 1997). Moreover, parental attachment insecurities are among the few individual difference variables that consistently predict children's suggestibility (Bruck & Melnyk, 2004).

While theoretical issues continue to be debated and researched, empirical evidence of children's memory for stressful events continues to mount and likely will constrain theory as the field moves forward. In the meantime, a number of factors have been found to affect children's memory for stressful events, some of which we turn to next.

CHILDREN'S AGE

Age at time of a stressful or traumatic event can affect how well it is remembered later on. Children can at times recall things that happened years before when they were quite young (Peterson, 2012). However, it is extremely unlikely that children (or adults) can remember and report events that occurred before age 1 (Peterson). Additionally, many individuals cannot remember traumatic or nontraumatic events that occurred before age 3 (Malloy & Quas, 2009; Terr, 1988), a phenomenon often referred to as infantile amnesia. Although the age at which the infantile amnesia barrier is considered to take hold has been gradually lowered by researchers (e.g., some older children and adults can recall information down to 2 years of age; Usher & Neisser, 1993), the offset of infantile amnesia usually is thought of as around 3 to 4 years of age, such that most older children and adults cannot accurately retrieve memories of events that occurred before those ages (Peterson, 2012). In any case, children's ability to remember and accurately report events continues to improve with age.

Peterson (2012) described results from her longitudinal study examining 2- to 13-year-old children's memories of traumatic injuries (e.g., broken bones or lacerations) as well as for the emergency room treatment that resulted. The children were interviewed at the time of the injury and emergency room treatment and again several years later. In terms of memory *completeness* regarding the *injury*, when interviewed after a 5-year delay, all children who were 2 years old or older at the time of the injury remembered it nearly as well as they had 5 years previously. However, older children's memory *accuracy* for their injuries was better compared to that of younger children, and all children's memory accuracy decreased over time (see Peterson). The biggest decrease in accuracy occurred during the year following the incident. Of interest, children's memories for *hospital treatment* declined in both completeness and accuracy. Still, the children's completeness and accuracy declined from about 92% to 85%, showing substantial robustness of memory even after a 2-year delay.

In a study of 3- to 10-year-olds, Goodman et al. (1997) found that age was positively related to the accuracy of memories of a stressful urinary catheterization procedure

involving genital penetration. In that study, memory was tested within a week to a month after the medical procedure. Results of a study evaluating children's memory for enucleation (removal of an eye due to retinoblastoma) showed that children who were older than 24 months talked more about the procedure than those who were younger. Unfortunately, memory accuracy was not assessed (Norgate & Littleton, 2011).

LANGUAGE AND PARENTAL COMMUNICATION

Also related to children's memory are language and parent/child communication factors. Some research has shown that children who remember an event up to 14 months after it occurred do not use language in their descriptions that was not in their vocabularies when the event occurred (Hayne & Simcock, 2009). Although such findings suggest that preverbal memories cannot be recalled verbally, more recent research indicates that some children can, at times, recall information for which they did not have those specific words earlier (Morris & Baker-Ward, 2007). These findings have fascinating legal implications, especially given the fact that children's competence to testify is assessed at the time of testimony rather than at the time of the alleged offense (Lyon, 2011).

In any case, once children develop language ability, opportunities increase for parent-child discussion about events. Peterson, Sales, Rees, and Fivush (2007) evaluated the relation between parent-child conversation and 2- to 5-year-old children's memories for an injury and subsequent emergency room treatment. Age and parental elaborative conversation style were the two most significant predictors for children's memory for both the injury and hospital treatment. In fact, parental elaboration was the most significant predictor of children's memory accuracy for hospital treatment, an event that appears to be more difficult to remember compared to an injury that prompts hospital treatment (Peterson, 1999, 2002). These findings suggest that discussions parents have with their children about traumatic events can assist with the encoding and storage processes necessary for memory retrieval (Chae, Ogle, & Goodman, 2009).

PARENTING STYLE

Another parental factor related to children's memory for traumatic and stressful events is parenting style. Children with parents who have a more traditional parenting style with a focus on obedience and parental authority remember less about a somewhat stressful event than children who have more authoritative parents (Burgwyn-Bailes, Baker-Ward, Gordon, & Ornstein, 2001). The researchers suggest that their findings may stem from children with authoritarian parents distrusting their own interpretation and memory of events. Furthermore, children may model their parents' coping behaviors, which may include avoidance or anxiety in response to stressful events (Brumariu & Kerns, 2010), both of which can affect memory (Chae et al., 2011).

CENTRAL VERSUS PERIPHERAL DETAILS

An important factor that plays a role in memory for events in general is the centrality of the information (i.e., how central or peripheral the details are that need to be remembered). For traumatic events, however, a “tunnel effect” can occur in memory, with heightened memory for central details and diminished memory for peripheral details (Christianson, 1992). Typically (albeit not always), in criminal investigations, central details are of most importance. As a general rule, crime witnesses are most likely to encode and remember central aspects of the crime better than more peripheral details. This trend holds for older and younger child witnesses, as least once children reach the age of approximately 5 years (e.g., Eisen, Goodman, Qin, Davis, & Crayton, 2007).

Of interest, the classification of a to-be-remembered detail as a central or peripheral event may differ depending on how relevant that detail is to an individual’s goals. For example, an individual whose goal is to suppress emotion may remember an emotional event less well than an individual not so motivated (Levine & Edelman, 2009). Contradictory findings about memory for central and peripheral details may result in part from a lack of consideration of individual goals as well as from differences in how centrality is defined across studies (Paz-Alonso, Goodman, & Ibabe, in press).

Even if children remember fewer peripheral compared to central details or remember peripheral details incorrectly, memory for central details still can be quite accurate. However, heightened memory for central versus peripheral details is not consistently found as a function of age across studies. In the Peterson (2011) study, which was briefly described earlier, younger children accurately remembered more central details relative to peripheral details, whereas older children remembered central and peripheral details equally well. In contrast, Eisen et al. (2007), in evaluating children’s memories for an anogenital exam and blood draw, reported that older children (ages 6 to 11) made fewer errors on central-specific questions compared to peripheral-specific questions, whereas this was not the case for younger children (ages 3 to 5). In a study of 9- to 12-year-old children’s memory for an impromptu speech and stressful math task, as self-reported stress increased, memory accuracy for central details increased, whereas memory accuracy for peripheral details decreased (Rush, Quas, & Yim, 2011). These contrasting findings may reflect not only differences in how researchers operationalize centrality distinctions but also differences in what children of various ages consider to be central versus peripheral to the main stressor.

PARTICIPANT VERSUS BYSTANDER CHILD WITNESSES

Many child eyewitness memory studies concern bystander witnesses—for example, children who view others performing actions. However, there is evidence to suggest that children who actively participate in events, more as a victim might, remember the event better than do bystander witnesses (e.g., Rudy & Goodman, 1991).

Pipe, Lamb, Orbach, and Esplin (2004) also note that people tend to recall events they experienced better than events they witnessed. Greenhoot, McCloskey, and Glisky (2005) evaluated adolescents' memory of family violence that had occurred 6 years previously. Adolescents remembered abuse that was committed against them better than they remembered abuse of their mothers. However, Christianson (1992) cited studies in which there were no differences in the accuracy of victims' and witnesses' memories. An important factor in the participant-over-bystander memory advantage may be activation of self-schema. That is, when self-schema are activated, a richly elaborated memory structure may help maintain storage of the memory. Although self-schema are likely to be activated when one is actively involved in an event, such schema may also be activated when watching an event unfold, which could then support accurate memory (Baker-Ward, Hess, & Flannagan, 1990; Howe & Otgaar, 2013), perhaps especially when the event has high personal relevance to a child's life (McWilliams et al., 2013).

REPEATED EVENTS

How frequently events are experienced is another factor likely to affect how well a stressful event is remembered. Unfortunately, little rigorous scientific research has examined children's memory for single versus repeated stressful or nonstressful events that were highly stressful. Goodman et al. (1997) found that children who experienced a single invasive medical procedure remembered it as accurately as children who experienced it multiple times. Discriminating between different instances of repeated events can be difficult for anyone (Pipe et al., 2004), and it is especially difficult for young children (Lyon & Saywitz, 2006; Pipe et al., 2004). For example, children who are victims of incest or repeated sexual assault are not always able to remember the details of each incident or whether a certain act occurred during the first or 50th assault (Bala et al., 2001). This is especially true if similar abuses are experienced repeatedly. Farrar and Goodman (1992) evaluated 4- and 7-year-old children's memories for repeated similar nonstressful events. The 4-year-olds merged the events more so than did the 7-year-olds. Relative to their older counterparts, the younger children had more difficulty separately recalling features of each event. Children may confuse details across events yet still may report the gist accurately (Pipe et al., 2004). More research is needed, however, on children's memory for repeated stressful events.

PHYSIOLOGICAL DISTRESS

Researchers are just beginning to evaluate children's physiological distress in relation to children's memory for stressful events. Quas, Yim, Edelstein, Cahill, and Rush (2011) uncovered a positive relation between 9- to 12-year-old children's cortisol levels and their memories for a stressful experience. Quas and Lench (2007) found an interaction between increased heart rate during encoding (watching a

frightening video) and nonsupportive interviewing for the video 1 week later. Specifically, the most memory errors were committed by 5- and 6-year-old children who had the largest increase in heart rate and were interviewed by an unsupportive interviewer. Of interest, in a study by Quas, Carrick, Alkon, Goldstein, and Boyce (2006), children's age and sympathetic nervous system reactivity (a measure of physiological distress) were significantly related to decreased memory accuracy for a stressful event (a 1-minute fire alarm).

CONCLUSION

Because attention is limited, people cannot encode everything about real-life events, particularly those as complex as most crimes. We have reviewed some of the factors that are related to how well children remember traumatic and stressful events. A complex multivariate model may be needed to create a clearer picture of children's memory for such experiences.

EFFECTS OF MALTREATMENT AND TRAUMA-RELATED PSYCHOPATHOLOGY ON MEMORY

Child abuse and neglect are major risk factors for anxiety, substance abuse, psychosis, and personality disorders (Felitti, 2002). Exposure to early abuse may affect the development of the hippocampus, a brain structure important for memory (Teicher, Anderson, & Polcari, 2012). Furthermore, when maltreated children are removed from their homes by protective services staff, the removal itself is likely to be traumatic (Baugerud & Melinder, 2012), as can be many of the social service and legal experiences that follow (e.g., Block, Oran, Oran, Baumrind, & Goodman, 2010; Quas, Goodman, et al., 2005).

Both nonmaltreated and maltreated children show age improvements in the accuracy and completeness of their reports as well as in their resistance to misleading questions (e.g., Eisen et al., 2007). Although maltreated children's basic memory processes are comparable to those of nonmaltreated children (Howe, Cicchetti, Toth, & Cerrito, 2004), abused children's memories of life events are at times overgeneralized, indicating difficulties with autobiographical memory (Valentino, Toth, & Cicchetti, 2009).

Maltreatment may affect children's emotion regulation strategies (Kim & Cicchetti, 2010). The emotion regulation strategies of avoidance of memory or of dissociation are predictors of memory deficits and/or greater suggestibility in general (Qin, Ogle, & Goodman, 2008; J. M. G. Williams, 1996) as well as in maltreated children. Eisen et al. (2007) examined maltreated children's memory for an anogenital examination and venipuncture (blood draw) conducted by doctors in a forensic hospital unit. Controlling for age and gender, children who scored highest on measures of dissociative tendencies, trauma symptoms, and cortisol increase after the medical experiences made more errors on misleading questions

relative to the remaining children. For highly traumatized children who were high dissociators, increased cortisol was related to poorer memory; for less traumatized children who were low dissociators, increased cortisol was associated with better memory. Melinder, Baugerud, Ovenstad, and Goodman (2013) examined memory in maltreated children who were being removed from home by protective services staff. When children's memory for the removal was studied, parents' avoidant attachment style predicted less accurate memory and increased suggestibility. Presumably, at least according to Bowlby's attachment theory, children had learned from their avoidant parents to avoid thinking or talking about stressful, negative experiences. In a more direct test of the effects of attachment avoidance on memory of child maltreatment victims, Edelstein et al. (2005) found that child sexual abuse victims who scored higher on a measure of avoidant attachment were particularly likely to have memory deficits for the abuse itself if the abuse was especially severe (e.g., rape at gunpoint or years of incest), whereas child sexual abuse victims who scored lower in attachment avoidance evinced the opposite trend: The more severe the sexual assault, the better was their memory. These results are consistent with those, mentioned earlier, from research on children's memory for medical procedures when children appear highly distressed by such experiences (Goodman et al., 1997). This consistency extending to child maltreatment victims demonstrates the broad generality of the avoidance and child memory findings and likely reflects the impact of emotion regulation strategies (Goodman, Quas, & Ogle, 2009).

Child maltreatment places children at risk of trauma-related psychopathology (e.g., posttraumatic stress disorder [PTSD], depression, dissociation). Evidence suggests that maltreated children's performance on eyewitness memory tests for negative information is more strongly associated with psychopathology than with maltreatment per se (McWilliams, Harris, & Goodman, 2012).

PTSD has been of special interest to those who study trauma and memory in maltreated children as well as to those who study traumatized adults (e.g., Rubin, Berntsen, & Bohni, 2008). It has been proposed that individuals with PTSD develop "fear networks," that is, semantic and episodic mental networks that store trauma-related information in a hyperactivated form and make victims particularly attentive to such information (Foa & Kozak, 1986). Fear networks may support particularly accurate memory for trauma in individuals with PTSD. Robust remembrance of trauma is reflected in flashbacks of traumatic information characteristic of PTSD, although avoidance of reminders of the trauma is also characteristic of the disorder. That said, memory monitoring problems in those with PTSD may increase commission errors for trauma-related or non-trauma-related information (Bremner, Shobe, & Kihlstrom, 2000; Windmann & Krüger, 1998).

CHILDREN'S SUGGESTIBILITY, FALSE REPORTS, AND FALSE MEMORY

Children's suggestibility and false memory are crucial issues in the study of children's eyewitness testimony. The devastating consequences of children making

false accusations were demonstrated during the 1980s in the McMartin child sexual abuse trial. Preschool-age children made accusations about sexual abuse against the nursery staff. Intense investigation ensued, yet no evidence was discovered to support the children's claims (e.g., that they were sexually abused in tunnels under the day-care center—no tunnels were found). The prosecution cost millions of dollars and caused great distress for the defendants and families involved.

Generally speaking, age is the strongest predictor of suggestibility and false memory reports; younger children are typically more suggestible and more prone to false memory reports than older children, adolescents, and adults (e.g., Goodman, Bottoms, Rudy, Davis, & Schwartz-Kenney, 2001; Malloy & Quas, 2009). That said, there are important individual differences in suggestibility and misinformation effects within any age-group. Although it is difficult to predict such individual differences, child forensic interviewers should be knowledgeable about the possibility that children may incorporate interviewer suggestions or misinformation and should have appropriate expectations for children relevant to the children's ages (Lamb, Malloy, & La Rooy, 2011; Malloy & Quas, 2009). It is important for investigators and interviewers to consider how children's suggestibility can influence their reports.

Suggestibility has been defined as "the degree to which encoding, storage, retrieval, and reporting of events can be influenced by a range of social and psychological factors" (Ceci & Bruck, 1993, p. 404). In the McMartin case, it is largely agreed within the scientific community that the police investigators and parents suggestively questioned the children, which ultimately may have implanted, through misinformation, abuse details in the children's memories or at least in the children's reports. This form of suggestibility—that of incorporating misinformation into one's own memory—not only has crucial legal consequences but it also has important theoretical implications for developmental and cognitive psychology (Ceci & Bruck, 2006; Johnson, Hashtroudi, & Lindsay, 1993; Loftus, 1975; Pezdek & Roe, 1995).

THEORETICAL ISSUES

Several theories have been proposed to account for the mechanisms associated with the form of suggestibility that can lead to memory report errors. Memory factors have been emphasized in most of these theoretical accounts. Such factors include trace alteration (Loftus, 1975), trace strength (Brainerd & Reyna, 1998), memory coexistence/retrieval blocking (Bekerian & Bowers, 1983; Eakin, Schreiber, & Sergent-Marshall, 2003), source misattribution (Johnson et al., 1993), and activation-based associative networks (Ayers & Reder, 1998). Although memory factors undoubtedly play a vital role, social factors (e.g., demand characteristics) are also important in producing misinformation effects (Roediger, Meade, & Bergman, 2011).

Cognitive and psychosocial mechanisms that develop throughout childhood bolster one's abilities to resist suggestion or misinformation. Cognitive and developmental theories assist in identifying the mechanisms that may be associated

with suggestibility's influence on children's memory reports (Chae et al., 2011; McWilliams, Bederian-Gardner, Hobbs, Bakanosky, & Goodman, 2012).

From a memory trace theoretical perspective, memories are preserved as traces, a consolidation of current features or attributes related to the person and event. When activated, these traces assist in recalling the details associated with that memory. Pezdek and Roe (1995) asserted that when memory traces are strong (i.e., they contain elaborative details, such as of time, place, individuals involved in the event) and are preserved during memory storage, they will be most resistant to suggestion. Children who have strong memory traces or representations can dismiss externally generated suggestions because they can directly compare information being suggested back to the trace that was recovered and conclude that the two accounts do not match.

However, when traces are weak, children may incorporate suggestions or misinformation because they can no longer counter with their own representations. Pezdek and Roe found support for this claim in their 1995 study examining memories for pictures in 4- and 10-year-old children. Some children viewed target pictures multiple times, which increased the strength of the memory trace through multiple presentations, whereas other children viewed the materials only once. Children then either did or did not receive misinformation by way of researchers reviewing the slides via a narrative. In the misinformation narrative condition, the researcher replaced items that were viewed with items that were not originally viewed (e.g., suggesting the picture item included a cup when in reality it was a dish). In addition to finding typical age differences, with older children recognizing pictures with greater accuracy than did younger children, increased viewing time (i.e., multiple exposures and increased memory strength) resulted in the 4- and 10-year-olds being more accurate and less suggestible than children who had viewed the pictures only once. The minimal exposure left weaker memory traces that enabled them to be tampered by the researcher's suggestions.

This idea of strong versus weak traces is also relevant to Brainerd and Reyna's fuzzy-trace theory (FTT, 2002), which stipulates a dual process model for memory encoding and retrieval processes. Memories are represented as either verbatim traces, which hold specific details about the memory, or gist traces, which hold the general meaning of the memory. An example of a verbatim trace would be remembering that you went to the movie *Titanic*, with your friend *Jon*, and ate *popcorn* at the theater; the gist trace would include the details that you went to see a *movie* with a *friend* and ate a *snack*. As verbatim traces hold more details that cannot be maintained for every memory experienced, these traces decay more quickly, often leaving only the gist trace behind. Gist traces are more susceptible to suggestion and misinformation as the original record of the event (i.e., verbatim trace) cannot be recovered to counter the suggestion. This effect is strongest when the suggestion is more similar to the gist trace and cannot be temporally discriminated from the original trace (Brainerd & Reyna, 2004; Reyna & Brainerd, 2011). Abilities to form verbatim and gist traces improve as children develop (Reyna & Brainerd,

2011) with gist trace mechanisms taking longer to develop over childhood than verbatim trace mechanisms (Brainerd, Forrest, Karibian, & Reyna, 2006). Although younger children are more reliant on verbatim than gist traces than are adults, older compared to younger children still have better verbatim traces (e.g., Reyna & Kiernan, 1994). Therefore, older children, who have stronger verbatim traces, should be less suggestible than younger children, according to FTT, although adults may be more subject to certain false memories than children, if the false memories are supported by gist traces (Brainerd, Reyna, & Ceci, 2008).

Source monitoring (SM) theory (Johnson et al., 1993) has also been used to account for children's suggestibility and misinformation effects. According to SM theory, details for memories are discriminated against one another via a decision process in which one attributes the source of these details using perceptual processes (i.e., perceiving a cue) and cognitive processes (e.g., retrieval strategies). During retrieval, individuals engage in decision processes regarding source information (*where, when, what, and with whom* details of events). Cues that are retrieved are evaluated with reality monitoring (i.e., deciding if the detail actually occurred in reality or if it was imagined), internal monitoring (i.e., deciding if details actually occurred or only were thought about), and external monitoring (i.e., deciding if details were from this event or another event) processes. The SM theoretical framework assumes that certain cognitive abilities are in place to assist retrieval during more difficult monitoring times (e.g., decision making, metamemory strategies). Such abilities change and improve in children as they develop (e.g., Bjorklund, Dukes, & Douglas-Brown, 2008; Ghetti, 2008; D. S. Lindsay, 2002). For example, Poole and Lindsay (1995) examined source monitoring in the context of children's suggestibility when children were exposed to misinformation introduced via parents. Their study examined memories of children between the ages of 3 and 7 for a one-on-one experience with "Mr. Science" wherein the confederate, Mr. Science, demonstrated several activities for the children. Following these interactions, each child was interviewed in a nonsuggestive manner (i.e., interviews began with free-recall prompts followed by open-ended prompts, "Can you tell me more about that?") about what had happened. In this immediate interview, children were highly accurate in their reports.

In the weeks that followed this initial interview, the parents of child participants were asked to read a story about Mr. Science to their children. This story included events that happened during each child's initial interaction with Mr. Science as well as new events that had not been originally experienced. Children were interviewed again, only this time interviews contained misleading questions that included information provided only in the story (e.g., "Did you make paper airplanes with Mr. Science? Tell me more about when you made paper airplanes with Mr. Science"). Children, particularly younger ones, incorporated information that they had heard, but had not experienced, into their free-recall reports (e.g., approximately 41% of 3- and 4-year-olds included details of false events) and succumbed to direct misleading questions. Children were also asked source monitoring questions about

whether they had actually witnessed the activities or heard about the events in the stories. The younger children had considerable difficulty with the source monitoring task. This study is often cited as an indication that children can be led into false reports through source monitoring errors. In this study and others, according to SM theory, younger children likely did not have the cognitive abilities to monitor the source of the information experienced in the event versus suggested by their parents or the interviewer) well enough to answer the questions correctly. However, it is important to note that even young children, despite making more errors than older children, appropriately reject many of the false event details in most of these studies (e.g., Goodman et al., 2001).

False memories of entire events also can be formed based on suggestibility. False memory formation has been explained by theories previously mentioned. Like suggestibility, the ease with which false memories can be implanted tends to decline as children age and acquire more cognitive abilities that allow them to create lasting memories and monitor intrusions (e.g., Ghetti, 2008; Otgaar & Candel, 2011). However, older children and adults succumb to false memories as well (e.g., Loftus & Pickrell, 1995). With ecologically valid false memory paradigms that parallel eyewitness abilities (see Otgaar & Candel, 2011), false memories are more frequently observed in younger compared to older children (but see articles on the Deese-Roediger-McDermott (DRM) false memory illusion; Brainerd, Reyna, & Zember, 2011). It is likely that older children's experiences and improvements in cognitive abilities permit them to evaluate the plausibility of suggested events.

For example, Pezdek and Hodge (1999) examined instances of false memory in a sample of younger children (5- to 7-year-olds) and older children (9- to 12-year-olds). Children were read four descriptions of events they were instructed they had experienced. In reality, they had only experienced two of these events. After hearing these descriptions, children were asked to remember every detail they could about these events. The latter false events consisted of a plausible event (i.e., getting lost in the mall) or implausible event (i.e., receiving a rectal enema). After hearing these descriptions of the four events, younger children were more likely than older children to recall details from both false events. Although some older children did recall details from the plausible event, none of the older children succumbed to false-memory intrusion of the implausible event. However, some younger children did report false memories of the rectal enema experience. It is possible that the older children were more likely than the younger children to evaluate the plausibility of the events.

So far, we have mainly discussed suggestibility and false memory trends as they relate to theoretical issues. However, it is important to review empirical evidence concerning misleading questions and repeated interviews specifically, because these topics are of considerable legal concern. Our discussion is not exhaustive of all the factors that affect children's memory, suggestibility, and false memory formation (for review, see Blandon-Gitlin & Pezdek, 2009; Bruck & Melnyk, 2004; Malloy, Johnson, & Goodman, 2013; Malloy & Quas, 2009).

MISLEADING QUESTIONS

Since the mid-1980s, children's suggestibility has been examined in relation to interviewer question type, specifically using interviews that include misleading questions about the event the child is recalling. In these paradigms, researchers have children (often preschool age) participate in controlled events and, after a specific period of delay, interview them suggestively. That is, questions asked by the interviewers presuppose or introduce false information about the event to examine whether children acquiesce to these suggestions or appropriately deny them (e.g., the question "Did you see the man knock over and break the lamp?" presumes that the man *did* knock over and break a lamp). Children's suggestibility is then scored or characterized by the likelihood or frequency of acquiescence to interviewer suggestions. Typical age trends emerge under this experimental paradigm; older children are less suggestible than younger children, as older children acquiesce less frequently, regardless of whether the event is distressing (e.g., Goodman et al., 1997; Peterson, 2011) or commonplace (e.g., Quas et al., 2007). However, it should be noted that children are often less suggestible about personally significant negative events (e.g., being hit, being naked, having their private parts touched) than about more mundane or positive experiences (Rudy & Goodman, 1991; Schaaf, Alexander, & Goodman, 2008). In some studies, even 4-year-old children's rates of false affirmation to abuse-related questions were extremely low (Rudy & Goodman, 1991).

A classic study by Leichtman and Ceci (1995) examined preschoolers' memories for a benign event. The children, ages 3 to 6 years old, experienced a visitor (i.e., a confederate) come into their day-care center who was introduced as "Sam Stone." He stayed a short time and then waved goodbye as he left. Children were subsequently interviewed weekly for 10 successive weeks to examine effects of misinformation (i.e., false events) and stereotypes (i.e., negative information about Sam) on children's reports. Children were assigned to one of four interview conditions: (1) a control condition (i.e., contained no misleading questions), (2) a suggestive condition (i.e., contained false misinformation, "Did you see Sam rip the book?"), (3) a stereotype condition (i.e., contained no misleading questions, but negative information about Sam was presented to the children before his visit: "Sam is always getting into accidents and breaking things"); and (4) a stereotype coupled with suggestion condition. Ten weeks after Sam Stone's visit, children were given a final interview in which they were asked to freely recall his visit and then to respond to probed questions that included previously suggested material (i.e., that Sam had ripped a book apart during his initial day-care visit).

Results indicated that, at the final interview, younger children acquiesced to interviewers' previous false suggestions more frequently than did older children, both in free recall and in response to probed questions. However, children in the control condition, who were not subjected to stereotypes or suggestions in the weeks following Sam's visit, were less likely to incorporate false events (e.g., Sam had ripped a book) or report stereotypes (e.g., Sam is clumsy) than were children

in the suggestion, stereotype, or suggestion and stereotyped conditions. The latter of these conditions proved to be where children made the most errors: They were more likely to include stereotyped information during their free recall accounts as well as acquiesce when asked probed questions.

Although these data indicate that children succumb to suggestion when misleading questions are asked, it is difficult to know whether *the memory* of the event has changed or whether *the report* of the memory has changed. That is, when children incorporate suggestions in their reports, does this occur because they are experiencing pressure from the interviewer or because their memory of the event has been distorted? This can be a crucial legal issue.

In a similar vein, do these studies accurately portray interviewer–interviewee conversational nuances that characterize forensic interviews with children? Gilstrap and Ceci (2005) addressed this concern by highlighting that most of the laboratory studies assessing children’s suggestibility do so by way of structured interviews in which all the questions are predetermined by the researchers; these interviews are imposed to ensure the scientific merit (i.e., internal validity) of the research. Results from studies that use structured interviews may not apply to forensic interviews wherein interviewers typically are not supplied with a standardized set of questions. Rather these interviews are driven not only by the interviewer’s agenda but also by the child’s report.

Gilstrap and Ceci (2005) argued that these child-to-adult influences are important to address. In their study, 3- to 7-year-olds were interviewed about an event the children had experienced in the previous week. Interviewers questioned the children about the event and were supplied with descriptions of things that might or might not have occurred, similar to the context in which a forensic interviewer might operate when interviewing a child suspected of being abused. In contrast to children succumbing or agreeing with interviewer suggestion as found in several studies (e.g., Leichtman & Ceci, 1995; Poole & Lindsay, 1995), children in the Gilstrap and Ceci (2005) study were more likely to respond to misleading questions with denial. Instead of interviewer bias predicting children’s acquiescence, the children’s own behavior preceding the misleading question was more strongly predictive of whether they succumbed to suggestion. These findings were obtained by a novel approach of analyzing children’s reports, as they occurred in a transactional exchange throughout the interview, rather than considering only the immediate antecedent (i.e., interviewer’s misleading question) of a child’s error. Such statistical designs appear to be particularly ecologically valid as applied to forensic interviews, although more research is needed to validate these findings and tease apart additional effects that children’s reports may have on the type of questions interviewers ask.

It would be an error to assume that empirical studies using structured interviews are flawed. Researchers should embrace multiple approaches to fully understand conditions that minimize or exacerbate children’s suggestibility. And there may be multiple suggestive influences on children. Garven and colleagues contended that it

is not only misleading questions that influence adults' and children's suggestibility but the additive factors of reinforcement, social pressure, and imagery (Garven, Wood, & Malpass, 2000).

In the study by Garven et al. (2000) that employed techniques modeled after those observed in the McMMartin case interviews, children were questioned about a visit from Paco Perez, a confederate they had witnessed the previous week. Children, ages 5 to 7 years, were interviewed about mundane and unusual events that allegedly occurred during Paco's visit using one of four approaches: (1) reinforcement (i.e., praise when they agreed and disappointment when they disagreed with interviewer suggestion), (2) co-witness (e.g., "The other kids say Paco took them to a farm. Did Paco take you to a farm?"), (3) reinforcement and co-witness, or (4) a control condition. Here additive factors of reinforcement and social pressure (i.e., effect of co-witnesses) were included to mimic the dynamics in the McMMartin preschool case. When interviewed about mundane events, children answered yes to 35% of the misleading questions when they received "reinforcement" from the interviewer compared to a 12% rate when children did not receive reinforcement. This gap between groups widened when children were asked about unusual events; children answered yes to 51% misleading questions compared to 5%, respectively. Moreover, children succumbed to suggestion when they were told that other children reported these same details, yet only for mundane events. Co-witness accounts and reinforcement did not significantly interact, but this research demonstrates that interviewer bias may be conveyed not only through misleading questions but also through such factors as reinforcement. Note that reinforcement that is contingent on a child's response is different from rapport building between interviewer and child eyewitness, which decreases children's suggestibility, particularly when recalling distressing events (e.g., Quas & Lench, 2007).

As researchers attempt to replicate real-world circumstances, some have acknowledged that the person to whom children most often disclose certain crimes (e.g., child sexual abuse) is a nonoffending parent, typically mothers. Few parents have training in interviewing child eyewitness, yet their collection of their children's statements holds forensic significance for whether children's reports will be seen as believable. Therefore, researchers should study the veracity of eyewitness statements when children disclose to a familiar person, such as a parent.

Goodman, Sharma, Thomas, and Considine (1995) addressed this concern by examining whether 4-year-olds' reports were influenced by who interviewed them: Mother- and stranger-interviewers were told that the children had participated in play activities with a researcher, with half of the interviewers receiving falsely biasing information about these activities. Interviewers who received biased information asked more misleading questions than those who did not receive such information. This in turn influenced the children's reports, but only when the interviewer was a stranger: Children made more errors and were less accurate in their reports to strangers but remained accurate in their reports to their mothers. Similar results have been reported elsewhere (e.g., S. Jackson & Crockenberg, 1998).

These findings suggest that children, when comfortable and familiar with the interviewer, correct errors and resist suggestion more easily than with a stranger. These findings offer further support for the importance of rapport building between the interviewer and child eyewitness as well as researchers examining the full range of ecologically valid factors that may influence children's suggestibility: Research on the effects of misleading questions should address not only *what* is asked but also *how* it is asked and *by whom*.

REPEATED INTERVIEWS

In the forensic context, children are often interviewed repeatedly. For example, first responders, police detectives, social workers, prosecuting and defense attorneys, clinicians, and judges may all need to question child eyewitnesses. It is therefore important to determine whether repetition has deleterious, harmless, or positive effects on the accuracy of children's reports.

There are several reasons to suspect that repeated interviews may increase errors in children's reports, especially if misinformation is included in the interviews. For example, SM theory predicts that with each additional interview, children will have more difficulty monitoring whether information or details were actually experienced or were suggested in previous interviews (e.g., Malloy & Quas, 2009; Poole & Lindsay, 1995). In the study described earlier (Leichtman & Ceci, 1995), children were repeatedly interviewed about Sam Stone's visit to their day care. Children who were given repeated erroneous information made more errors than children in the control condition who were only questioned suggestively during the last interview. The researchers interpreted these findings as evidence that repeated erroneous information (whether encountered during the interview process or elsewhere) can negatively affect children's reports, whereas children who are interviewed without suggestions or stereotypes can be capable of providing accurate reports.

In contrast, however, others argue that repeated interviews (even those with misleading questions) do not necessarily have negative effects on children's reports and, under certain conditions, that they actually may assist children in denying new false information by solidifying accurate memories reported previously (e.g., Goodman & Quas, 2008). Repetition of a report may permit children to rehearse their accounts and strengthen or create lasting memory traces, thereby decreasing their chances of succumbing to suggestion or answering misleading questions incorrectly. Quas et al. (2007) found support for this notion: Children, ages 3 and 5, experienced a nonstressful event (playing alone in a room) and were subsequently interviewed either once (after 3 weeks) or three times (weekly up to 3 weeks) about the event. Additionally, biased interviewer manipulations (e.g., providing misinformation through comments and misleading questions) were included in the experimental design. Children who were interviewed once after a 3-week delay made the most errors, particularly when questioned by a biased interviewer.

Of special interest, children who were interviewed multiple times, even when misleading comments were provided and misleading questions were asked, maintained accuracy and reported *fewer* errors compared to children who received one misleading interview after a delay. This finding suggests that suggestibility effects are more problematic when children's initial memories are weak. Researchers should therefore avoid overgeneralized assumptions that repeated interviews compromise children's memory accuracy; instead, these findings should enlighten debates on the complexity of factors influencing children's reports and their interactive or culminating effects (delay since the event, number of previous interviews, exposure to misinformation, etc.)

INDIVIDUAL DIFFERENCES IN CHILDREN'S MEMORY AND SUGGESTIBILITY

Considerable attention has been paid to individual difference predictors, aside from age, of children's memory and suggestibility. In legal cases, the question is typically whether the child witness before the court is likely to be accurate, not whether children of a certain age in general tend to be accurate. Thus, being able to determine whether a particular child is accurate is of considerable legal interest. Unfortunately, in research studies, even when significant correlations are uncovered, the predictors account for relatively little variability in performance and thus are not particularly informative for the courts in evaluating a specific child's accuracy. Here we review a few individual differences that have been examined in child witness research.

INTELLIGENCE

As a possible individual difference that might be related to the accuracy of children's eyewitness memory, intelligence has captured empirical attention, although the findings are somewhat mixed. Some studies uncover support for intelligence as a predictor of suggestibility (e.g., Chae & Ceci, 2005) whereas others find no support of such a relation (e.g., Geddie, Fradin, & Beer, 2000). Geddie et al. (2000) reported that intelligence was positively associated with children's accuracy and negatively associated with children's suggestibility. However, after controlling for additional factors, such as children's race, age, and socioeconomic status (SES), the relation between intelligence and suggestibility was no longer statistically significant. However, some studies that included both intelligence and other sociodemographic variables still found that intelligence significantly predicted children's suggestibility (e.g., McFarlane, Powell, & Dudgeon, 2002). For example, McFarlane et al. (2002) examined 220 preschoolers' suggestibility in relation to their intelligence using a standardized suggestibility measure (i.e., Video Suggestibility Scale for Children). Children with higher intelligence were less suggestible than children with lower

intelligence. Intelligence accounted for 6% of the variance observed in children's suggestibility scores, whereas SES, although still significant, accounted for less than 3% of the variance.

In their meta-analytic review of individual difference factors that may influence children's suggestibility, Bruck and Melnyk (2004) concluded that studies that report a significant association between intelligence and suggestibility include a wide range of variability in intelligence scores among young children, mostly preschool age (e.g., Chae & Ceci, 2005; McFarlane et al., 2002; Young, Powell, & Dudgeon, 2003), whereas studies finding no such relation tend to include primarily older children of average or above-average intelligence (e.g., Burgwyn-Bailes et al., 2001; Eisen et al., 2007; but see Roebbers & Schneider, 2001). Additionally, Bruck and Melnyk (2004) argued that studies that have found that intelligence relates to children's memory reports and suggestibility include individuals with developmental delays. Thus, intelligence appears to be somewhat predictive of the accuracy of children's reports, but primarily when studies include developmentally delayed individuals compared to individuals scoring in the normal ranges of intelligence.

It may be that intelligence can influence a younger child's report but be less influential in the case of older children of average or above-average intelligence (Bruck & Melnyk, 2004). For older children at least, this suggests that the relations between intelligence and memory and suggestibility are nonlinear: Intelligence levels that are below average will be associated with increased suggestibility whereas intelligence levels that are average or above average will have no significant relation to an individual's suggestibility. This idea was tested by Gignac and Powell (2006), who examined 158 children's intelligence and suggestibility. There was a nonlinear quadratic relation between children's intelligence and their suggestibility performance. That is, the relation between intelligence and suggestibility was linear until a certain point (i.e., an IQ score of 105). Around this mid-average range, the relation plateaued. Intelligence was significantly associated with suggestibility only when intelligence was low. Individuals with particularly lower intelligence may be more suggestible; however, intelligence is unrelated to suggestibility in persons of average to above-average intelligence.

VERBAL ABILITY

Although age is linked with verbal abilities, there are wide variations in verbal abilities even when controlling for age statistically or comparing children who are the same age. One might expect that because reporting of past events and responding to interviewer questions in the forensic context are, in effect, verbal conversations, children who have a better understanding of communicative nuances, receptive and expressive language skills, and bigger vocabularies may be better able to articulate their experiences than children who have more limited verbal abilities. Furthermore, verbally skilled children should be able to identify incorrect suggestions put forth by the interviewer (Pezdek & Roe, 1995).

Generally, research reveals that children's proficiencies in communication assist them in being more accurate in recalling past experiences and more resistant to suggestions from others. For example, Clarke-Stewart, Malloy, and Allhusen (2004) examined 5-year-olds' verbal abilities in relation to the suggestibility of their reports about a staged event witnessed approximately 9 months previously. Children were interviewed by a "leading" interviewer, who asked leading questions (e.g., "Where did he touch you?") and accepted any response put forth, or by a "suggestive" interviewer who asked follow-up questions when the children's answers did not conform to the interviewer's suggestions (e.g., Interviewer: "And did he touch your bottom?" Child: "Nope." Interviewer: "No? Didn't he touch you on your behind?"). Children's verbal abilities (i.e., a composite score of their receptive and expressive language abilities) and aptitude for effectively communicating with adults (i.e., the child's score from the Adaptive Language Inventory questionnaire) were inversely related to overall suggestibility, as indexed by children's responses to misleading interview questions. Additionally, children's abilities to effectively communicate with adults were also inversely related to whether they succumbed to interviewer suggestions in both the "leading" and "suggestive" experimental groups. These results imply that children with greater verbal skills were more accurate and less suggestible than their peers.

However, in other studies, no significant associations emerged between verbal skill and suggestibility (e.g., Bright-Paul & Jarrold, 2009; Quas & Lench, 2007), and the opposite effect has even been reported, with verbal skills being positively associated with children's increased suggestibility (e.g., Kulkofsky & Klemfuss, 2008). For example, in Kulkofsky and Klemfuss's (2008) study, 3- to 5-year-olds who produced high-quality verbal autobiographic memory narratives were more suggestible during their memory interview. That is, children who provided more elaborative details about their own past experiences incorporated more interviewer suggestions when reporting details about a staged event.

This inconsistency could in part be due to methodological differences in how the *type* of verbal ability (e.g., vocabulary, receptive language, narrative quality) was assessed. For example, Quas and Lench (2007) examined 5- to 6-year-olds' memories of a fear-eliciting video clip they watched the preceding week. Children's verbal abilities were assessed using a receptive vocabulary test (i.e., the Peabody Picture Vocabulary Test [PPVT]). No significant relations were observed between children's scores on the PPVT and their memory accuracy or suggestibility. However, only a single measure of language assessment, primarily one that tapped receptive vocabulary, was employed. Studies that do find significant associations (e.g., Clarke-Stewart et al., 2004) tend to include more global assessments of children's verbal competencies. One such type of assessment is the quality of the narrative reports supplied by children as they recall their past. In this approach, narrative quality may provide a comprehensive view of children's verbal fluency as it pertains to the accuracy of their reports (Peterson, 2012).

DISCLOSURE OF ABUSE

In the following section, we discuss various factors associated with disclosure. These include reasons children may delay or avoid disclosing abuse, types of emotions children typically express during disclosure, and possible determinants of lying during disclosure.

FACTORS AFFECTING DISCLOSURE

Children often delay disclosing sexual abuse (London, Bruck, Ceci, & Shuman, 2005). In fact, in an analysis of 10 retrospective studies on the topic, London et al. (2005) reported that an average of only 39% of adults who reported being sexually abused indicated they had disclosed the abuse during childhood. Children may be hesitant to disclose even when there is photographic or video evidence (Paz-Alonso, Ogle, et al., 2013). Research on children's disclosure and factors that impede it have focused on socio-motivational factors rather than cognitive characteristics, and reveal that sexually abused children delay disclosure for a number of reasons, which may vary as a function of age (e.g., Goodman-Brown, Edelstein, Goodman, Jones, & Gordon, 2003; Malloy, Lyon, & Quas, 2007).

Children may be less likely to disclose if a parent or caregiver is the perpetrator (Lyon, Ahern, Malloy, & Quas, 2010). Lyon and his colleagues examined the responses of 299 maltreated and nonmaltreated 4- to 9-year-olds to hypothetical vignettes of children experiencing misdeeds by an authority figure (i.e., parent, teacher, or stranger). After each vignette, children were asked whether the child in the vignette should report the wrongdoing to another adult or comply with the authority figure to keep it a secret. Children were more likely to disclose when the authority figure was a stranger than a parent. Additionally, younger maltreated children endorsed disclosure less frequently than nonmaltreated children, particularly when a parent was described as the perpetrator.

Correlates of delays in disclosure have been identified in studies of maltreated children (e.g., Goodman-Brown et al., 2003). In child sexual abuse cases, older children are more likely to delay disclosure when they fear that negative consequences (e.g., punishment, embarrassment) would ensue for them or a third party (who was not an offender), such as a sister. Older children were also more likely to report that they perceived themselves to be partially responsible for the abuse. Malloy, Brubacher, and Lamb (2011) uncovered a similar finding in their examination of factors predicting children's and adolescents' (5 to 13 years old) disclosure of sexual abuse: Children who mentioned fear of negative consequences (e.g., punishment, embarrassment) delayed their disclosures, again only when these consequences would have affected them or innocent others (e.g., siblings), not the perpetrator. However, according to the London et al. (2005) review, although it may seem logical that when children are more fearful after sexual abuse, they are less likely to disclose, in fact, children are more likely to disclose when sexual abuse involves fear and physical injury.

EMOTIONAL EXPRESSION DURING DISCLOSURE

There are apparently numerous misunderstandings among laypeople about how children disclose sexual abuse. For example, demeanor during disclosure often is used to assess the credibility of child victims (Myers, Redlich, Goodman, Prizmich, & Imwinkelried, 1999; Regan & Baker, 1998). Yet research indicates that, during forensic interviews, children appear less upset than might be expected. Wood, Orsak, Murphey, and Cross (1996) analyzed child sexual abuse interviews of children age 2 to 11 years for emotions and attentiveness. Children were most often rated as relaxed or neutral; however, girls did display more sadness than boys. Sayfan, Mitchell, Goodman, Eisen, and Qin (2008) rated 3- to 16-year-olds' affect while they were being interviewed regarding maltreatment (for which there was corroborating evidence). Although children alleging sexual abuse were more likely to be upset than children alleging other forms of abuse, most children were not rated as emotional, and most (98%) did not cry. Compared to children who suffered less severe abuse, children who suffered more severe maltreatment (e.g., over a longer period of time) were less likely to display intense emotion. In a direct comparison of disclosers and nondisclosers of abuse, Katz et al. (2012) examined interviews of 40 victims of abuse for whom there was substantial external evidence that the abuse had occurred. Although positive emotions decreased as the interview progressed, there was no difference in positive emotions displayed by disclosers and nondisclosers. Castelli, Sayfan, Mitchell, Culver, and Goodman (2005) also found that, during forensic interviews, positive emotions decrease from the rapport building to the disclosure phase of the interview. Additionally, both disclosers and nondisclosers have shown instances of negative emotions such as shame and guilt when interviewed as adults about child sexual abuse (Bonanno, Noll, Putnam, O'Neill, & Trickett, 2003). The overall picture indicates that, during abuse interviews, children show less emotion than possibly expected. However, they do, on average, show some negative emotions, and their displays of emotion vary over the course of the interviews and as a function of abuse severity.

LYING

When a child discloses information to authorities, concerns may be raised about the child's honesty. There are many legal situations in which children may be motivated to lie (e.g., if coached not to reveal a parental transgression). An antisocial lie is specifically meant to protect oneself from harm or to provide oneself with personal gain (Talwar & Lee, 2008a). Although children's antisocial lies can certainly play a role in legal cases (e.g., when the child is accused of delinquent acts), when the child is a witness or victim, concerns usually center on the child being coached to knowingly make a false allegation (e.g., in a custody case, to accuse the father of sexual abuse so that the child can stay with the mother) or protect a culprit who has asked the child to lie or keep a secret.

Lying appears to develop through three main stages: (1) beginning to make untrue statements at around 2 to 3 years of age, (2) lying to conceal one's own transgressions at 3 to 4 years of age, and (3) being able to maintain lies at 7 to 8 years of age (Talwar & Lee, 2008a). The development of children's lie-telling is related to Theory of Mind ability (Talwar, Gordon, & Lee, 2007) and executive functioning (Talwar & Lee, 2008b). Of interest, most research has not shown a relation between understanding of lying and actual lying to conceal a transgression (London & Nuñez, 2002; Talwar, Lee, Bala, & Lindsay, 2002). However, understanding of lying was related to lie telling when it involved concealing a transgression by a parent (Talwar, Lee, Bala, & Lindsay, 2004) or when lie telling involved maltreated children making a false allegation (Lyon & Dorado, 2008).

A forensically relevant question with respect to children's lying is whether the lie is to conceal a transgression committed by someone emotionally close to the children. Children may be unlikely to lie to conceal the transgression of a relative stranger, although younger children are more likely to do so than older children (Pipe & Wilson, 1994). Talwar et al. (2004) examined whether 3- to 11-year-old children would cover up their parents' transgression and found that most children would disclose the transgression even with the parent in the room at the time of the interview. However, the children were more likely to lie to protect a parent if they themselves could not be blamed for the transgression. In a study of 3- to 6-year-olds, Bottoms, Goodman, Schwartz-Kenney, and Thomas (2002) reported that older children told to keep a secret by their parents were more likely to withhold information about their parents' transgressions than older children who were not told by their parents to keep the acts a secret. Although such findings provide important insight about children's lying behavior, it should be noted that the transgressions in these studies were quite mild (e.g., breaking a toy). The dynamics could well change for lies about more serious acts, such as child maltreatment and other types of violent crime.

Children can and do lie to protect themselves and to protect others. When children are lying in such a manner, can these lies be detected? Most studies indicate that adults are not accurate at detecting children's lies (Crossman & Lewis, 2006; Goodman et al., 2006) and that they are no better at detecting children's lies than adults' lies (Goodman et al., 2006). Coached lies by older children may be particularly difficult to detect (K. L. Warren, Dodd, Raynor, & Peterson, 2012). However, Nysse-Carris, Bottoms, and Salerno (2011) found that adults could detect 3- to 6-year-old children's lies about their parents' transgressions at above chance levels. A goal for future research is to better explain the difficulty in detecting children's lying.

FACE RECOGNITION AND CHILDREN'S EYEWITNESS IDENTIFICATIONS

Eyewitness identifications are crucial in the forensic context. Legal authorities need to know who committed the crime in question. Often when children are victims of or bystanders to crime, they may be presented with a photo lineup or a live lineup

and asked to identify the culprit. Considerable research has examined factors that affect children's eyewitness identification accuracy.

At the start of this chapter, we described the case of David Wiggins, who had been convicted of raping a 14-year-old girl and who spent over 20 years in prison before being DNA exonerated. The girl had identified him after seeing his picture in a mugshot booklet. Research psychologists would say that the identification procedure was problematic because of phenomena termed "unconscious transference" and "mugshot commitment" (Deffenbacher, Bornstein, & Penrod, 2006). The victim tentatively identified Wiggins in a photo lineup, then subsequently identified him in a live lineup. Because he was the only person who was in both lineups, it is likely that Wiggins looked familiar to the victim because she had already seen him in the photo lineup and unconsciously (mentally) transferred him from the photo lineup to the crime. When she saw him in the courtroom and identified him again, he was even more familiar to her, because she had identified him twice before. (For further discussion of these and related issues, see Ross, Tredoux, and Malpass, Chapter 17 this volume.)

Child bystanders (as opposed to child victims) may also be asked to make identifications. In Northern California, a 10-year-old girl heard some men walking up the stairs in her apartment building and caught a glimpse of their faces as she peeked out the window. She looked again about 30 minutes later when she heard them leaving a neighbor's apartment. The neighbor was later found dead. The police investigation revealed several suspects, so the officers showed the 10-year-old girl photo lineups that included the suspects. She readily identified them. Her testimony, which was corroborated by other evidence, was crucial at the murder trial.

FACE PROCESSING AND RECOGNITION

Before discussing how research can inform police lineup procedures for child witnesses, we first briefly explain the theoretical underpinnings and mechanisms for face processing and face recognition in children and adults, which can affect crucial cognitive processes involved in picking out a suspect from a lineup.

As is true for memory generally, facial recognition improves as children age (Lawrence et al., 2008). Lawrence et al. investigated the relation between face recognition and age in 6- to 16-year-olds. Participants viewed 50 pictures and rated whether faces were "Nice or Not Nice." Then participants were shown the faces again, only this time each face seen previously was paired with a new, yet similar, face. Participants indicated which face was previously viewed. Facial identification improved from age 6 to 10 years, remained stable up to age 13 years, and then increased again. This age effect remained even after controlling for intelligence, which was also related to face identification accuracy.

What if the 10-year-old girl described earlier was Caucasian and the men coming up the stairs were of a different ethnicity? In fact, that was the case. With age, cross-racial face identification becomes less accurate than same-race facial identification.

Goodman et al. (2007) reported that 5- to 7-year-olds did not show a cross-racial decrement but that older children and adults recognized own-race faces better than cross-race faces. However, infants as young as 9 months old may exhibit an own-race bias (Liu et al., 2011), suggesting that even within the first year, humans may have conceptual models specialized for processing in-group faces. The result is that an out-group bias for facial processing may begin quite early in life.

There has been considerable theorizing about the cross-racial effect. Of interest to this debate, similar effects have been found for cross-gender and cross-age face recognition accuracy.

One of the theoretical accounts for facial biased processing, the contact hypothesis, states that having face-to-face contact with social partners (e.g., caregivers, siblings) enables the construction of mental models for processing features (Scherf & Scott, 2012; Sporer, 2001). Commonalities among faces lead to effective processing strategies when similar features are observed in new social partners. As a result, a Caucasian individual, who has been primarily exposed to other Caucasian faces, will be able to strategically process facial features of individuals who are Caucasian but would struggle processing an Asian individual's face. This difficulty in out-group face processing presumably occurs because mental models were previously established for the Caucasian face, permitting additional elaboration of individualized facial characteristics; however, the Asian face required additional processing. A similar process is believed to occur for same gender and same race faces. For the child eyewitness, these results suggest that, for older children and perhaps younger ones as well, race effects may influence eyewitness testimony if the victim and perpetrator are of different races and the victim has not been meaningfully and sufficiently exposed to members of the perpetrator's race. Similar influences are also at play for identification of faces representing different genders and ages from the eyewitness (Scherf & Scott, 2012).

EYEWITNESS IDENTIFICATION AND LINEUP FAIRNESS

In face identification studies, where theoretical issues are tested, children and adults typically are briefly exposed to photographs of faces both at study and at test. However, in reality, eyewitnesses observe actual people live over extended periods of time, which likely affects encoding and memory. It has therefore been important to examine eyewitness identification in more realistic studies. Such research reveals that, by the age of about 5 or 6, children are often as accurate as adults in identifying people with whom they have interacted when presented with target-present lineups (i.e., lineups that include the target person—the “culprit”). However, when the actual culprit is not in the lineup (i.e., “target-absent” lineups), even older children (e.g., 10-year-olds) are more likely than adults to falsely identify an individual and less likely to report that the target person is not included in the lineup (Pozzulo & Lindsay, 1999). Target-absent lineups are generally more difficult compared to target-present lineups, even for adults (e.g., R. C. L. Lindsay,

Pozzulo, Craig, Lee, & Corber, 1997; Pozzulo & Dempsey, 2006), as individuals may assume that, because they are viewing a lineup, the perpetrator is included. Some individuals have a tendency to guess. This is a serious concern for criminal investigators, as children and adults may assume that the task is to identify one of the choices rather than to judge whether the perpetrator is present at all (Beresford & Blades, 2006; Humphries, Holliday, & Flowe, 2012).

Eyewitness identification procedures have received heavy criticism for improper or suggestive methods that could taint an eyewitness's memory (e.g., Wells & Loftus, 2003; Wells & Quinlivan, 2009). Research has identified several factors that promote the fairness of lineups, such as foils appearing similar to the suspect, clear pre-lineup instructions (e.g., "The perpetrator may or may not appear here"), and avoiding use of authority approval or confirmation (Wells & Loftus, 2003). Given children's greater suggestibility compared to adults, such factors may be particularly important when children are subjected to lineup procedures.

This research has also revealed that simultaneous lineups, wherein the suspect is viewed simultaneously among other foils, have the potential to be suggestive. For example, if the "suspect" is truly innocent yet the foils in a simultaneous lineup do not match the eyewitness description, then the "suspect" is more likely to be falsely identified by mere fact that the he or she looks *the most* like the actual suspect compared with foils who do not match the eyewitness's description.

Instead of simultaneous lineups, it is suggested that investigators show eyewitnesses a sequential lineup, with the eyewitness making a yes/no judgment for each person. In this way, eyewitnesses are more likely to compare the photograph to their memory representation of the culprit and make an "absolute" rather than "relative" judgment. This method decreases rates of picking the person who looks most like the perpetrator (Wells & Loftus, 2003). Similar success has been demonstrated using an elimination paradigm wherein individuals are asked to eliminate individuals from the lineup who do not match the target (Pozzulo & Lindsay, 1999). This approach has demonstrated some success, with decreasing instances of false identifications in both target-present and target-absent lineups, particularly for children and adolescents (e.g., Pozzulo, Dempsey, & Crescini, 2009; Pozzulo & Lindsay, 1999).

A study by Beresford and Blades (2006) tested whether 6- to 10-year-olds would benefit from receiving instructions prior to presentation of both target-present and target-absent standard lineups. After witnessing a staged theft, children were instructed that they would view a series of individuals (with format of individuals varying between either static pictures or live videos) and that the thief *may or may not* be included. This latter portion of the instructions was heavily emphasized, with children receiving instructions that false identifications would lead to negative consequences for that individual. Children who received such cautions, regardless of age, made fewer false identifications than children not so instructed. Additionally, these instructions did not decrease the rate of accuracy for children in target-present conditions. That is, the instructions decreased children's false identifications on

target-absent lineups while at the same time did not adversely influence children's identifications on target-present arrays. These results are quite promising. However, beneficial effects of instruction were not observed for elimination lineups.

Researchers should address such discrepancies to identify the most effective means of administering lineups to children. Moreover, instructions to improve lineup performance in young preschoolers (e.g., 3-year-olds) still are sorely needed.

JURORS' REACTIONS TO CHILD EYEWITNESSES

When children testify in court at jury trials, judges and jurors have the difficult task of assessing the accuracy of the children's testimony. Characteristics of children and of the jurors themselves may affect whether children are believed or not. In some types of cases, such as in child sexual abuse trials, jurors claim they consider child-victim characteristics to be the most important evidence (Myers et al., 1999). It is thus important to understand legal decision makers' reactions to child witnesses. Much of the research in this area has focused on child victim-witnesses in sexual abuse trials. This is in part because, at least in the United States, children are most likely to testify in criminal proceedings when they are victims of sexual abuse (Goodman, Quas, Bulkley, & Shapiro, 1999).

In mock jury research, two of the most widely studied victim characteristics have been age and gender of the victim. The effects of victim age on jury decisions differ depending on whether witness competence or witness honesty is emphasized (Bottoms, Golding, Stevenson, Wiley, & Yozwiak, 2007). Young children are viewed as less credible than older children and adults when issues of competence (e.g., memory or suggestibility) are stressed because young children are viewed as less accurate in remembering (Leippe & Romanczyk, 1989). In contrast, younger children are viewed as more credible when issues of honesty and sincerity are emphasized (Ross, Miller, & Moran, 1987) as well as issues of sexual naïveté (Bottoms & Goodman, 1994). As such, younger children may be viewed as more credible than older ones when they are the victims of sexual abuse (Ross, Jurden, Lindsay, & Keeney, 2003). The perception of sexual naïveté also leads mock jurors to view children as more credible when they are testifying about sexual abuse as opposed to nonsexual offenses (McCauley & Parker, 2001). Nightingale (1993, experiment 2) varied the age of the victim in a corroborated sexual abuse scenario from age 6 to age 14, and, as age increased, victims were viewed as less credible. By manipulating the age of the victims incrementally, Wright, Hanoteau, Parkinson, and Tatham (2010) were able to determine more precisely at what ages perceptions of children's honesty and cognitive abilities changed. Observers' perceptions of memory reliability increased from ages 3 to 6 but then plateaued whereas perceptions of honesty increased until age 11 but then decreased.

In real trials, child victim gender has not been consistently found to be as influential on jury decision making as child age (Myers et al., 1999). However, there

is some evidence that law enforcement officers and rape crisis counselors once believed male victims of sexual abuse to be weak and effeminate and likely to enjoy sexual assault (Donnelly & Kenyon, 1996). It is hoped that such views have changed in the years since that study was undertaken. Most statistically significant effects of child victim gender on mock jurors' decisions are through interactions with either defendant gender or juror gender. For example, with teenage victims, opposite-gender sexual abuse is viewed as less abusive than same-gender sexual abuse (e.g., Dollar, Perry, Fromuth, & Holt, 2004). Male but not female mock jurors are affected by victim gender (e.g., Clark & Nightingale, 1997), and women mock jurors favor girl victims while male mock jurors favor boy victims (ForsterLee, Horowitz, Ho, ForsterLee, & McGovern, 1999). In some cases, victim gender interacts with both mock juror gender and defendant gender. Quas, Bottoms, Haecerich, and Nysse-Carris (2002) found that female defendants were less likely to be convicted by male mock jurors when the victims of the abuse were boys. However, regarding juror gender, numerous studies reveal that female mock jurors are more empathetic to child victims overall and more likely to believe them in child sexual abuse cases (Bottoms et al., 2007).

Interest in the effects of race and ethnicity has increased in psychology and the law, and this is also true in research on jurors' reactions to child victims and witnesses. A common stereotype of minorities is of increased sexual promiscuity and experience (Alley, 2012). As a result, jurors may view sexual abuse of minority children as less heinous and might hold the victim more responsible. There have been few studies examining these questions directly, but evidence so far has shown that mock jurors hold Caucasian victims compared to African American or Hispanic American victims as less responsible for their abuse (Bottoms, Davis, & Epstein, 2004).

Victim demeanor is especially important in jurors' impressions of witnesses, including children. It is considered so relevant by the courts that jury instructions frequently direct jurors to consider facial expressions when judging the credibility of a witness (A. J. Williams, 2008). Adults who had just served jury duty in child sexual abuse trials rated facial expressions and demeanor as being important in forming impressions regarding the child victims' believability when providing testimony (Myers et al., 1999). In the first part of a two-part study by Regan and Baker (1998), most undergraduate mock jurors expected a 6-year-old female victim of paternal sexual abuse to cry (81%) and show fear (67%) when confronting the defendant. In the second investigation, undergraduates read a scenario in which the 6-year-old female victim was either calm or crying. The crying victim was rated as significantly more credible than the calm victim, and the defendant was more likely to be judged guilty in the crying scenario.

Golding, Fryman, Marsil, and Yozwiak (2003) varied scenarios of female child sexual abuse by the age of the victim (6 or 15 years old) and the emotion displayed (calm, teary, or hysterical crying) in both the text and in line drawings. Participants found the teary victim to be more believable than the calm or the hysterical victim

and rendered more guilty verdicts in this scenario. The findings point to mild negative emotion as being considered an appropriate level of distress for female victims of child sexual abuse to display. Wessel and Melinder (2012) also varied the age of the female victim (11 or 13 years old) while varying the type of emotion using child actors who portrayed emotions classified as sad, neutral, angry, or positive during a mock police interview. Participants who viewed the videos perceived the crying victim to be the most credible, followed by the neutral victim, the positive victim, and then the angry victim. Overall research on emotions in legal contexts indicates that adult expectations of children's emotional displays influence how children are judged.

Most studies of jury decision making involve mock jurors, and, as such, methodological issues limit the generalizability of the findings. Although participants are often undergraduate students who are not representative of actual jurors, the use of undergraduate students as opposed to community members has been validated (Bornstein, 1999), and investigators have found little variation between the decisions made by students and jurors (Diamond, 1997). These studies have not examined the differences between undergraduates' and actual jurors' decisions when children's emotion was the variable of interest. A potentially important difference between students and actual jurors or community samples is that students are less likely to be parents, and being a parent might influence the perceptions of children who are victims of abuse.

Presentation medium is also a concern in jury decision-making research. Methods of presentation vary from written transcripts, to videos, to live performances, and although presentation medium has not led to widely different outcomes in mock jury trials (Bornstein, 1999), the use of video or live performances might be of more importance for research examining the impact of emotions on decisions as a video might better convey the victims' emotional displays. Research that relies on surveys of mock or real jurors does not usually include the deliberation process that takes place in actual trials. Furthermore, factors that influence individual judgments do not always occur in studies of jury decisions (Devine, Clayton, Dunford, Seying, & Pryce, 2001). The methodological limitations of jury decision-making research should temper the interpretations of the results and their extrapolations to the real world. However, this line of research has been invaluable in both identifying the factors that are most likely to influence actual jurors and the areas in which juror expectations contrast with actual child behaviors.

JURORS' REACTIONS TO EXPERT WITNESSES IN CHILD ABUSE CASES

Under certain conditions, psychologists and other professionals may be asked to provide expert testimony in child witness cases (Myers, 1993b). There is growing consensus that expert witnesses can help jurors evaluate the accuracy of children's testimony (e.g., Bottoms et al., 2007; Quas, Thompson, & Clarke-Stewart, 2005). Nonetheless, it is still a matter of controversy as to the conditions under which

expert witnesses significantly affect jurors' decision making and verdicts (e.g., Lyon, 2002) and whether their testimony is simply unnecessary or even detrimental. For example, if jurors already know about children's suggestibility and the factors that lead them to make errors, expert testimony on these topics may result in unwarranted levels of skepticism on the part of jurors (e.g., Lyon, 2002; see also Buck, London, & Wright, 2011).

Most of the studies on expert testimony that we discuss here concern child sexual abuse cases or "repressed memory" cases involving allegations of past child sexual abuse. These studies typically present undergraduate students with vignettes of trials. However, in a few cases, the researcher analyzed actual legal cases (e.g., Read, Connolly, & Welsh, 2006). In an analysis of 29 studies on expert witnesses, Kwartner (2007) demonstrated that there was a small but significant effect on jurors' verdicts when the testimony was evaluative (pertaining to the witness at hand) rather than educative (pertaining to general scientific information that could assist jurors in their decision making) in nature. Additionally, Gabora, Spanos, and Joab (1993) found that jurors were more in favor of conviction when expert psychological testimony was specific to the child sexual abuse case at issue (e.g., the expert submitted clinical evidence based on an assessment of the alleged victim) rather than general (e.g., the expert offered social framework testimony about rates of child sexual abuse more generally; see also Kovera, Gresham, Borgida, Gray, & Regan, 1997). Conversely, the findings of Crowley, Callaghan, and Ball (1994) suggest that jurors who hear social framework testimony based on scientific literature rate child victims of sexual abuse significantly more favorably in terms of their memory ability, resistance to suggestion, and ability to differentiate fact from fiction, and therefore are more likely to reach a verdict to convict, relative to those jurors who do not hear such testimony.

There are numerous additional ways that expert witnesses might influence jurors' decision making. For example, Klettke, Graesser, and Powell (2010) found that the coherence of an expert witness's testimony and the strength of the evidence presented positively affected the credibility of child sexual abuse victims and the likelihood of a guilty verdict. Of interest, the credentials of the expert had no effect on the mock jurors' perceptions or decision making. Nuñez, Gray, and Buck (2012) reported that providing mock jurors with multiple reasons to doubt hearsay evidence in a child sexual abuse case influenced perceptions of hearsay witness credibility and verdict decisions more than providing reasons for doubting such testimony. Thus, jurors may need more than one reason to alter their verdict behavior.

Expert testimony could also counteract jurors' misunderstanding of children's memory and suggestibility. Quas, Thompson, et al. (2005) examined whether expert witnesses are needed to educate jury-eligible adults or if such adults already have adequate knowledge about children's memory and suggestibility. Participants did not recognize the powerful influence of stereotypic inductions on children's accuracy as eyewitnesses. It may be that, even if individuals are knowledgeable and skeptical about some aspects of children's suggestibility, they are less aware

of adverse effects of subtle but still-influential interview manipulations. There was considerable variability in individuals' knowledge about children's eyewitness abilities; individuals had both inaccurate and accurate beliefs, which could indicate that expert testimony is potentially important (Quas, Thompson, et al., 2005).

Buck et al. (2011) evaluated whether expert testimony helped mock jurors distinguish between well-conducted interviews and poorly conducted interviews of children in descriptions of sexual abuse investigations. Mock jurors who were provided expert testimony were more likely to render guilty verdicts if the interview quality was good versus poor. However, without such testimony, the mock jurors did not consider forensic interview quality when reaching their verdicts. These findings suggest that expert testimony on interview methods may help laypeople make more informed decisions about the reliability of children's reports.

Finally, although expert testimony might influence the outcomes of trials involving child witnesses, the effects seem to fluctuate depending on the party that uses the testimony and the facts of the case at hand: defense alone, prosecutor alone, or concurrent opposing experts. Some researchers claim that expert testimony in a trial is associated with decreased rates of convictions as compared to trials in which there were no experts at all (Griffith, Libkuman, Dodd, Shafir, & Dickenson, 2002; Read et al., 2006). When defense experts alone testify, there are even more reductions in guilty verdicts, as studied in "historic" child sexual abuse cases involving adults who testify about their childhood abuse experiences (Read et al., 2006). This is probably explained by the fact that defense experts most likely raise issues of reasonable doubt regarding the reliability of eyewitness memory (Connolly, Price, & Read, 2006). However, when the juries are exposed to competing experts, there is no overall effect on trial verdicts (Read et al., 2006).

ACCOMMODATIONS FOR CHILD WITNESSES

Concern about child witnesses experiencing secondary trauma while testifying has resulted in the development of court modifications and system interventions to reduce such trauma (Hall & Sales, 2008). Protective services and legal interventions to ameliorate child witness trauma alleviate children's emotional distress, promote the well-being of child victims, and support children in providing reliable testimony (Malloy, Mitchell, Block, Quas, & Goodman, 2006; Troxel et al., 2009). These services and interventions include out-of-court testimony, such as closed-circuit television (CCTV) and hearsay testimony via third party interviewers, as well as the use of multidisciplinary service centers or child advocacy centers (CACs).

OUT-OF-COURT TESTIMONY

Like adults, children experience both pre- and posttestimony anxiety, especially if they have to give testimony in front of defendants in open court in criminal actions (e.g., Goodman et al., 1992). To help alleviate potential trauma for child witnesses,

statements made outside of the courtroom (e.g., through interviews with third parties such as forensic interviews, video recordings, or CCTV) are sometimes permitted. Hearsay testimony allows children's out-of-court statements (e.g., to their mothers or other family members) to be considered evidence in court proceedings on behalf of child victims, at least under certain conditions. In some cases, forensic interviews with child witnesses may be video recorded and presented as hearsay evidence to the court. CCTV allows a child to give evidence outside the courtroom in front of a camera, with the image and sound immediately relayed to the courtroom for viewing while the child undergoes direct and cross-examination.

Hearsay. Hearsay evidence is defined as "a statement, other than one made by the declarant while testifying at the trial or hearing, offered in evidence to prove the truth of the matter asserted" (Federal Rules of Evidence 801, 1975). Although use of hearsay evidence may help protect a child witness from secondary trauma and reduce anxiety associated with confronting an alleged perpetrator, there are risks of adults misrepresenting the words and testimony of child witnesses (Lamb, Orbach, Sternberg, Hershkowitz, & Horowitz, 2000). Thus, presentation in court of video-recorded child forensic interviews, a structured form of hearsay, permits triers of fact to hear children's out-of-court statements directly.

Nevertheless, because the eyewitness statements are made outside of the courtroom, the eyewitness is not subject to cross-examination, and thus such hearsay may not meet the court's "indicia of reliability" (Goodman et al., 2006). In fact, a United States Supreme Court decision (*Crawford v. Washington*, 2004) challenged the admissibility at trial of out-of-court testimonials (e.g., children's videotaped statements to law enforcement), unless the children also appeared as witnesses. This decision results in greater judicial pressure for children to testify live in court before video-recorded forensic interviews can be admitted. Related concerns center on several assumptions about the value and significance of defendants' abilities to confront witnesses, including (a) the stress of testifying on the stand and facing the accused improves the accuracy of witness testimony; (b) the jury's ability to detect deception is impeded unless the witness testifies live in court; and (c) the introduction of out-of-court statements may negatively bias the jury's perception of the defendant and adversely affect case outcome. Using mock trial and juror interview studies, researchers continue to examine these issues in attempts to find a reasonable balance between the rights of child witnesses and the accused (e.g., Landström, Granhag, & Hartwig, 2007; McAuliff & Kovera, 2012).

The assumption that jurors can best detect the truthfulness or deceptiveness of a witness when a witness is testifying live in front of them is not supported by the prevailing research literature. As mentioned previously, the ability of adults to distinguish between deceptive and truthful adults (e.g., Malone & DePaulo, 2001) and children (e.g., Edelman, Lutten, Ekman, & Goodman, 2006) is often not much better than chance. Further, a meta-analysis by Aamodt and Custer (2006) suggests

that most legal professionals, such as judges and law enforcement officers, may be no more accurate in detecting deception than untrained individuals. In an examination of live testimony, videotaped testimony, and adult hearsay testimony, Goodman et al. (2006) found that mock jurors had difficulty discerning between accurate and deceptive statements from child witnesses regardless of live or out-of-court testimony.

The format or mode of testimony may be an important determinant of perceived child witness credibility and truthfulness as children who testify live are generally seen more positively or truthful than children who testify outside of court (Landström et al., 2007). In an examination of prospective jurors' expectancies for a child sexual abuse case, McAuliff and Kovera (2012) reported that jurors believed it was easier to determine a child's truthfulness, and fairest to the defendant, when testimony was live in court. These findings support previous research where children testifying live, or more proximal to adult observers, were seen more positively and given greater credibility than children testifying out of court in more distal locations (Goodman et al., 2006; Landström et al., 2007).

For hearsay testimony, Warren, Nuñez, Keeney, Buck, and Smith (2002) found that adults who appear in court to repeat children's statements were viewed as more accurate than children giving firsthand, live testimony. In that regard, the hearsay testimony effectiveness may depend on the status or perceived credibility of the adult (e.g., doctor, law enforcement officer) who testifies about the child's out-of-court statements (Ross, Lindsay, & Marsil, 1999). Further research is warranted to determine the impact of hearsay evidence on judicial processes as well as on the well-being of child witnesses.

CCTV. The use of out-of-court testimony for child witnesses is widely accepted and established in a number of countries. In Australia, New Zealand, and the United Kingdom, a two-way closed circuit television (CCTV) approach is employed, allowing interactive testimony between attorneys and the judge while a child witness is outside of court in a separate room. In the United Kingdom, the videotaped forensic interview serves as direct examination in court, and CCTV is used for cross-examination purposes. In other countries, such as Finland, Norway, and Sweden, child witnesses are video-recorded during preliminary police interviews, and those recordings serve as direct and cross-examination. One-way CCTV is employed at times in the United States although it remains controversial as some argue that it violates the 6th and 14th Amendments of the U.S. Constitution, which provide defendants the right to confront their accusers during criminal trials and to due process, respectively (Hall & Sales, 2008). Following a landmark case in which the U.S. Supreme Court decided in favor of the use of one-way CCTV in child sexual abuse cases under certain conditions (*Maryland v. Craig*, 1990), courts in the United States are being asked to rule on the use of one-way CCTV.

Although the ability to confront a witness is believed to produce more accurate testimony, research has not supported this belief. In examining the effects of

CCTV on mock jurors' perceptions of child witnesses, Goodman et al. (1998) found that 8- to 9-year-olds generally provided more accurate information than 5- to 6-year-olds in both CCTV and open court but that CCTV was associated with reduced suggestibility for the younger children. It has also been argued that CCTV might hamper jurors' abilities to determine truthfulness in child witnesses, but this has not held up in the research literature. Orcutt, Goodman, Tobey, Batterman-Faunce, and Thomas's (2001) research revealed that mock jurors were no better at determining deception when children testified in open court or through CCTV.

One concern about child witnesses testifying through CCTV is the perception of less emotional impact compared to live court testimony (McAuliff & Kovera, 2012). The emotional impact appears to be even less with video-recorded child testimony (Landström, 2008). Orcutt et al. (2001) reported that children testifying via CCTV were seen as less accurate, less believable, less consistent, less confident, less attractive, and less intelligent than children who testified in open court. One rationale for jurors perceiving children in more negative terms when testifying via CCTV is the *vividness effect* (Nisbett & Ross, 1980), which suggests the closer the witness is in proximity and time, the more positively jurors' evaluate the witness. McAuliff and Kovera (2012) propose that negative evaluations of children's testimony given in alternative forms, such as CCTV and video-recording, may be the result of expectancy violations, meaning jurors expect differences in children's verbal and nonverbal behavior as a result of the CCTV accommodation, but those differences actually may not occur.

CHILD ADVOCACY CENTERS

The child advocacy center (CAC) multidisciplinary approach to child forensic interviews is designed to reduce secondary victimization in children by (a) facilitating collaboration between relevant agencies (e.g., child protective services, law enforcement, prosecution, mental health, and medicine), (b) providing child-sensitive interview settings, and (c) limiting the number of interviews a child victim experiences. By providing supportive services to child witnesses, CACs aim to reduce trauma associated with the investigative and legal processes. The 10 core components of a CAC are:

1. Multidisciplinary team
2. Cultural competency and diversity
3. Child forensic interview
4. Victim support and advocacy
5. Medical evaluation
6. Mental health services
7. Case review
8. Case tracking

9. Organizational capacity
10. Child-focused setting (Connell, 2009; S. L. Jackson, 2004, 2012; Newman, Dannenfelser, & Pendleton, 2005)

Evaluations of CACs are promising and suggest they decrease delays between law enforcement reports and indictment dates (Walsh, Lippert, Cross, Maurice, & Davison, 2008), increase access to medical examinations, improve the experience of nonoffending parents during the investigation process, and decrease the level of fear experienced by children during interviews (L. M. Jones, Cross, Walsh, & Simone, 2007). Although the improvements in the treatment of child witnesses and their families are encouraging, the effects of CACs on prosecution outcomes, false allegations, children's disclosure rates, and children's stress reduction are as yet unclear (Saywitz & Camparo, 2009). Data are still emerging relevant to the efficacy of CACs, but the accumulating research suggests CACs are likely to be helpful to child witnesses and families involved in criminal proceedings. CACs are beginning to spread worldwide in countries' efforts to combat crimes against children and ease children's involvement in the legal process (Rodrigues dos Santos & Batista Gonçalves, 2009).

CONCLUSIONS

Children pose many dilemmas for the legal system. Yet to protect children and others from harm and ensure justice, society has little choice but to include child witnesses in legal cases, especially when other evidence is lacking or when the children's testimony plays a key role in a prosecution. The United Nations Convention on the Rights of the Child (UNCRC), which has been ratified by every country in the United Nations except three (including not by the United States), specifies that:

1. [Countries] shall assure to the child who is capable of forming his or her own views the right to express those views freely in all matters affecting the child, the views of the child being given due weight in accordance with the age and maturity of the child.
2. For this purpose the child shall in particular be provided the opportunity to be heard in any judicial and administrative proceedings affecting the child, either directly, or through a representative or an appropriate body, in a manner consistent with the procedural rules of national law (Convention on the Rights of the Child, 1989).

It is clear that many countries in the world are—or soon will be—struggling with how and when to listen to child witnesses in the legal context. Fortunately, psychological science is in an excellent position to make a meaningful and important contribution to this effort.

REFERENCES

- Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar? *Forensic Examiner, 15*, 6–11.
- Alley, D. (2012). *Priming a sexual stereotype: Influences on adults' perceptions of child sexual abuse victims* (Unpublished doctoral dissertation). University of California, Davis.
- Ayers, M. S., & Reder, L. M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. *Psychonomic Bulletin & Review, 5*, 1–21.
- Baker-Ward, L., Hess, T. M., & Flannagan, D. A. (1990). The effects of involvement on children's memory for events. *Cognitive Development, 5*, 55–69.
- Bala, N., Lee, J., & McNamara, E. (2001). Children as witnesses: Understanding their capacities, needs, and experiences. *Journal of Social Distress and the Homeless, 10*, 41–68. doi:10.1023/A:1009429602266
- Baugerud, G. A., & Melinder, A. (2012). Maltreated children's memory of stressful removals from their biological parents. *Applied Cognitive Psychology, 26*, 261–270. doi:10.1002/acp.1817
- Bekerian, D. A., & Bowers, J. M. (1983). Eyewitness testimony: Were we misled? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*, 139–145.
- Beresford, J., & Blades, M. (2006). Children's identification of faces from lineups: The effects of lineup presentation and instructions on accuracy. *Journal of Applied Psychology, 91*, 1102–1113. doi:10.1037/0021-9010.91.5.1102
- Bjorklund, D. F., Dukes, C., & Douglas-Brown, R. (2008). The development of memory strategies. In M. Courage & N. Cowan (Eds.), *The development of memory in infancy* (pp. 145–176). New York, NY: Psychology Press.
- Blandon-Gitlin, I., & Pezdek, K. (2009). Children's memory in forensic contexts: Suggestibility, false memory, and individual differences. In B. L. Bottoms, C. J. Najdowski, & G. S. Goodman (Eds.), *Children as victims, witnesses, and offenders* (pp. 57–80). New York, NY: Guilford Press.
- Block, S. D., Oran, H. S., Oran, D., Baumrind, N., & Goodman, G. S. (2010). Abused and neglected children in court: Knowledge and attitudes. *Child Abuse & Neglect, 34*, 659–670. doi:10.1016/j.chiabu.2010.02.003
- Bonanno, G. A., Noll, J. G., Putnam, F. W., O'Neill, M., & Trickett, P. K. (2003). Predicting the willingness to disclose childhood sexual abuse from measures of repressive coping and dissociative tendencies. *Child Maltreatment, 8*, 302–318. doi:10.1177/1077559503257066
- Bornstein, B. H. (1999). The ecological validity of jury simulation: Is the jury still out? *Law and Human Behavior, 23*, 75–91.
- Bottoms, B. L., Davis, S. L., & Epstein, M. A. (2004). Effects of victim and defendant race on jurors' decisions in child sexual abuse cases. *Journal of Applied Social Psychology, 34*, 1–33. doi:10.1111/j.1559-1816.2004.tb02535.x
- Bottoms, B. L., Golding, J. M., Stevenson, M. C., Wiley, T. R. A., & Yozwiak, J. A. (2007). A review of factors affecting jurors' decisions in child sexual abuse cases. In J. D. Read, D. Ross, M. Togli, & R. Lindsay (Eds.), *The psychology of eyewitness testimony* (pp. 509–544). Mahwah, NJ: Erlbaum.
- Bottoms, B. L., & Goodman, G. S. (1994). Perceptions of children's credibility in sexual assault cases. *Journal of Applied Social Psychology, 28*, 702–732.
- Bottoms, B. L., Goodman, G. S., Schwartz-Kenney, B. M., & Thomas, S. N. (2002). Understanding children's use of secrecy in the context of eyewitness reports. *Law and Human Behavior, 26*, 285–313.

- Bowlby, J. (1982). Attachment and loss: Retrospect and prospect. *American Journal of Orthopsychiatry*, 52, 664–678.
- Brainerd, C. J., Forrest, T. J., Karibian, D., & Reyna, V. F. (2006). Development of the false-memory illusion. *Developmental Psychology*, 42(5), 962–979. doi:10.1037/0012-1649.42.5.962
- Brainerd, C. J., & Reyna, V. F. (1998). Fuzzy-trace and children's false memory. *Journal of Experimental Child Psychology*, 71, 81–129.
- Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, 11, 164–169. doi:10.1111/1467-8721.00192
- Brainerd, C. J., & Reyna, V. F. (2004). Fuzzy-trace theory and memory development. *Developmental Review*, 24, 396–439. doi:10.1016/j.dr.2004.08.005
- Brainerd, C. J., Reyna, V. F., & Ceci, S. J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin*, 134, 343–382. doi:10.1037/0033-2909.134.3.343
- Brainerd, C. J., Reyna, V. F., & Zember, E. (2011). Theoretical and forensic implications of developmental studies of the DRM illusion. *Memory & Cognition*, 39, 365–380. doi:10.3758/s13421-010-0043-2
- Bremner, J. D., Shobe, K. K., & Kihlstrom, J. F. (2000). False memories in women with self-reported childhood sexual abuse: An empirical study. *Psychological Science*, 11, 333–337. doi:10.1037/a0015371
- Bright-Paul, A., & Jarrold, C. (2009). A temporal discriminability account of children's eyewitness suggestibility. *Developmental Science*, 12, 647–661. doi:10.1111/j.1467-7687.2009.00811.x
- Bruck, M., & Melnyk, L. (2004). Individual differences in children's suggestibility: A review and synthesis. *Applied Cognitive Psychology*, 18, 947–996. doi:10.1002/acp.1070
- Brumariu, L. E., & Kerns, K. A. (2010). Parent-child attachment and internalizing symptoms in childhood and adolescence: A review of empirical findings and future directions. *Development and Psychopathology*, 22, 177–203. doi:10.1017/S0954579409990344
- Buck, J. A., London, K., & Wright, D. B. (2011). Expert testimony regarding child witnesses: Does it sensitize jurors to forensic interview quality? *Law and Human Behavior*, 35, 1521–64. doi:10.1007/s10979-010-9228-2
- Burgwyn-Bailes, E., Baker-Ward, L., Gordon, B. N., & Ornstein, P. A. (2001). Children's memory for emergency medical treatment after one year: The impact of individual difference variables on recall and suggestibility. *Applied Cognitive Psychology*, 15, S25–S48. doi:10.1002/acp.833
- Castelli, P., Sayfan, L., Mitchell, E. B., Culver, M., & Goodman, G. S. (2005). Children's emotional expressions when disclosing abuse. In L. C. Malloy & J. A. Quas (Chairs), *Disclosing child abuse: Insight from forensic and courtroom settings*. Paper presented at the American Psychology–Law Society Conference, La Jolla, CA.
- Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*, 113, 403–439.
- Ceci, S. J., & Bruck, M. (2006). Children's suggestibility: Characteristics and mechanisms. *Advances in child development and behavior* (Vol. 34, pp. 247–284). San Diego, CA: Elsevier Academic Press.
- Ceci, S. J., Papierno, P. B., & Kulkofsky, S. (2007). Representational constraints on children's suggestibility. *Psychological Science*, 6, 503–509. doi:10.1111/j.1467-9280.2007.01930.x

- Ceci, S. J., Ross, D. F., & Toglia, M. P. (1987). Suggestibility of children's memory: Psycholegal implications. *Human Development, 116*, 38–49. doi:10.1037/0096-3445.116.1.38
- Cederborg, A.-C., Lamb, M. E., & Laurell, O. (2007). Delay of disclosure, minimization, and denial when the evidence is unambiguous: A multi-victim case. In M.-E. Pipe, M. E. Lamb, Y. Orbach, & A.-C. Cederborg (Eds.), *Child sexual abuse: Disclosure, delay, and denial* (pp. 159–174). Hillsdale, NJ: Erlbaum.
- Chae, Y., & Ceci, S. J. (2005). Individual differences in children's recall and suggestibility: The effect of intelligence, temperament, and self-perceptions. *Applied Cognitive Psychology, 19*(4), 383–407. doi:10.1002/acp.1094
- Chae, Y., Goodman, G. S., & Edelstein, R. S. (2011). Autobiographical memory development from an attachment perspective: The special role of negative events. In J. Benson (Ed.), *Advances in child development and behavior* (Vol. 40, pp. 1–49). Burlington, MA: Academic Press. doi:10.1016/B978-0-12-386491-8.00001-3
- Chae, Y., Ogle, C. M., & Goodman, G. S. (2009). Remembering negative childhood experiences: In J. A. Quas & R. Fivush (Eds.), *Emotion and memory development: Biological, cognitive, and social considerations* (pp. 3–27). New York, NY: Oxford University Press.
- Christianson, S.-Å. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin, 112*, 284–309. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1454896>
- Clark, H. L., & Nightingale, N. N. (1997). When jurors consider recovered memory cases: Effects of victim and juror gender. *Journal of Offender Rehabilitation, 25*, 87–104. doi:10.1300/J076v25n03_06
- Clarke-Stewart, K. A., Malloy, L. C., & Allhusen, V. D. (2004). Verbal ability, self-control, and close relationships with parents protect children against misleading suggestions. *Applied Cognitive Psychology, 18*(8), 1037–1058. doi:10.1002/acp.1076
- Connell, M. (2009). The Child Advocacy Center model. In K. Kuehnle & M. Connell (Eds.), *The evaluation of child sexual abuse allegations: A comprehensive guide to assessment and testimony* (pp. 423–449). Hoboken, NJ: Wiley.
- Connolly, D. A., Price, H. L., & Read, J. D. (2006). Predicting expert social science testimony in criminal prosecutions of historic child sexual abuse. *Legal and Criminological Psychology, 11*, 55–74. doi:10.1348/135532505X50312
- Cordón, I. M., Pipe, M.-E., Sayfan, L., Melinder, A., & Goodman, G. S. (2004). Memory for traumatic experiences in early childhood. *Developmental Review, 24*, 101–132. doi:10.1016/j.dr.2003.09.003
- Crawford v. Washington, 541 US 36. (2004).
- Crossman, A. M., & Lewis, M. (2006). Adults' ability to detect children's lying. *Behavioral Sciences and the Law, 24*, 703–715. doi:10.1002/bsl
- Crowley, M. J., Callaghan, M. G. O., & Ball, P. J. (1994). The juridical impact of psychological expert testimony in a simulated child sexual abuse trial. *Law and Human Behavior, 18*, 89–105.
- Deffenbacher, K. A., Bornstein, B. H., & Penrod, S. D. (2006). Mugshot exposure effects: Retroactive interference, mugshot commitment, source confusion, and unconscious transference. *Law and Human Behavior, 30*, 287–307. doi:10.1007/s10979-006-9008-1
- Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior, 28*(6), 687–706. doi:10.1007/s10979-004-0565-x

- Dent, H. R., & Stephenson, G. M. (1979). An experimental study of the effectiveness of different techniques of questioning child witnesses. *British Journal of Social & Clinical Psychology, 18*, 41–51. doi:10.1111/j.2044-8260.1979.tb00302.x
- Devine, D. J., Clayton, L. D., Dunford, B. B., Seying, R., & Pryce, J. (2001). Jury decision making: 45 years of empirical research on deliberating groups. *Psychology, Public Policy, and Law, 7*, 622–727. doi:10.1037//1076-8971.7.3.622
- Diamond, S. S. (1997). Illuminations and shadows from jury simulations. *Law and Human Behavior, 21*, 561–571. doi:10.1023/A:1024831908377
- Dollar, K. M., Perry, A. R., Fromuth, M. E., & Holt, A. R. (2004). Influence of gender roles on perceptions of teacher/adolescent student sexual relations. *Sex Roles, 50*, 91–101.
- Donnelly, D. A., & Kenyon, S. (1996). “Honey, we don’t do men”: Gender stereotypes and the provision of services to sexually assaulted males. *Journal of Interpersonal Violence, 11*, 441–448. doi:10.1177/088626096011003009
- Eakin, D. K., Schreiber, T. A., & Sergent-Marshall, S. (2003). Misinformation effects in eyewitness memory: The presence and absence of memory impairment as a function of warning and misinformation accessibility. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 813–825.
- Edelstein, R. S., Ghetti, S., Quas, J. A., Goodman, G. S., Alexander, K. W., Redlich, A. D., & Cordón, I. M. (2005). Individual differences in emotional memory: Adult attachment and long-term memory for child sexual abuse. *Personality and Social Psychology Bulletin, 31*, 1537–1548. doi:10.1177/0146167205277095
- Edelstein, R. S., Luten, T. L., Ekman, P., & Goodman, G. S. (2006). Detecting lies in children and adults. *Law and Human Behavior, 30*, 1–10. doi:10.1007/s10979-006-9031-2
- Eisen, M. L., Goodman, G. S., Qin, J., Davis, S. L., & Crayton, J. (2007). Maltreated children’s memory: Accuracy, suggestibility, and psychopathology. *Developmental Psychology, 43*, 1275–1294. doi:10.1037/0012-1649.43.6.1275
- Farrar, M. J., & Goodman, G. S. (1992). Developmental changes in event memory. *Child Development, 63*, 173–187.
- Federal Rules of Evidence 801 (1975). Pub L. 93-595, 88 Stat 1926.
- Felitti, V. J. (2002). The relationship of adverse childhood experiences to adult health: Turning gold into lead. *The Permanente Journal, 6*, 44–47.
- Foa, E. G., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99*, 20–35.
- ForsterLee, R., Horowitz, I. A., Ho, R., ForsterLee, L., & McGovern, A. (1999). Community members’ perceptions of evidence: The effects of gender in a recovered memory civil trial. *Journal of Applied Psychology, 84*, 484–95. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10504892>
- Gabora, N. J., Spanos, N. P., & Joab, A. (1993). The effects of complainant age and expert psychological testimony in a simulated child sexual abuse trial. *Law and Human Behavior, 17*, 103–119. doi:10.1007/BF01044540
- Garven, S., Wood, J. M., & Malpass, R. S. (2000). Allegations of wrongdoing: The effects of reinforcement on children’s mundane and fantastic claims. *Journal of Applied Psychology, 85*, 38–49. doi:10.1037//0021-9010.85.1.38
- Geddie, L., Fradin, S., & Beer, J. (2000). Child characteristics which impact accuracy of recall and suggestibility in preschoolers: Is age the best predictor? *Child Abuse & Neglect, 24*, 223–35. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10695517>

- Ghetti, S. (2008). Rejection of false events in childhood: A metamemory account. *Current Directions in Psychological Science*, *17*, 16–20. doi:10.1111/j.1467-8721.2008.00540.x
- Gignac, G. E., & Powell, M. B. (2006). A direct examination of the nonlinear (quadratic) association between intelligence and suggestibility in children. *Applied Cognitive Psychology*, *20*, 617–623. doi:10.1002/acp.1213
- Gilstrap, L. L., & Ceci, S. J. (2005). Reconceptualizing children's suggestibility: Bidirectional and temporal properties. *Child Development*, *76*, 40–53. doi:10.1111/j.1467-8624.2005.00828.x
- Golding, J. M., Fryman, H. M., Marsil, D. F., & Yozwiak, J. A. (2003). Big girls don't cry: The effect of child witness demeanor on juror decisions in a child sexual abuse trial. *Child Abuse & Neglect*, *27*, 1311–1321. doi:10.1016/j.chiabu.2003.03.001
- Goodman, G. S., Bottoms, B. L., Rudy, L., Davis, S. L., & Schwartz-Kenney, B. M. (2001). Effects of past abuse experiences on children's eyewitness memory. *Law and Human Behavior*, *25*, 269–298. doi:10.1023/A:1010797911805
- Goodman, G. S., Bottoms, B. L., Schwartz-Kenney, B. M., & Rudy, L. (1991). Children's testimony about a stressful event: Improving children's reports. *Journal of Narrative & Life History*, *1*, 69–99.
- Goodman, G. S., Hirschman, J. E., Hepps, D., & Rudy, L. (1991). Children's memory for stressful events. *Merrill-Palmer Quarterly*, *37*, 109–157.
- Goodman, G. S., Myers, J. E. B., Qin, J., Quas, J. A., Castelli, P., Redlich, A. D., & Rogers, L. (2006). Hearsay versus children's testimony: Effects of truthful and deceptive statements on jurors' decisions. *Law and Human Behavior*, *30*, 363–401. doi:10.1007/s10979-006-9009-0
- Goodman, G. S., & Quas, J. A. (2008). Repeated interviews and children's memory: It's more than just how many. *Current Directions in Psychological Science*, *17*, 386–390. doi:10.1111/j.1467-8721.2008.00611.x
- Goodman, G. S., Quas, J. A., Batterman-Faunce, J. M., Riddlesberger, M. M., & Kuhn, J. (1997). Children's reactions to and memory for a stressful event: Influences of age, anatomical dolls, knowledge, and parental attachment. *Applied Developmental Science*, *1*, 54–75. doi:10.1207/s1532480xads0102
- Goodman, G. S., Quas, J. A., Bulkley, J., & Shapiro, C. (1999). Innovations for child witnesses: A national survey. *Psychology, Public Policy, and Law*, *5*, 255–281. doi:10.1037//1076-8971.5.2.255
- Goodman, G. S., Quas, J. A., & Ogle, C. M. (2009). Child maltreatment and memory. *Annual Review of Psychology*, *61*, 325–351. doi:10.1146/annurev.psych.093008.100403
- Goodman, G. S., & Reed, R. S. (1986). Age differences in eyewitness testimony. *Law and Human Behavior*, *10*, 317–332.
- Goodman, G. S., Sayfan, L., Lee, J. S., Sandhei, M., Walle-Olsen, A., Magnussen, S., Pezdek, K., & Arredondo, P. (2007). The development of memory for own- and other-race faces. *Journal of Experimental Child Psychology*, *98*, 233–242. doi:10.1016/j.jecp.2007.08.004
- Goodman, G. S., Sharma, A., Thomas, S. F., & Considine, M. G. (1995). Mother knows best: Effects of relationship status and interviewer bias on children's memory. *Journal of Experimental Child Psychology*, *60*, 195–228.
- Goodman, G. S., Taub, E. P., Jones, D.P.H., England, P., Port, L., K., Rudy, L., & Prado, L. (1992). Testifying in criminal court: Emotional effects on child sexual assault victims. *Monographs of the Society for Research in Child Development*, *57*, i, 1–159.
- Goodman, G. S., Tobey, A. E., Batterman-Faunce, J. M., Orcutt, H. K., Thomas, S. F., Shapiro, C., & Sachsenmaier, T. (1998). Face-to-face confrontation: Effects of closed-circuit

- technology on children's eyewitness testimony and juror's decisions. *Law and Human Behavior*, 22, 165–203.
- Goodman-Brown, T. B., Edelman, R. S., Goodman, G. S., Jones, D.P.H., & Gordon, D. S. (2003). Why children tell: A model of children's disclosure of sexual abuse. *Child Abuse & Neglect*, 27(5), 525–540. doi:10.1016/S0145-2134(03)00037-1
- Greenhoot, A. F., McCloskey, L., & Glisky, E. (2005). A longitudinal study of adolescents' recollections of family violence. *Applied Cognitive Psychology*, 19, 719–743. doi:10.1002/acp.1103
- Griffith, J. D., Libkuman, T. M., Dodd, J., Shafir, Z., & Dickenson, J. J. (2002). The effects of expert testimony on mock jurors' decision making and memory. *American Journal of Forensic Psychology*, 2, 69–80.
- Hall, S. R., & Sales, B. D. (2008). *Courtroom modification for child witnesses*. Washington, DC: American Psychological Association. doi:10.1037/11808-000
- Harris, L. S., Goodman, G. S., Augusti, E. M., Chae, Y., & Alley, D. (2009). Children's resistance to suggestion. In K. Kuehnle & M. Connell (Eds.), *The evaluation of child sexual abuse allegations* (pp. 181–202). Hoboken, NJ: Wiley.
- Hayne, H., & Simcock, G. (2009). Memory development in toddlers. In M. L. Courage & N. Cowan (Eds.), *The development of memory in infancy and childhood* (pp. 43–68). New York, NY: Psychology Press.
- Hirst, W., Phelps, E. A., Buckner, R. L., Budson, A. E., Cuc, A., Gabrieli, J. D. E., Johnson, M. K., . . . Vaidya, C. J. (2009). Long-term memory for the terrorist attack of September 11: Flashbulb memories, event memories, and the factors that influence their retention. *Journal of Experimental Psychology: General*, 138, 161–176. doi:10.1037/a0015527
- Howe, M. L. (2011). *The nature of early memory: An adaptive theory of the genesis and development of memory*. New York, NY: Oxford University Press.
- Howe, M. L., Cicchetti, D., Toth, S. L., & Cerrito, B. M. (2004). True and false memories in maltreated children. *Child Development*, 75, 1402–1417. doi:10.1111/j.1467-8624.2004.00748.x
- Howe, M. L., & Otgaar, H. (2013). Proximate mechanisms and the development of adaptive memory. *Current Directions in Psychological Science*.
- Humphries, J. E., Holliday, R. E., & Flowe, H. D. (2012). Faces in motion: Age-related changes in eyewitness identification performance in simultaneous, sequential, and elimination video lineups. *Applied Cognitive Psychology*, 26, 149–158. doi:10.1002/acp.1808
- Innocence Project, The. (2012). Retrieved from <http://www.innocenceproject.org>
- Jackson, S., & Crockenberg, S. (1998). A comparison of suggestibility in 4-year-old girls in response to parental or stranger misinformation. *Journal of Applied Developmental Psychology*, 19, 527–542. doi:10.1016/S0193-3973(99)80054-9
- Jackson, S. L. (2004). A USA national survey of program services provided by child advocacy centers. *Child Abuse & Neglect*, 28, 411–421. doi:10.1016/j.chiabu.2003.09.020
- Jackson, S. L. (2012). Results from the Virginia Multidisciplinary Team Knowledge and Functioning Survey: The importance of differentiating by groups affiliated with a child advocacy center. *Children and Youth Services Review*, 34, 1243–1250. doi:10.1016/j.childyouth.2012.02.015
- Johnson, M. K., & Foley, M. A. (1984). Differentiating fact from fantasy: The reliability of children's memory. *Journal of Social Issues*, 40, 33–50. doi:10.1111/j.1540-4560.1984.tb01092.x
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114, 3–28. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8346328>

- Jones, D.P.H., & Krugman, R. D. (1986). Case report: Can a three-year-old child bear witness to her sexual assault and attempted murder? *Child Abuse & Neglect*, *10*, 253–258.
- Jones, L. M., Cross, T. P., Walsh, W. A., & Simone, M. (2007). Do children's advocacy centers improve families' experiences of child sexual abuse investigations? *Child Abuse & Neglect*, *31*, 1069–1085. doi:10.1016/j.chiabu.2007.04.006
- Katz, C., Hershkowitz, I., Malloy, L. C., Lamb, M. E., Atabaki, A., & Spindler, S. (2012). Non-verbal behavior of children who disclose or do not disclose child abuse in investigative interviews. *Child Abuse & Neglect*, *36*(1), 12–20. doi:10.1016/j.chiabu.2011.08.006
- Keeney, K. S., Amacher, E., & Kastanakis, J. A. (1992). The court prep group: A vital part of the court process. In H. Dent & R. Flin (Eds.), *Children as witnesses* (pp. 201–209). West Sussex, England: Wiley.
- Kensinger, E. A. (2009). Remembering the details: Effects of emotion. *Emotion Review*, *1*, 99–113. doi:10.1177/1754073908100432
- Kim, J., & Cicchetti, D. (2010). Longitudinal pathways linking child maltreatment, emotion regulation, peer relations, and psychopathology. *Journal of Child Psychology and Psychiatry*, *51*, 706–716. doi:10.1111/j.1469-7610.2009.02202.x
- Klettke, B., Graesser, A. C., & Powell, M. B. (2010). Expert testimony in child sexual abuse cases: The effects of evidence, coherence and credentials on juror decision making. *Applied Cognitive Psychology*, *24*, 481–494. doi:10.1002/acp
- Kovera, M. B., Gresham, A. W., Borgida, E., Gray, E., & Regan, P. C. (1997). Does expert psychological testimony inform or influence juror decision making? A social cognitive analysis. *Journal of Applied Psychology*, *82*, 178–191.
- Kulkofsky, S., & Klemfuss, J. Z. (2008). What the stories children tell can tell about their memory: Narrative skill and young children's suggestibility. *Developmental Psychology*, *44*, 1442–1456. doi:10.1037/a0012849
- Kwartner, P. P. (2007). *The impact of educative and evaluative expert witness testimony on trial outcome: A meta-analytic review* (Unpublished doctoral dissertation). Sam Houston State University, Texas.
- Lamb, M. E., Malloy, L. C., & La Rooy, D. J. (2011). Setting realistic expectations: Developmental characteristics, capacities, and limitations. In M. E. Lamb, D. J. La Rooy, L. C. Malloy, & C. Katz (Eds.), *Children's testimony: A handbook of psychological research and forensic practice*. Chichester, UK: Wiley. doi:10.1002/9781119998495.ch2
- Lamb, M. E., Orbach, Y., Sternberg, K. J., Hershkowitz, I., & Horowitz, D. (2000). Accuracy of investigators' verbatim notes of their forensic interviews with alleged child abuse victims. *Law and Human Behavior*, *24*, 699–708. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11105480>
- Landström, S. (2008). *CCTV, live, and videotapes: How presentation mode affects the evaluation of witnesses*. Doctoral dissertation, University of Gotenborg, Sweden.
- Landström, S., Granhag, P. A., & Hartwig, M. (2007). Children's live and videotaped testimonies: How presentation mode affects observers' perception, assessment and memory. *Legal and Criminological Psychology*, *12*, 333–347. doi:10.1348/135532506X133607
- Lawrence, K., Bernstein, D., Pearson, R., Mandy, W., Campbell, R., & Skuse, D. (2008). Changing abilities in recognition of unfamiliar face photographs through childhood and adolescence. *Journal of Neuropsychology*, *2*, 27–45. doi:10.1348/174866407X231074
- Leander, L., Christianson, S.-Å., & Granhag, P. A. (2007). A sexual abuse case study: Children's memories and reports. *Psychiatry, Psychology and Law*, *14*, 120–129. doi:10.1375/pplt.14.1.120

- Leander, L., Granhag, P. A., & Christianson, S.-Å. (2005). Children exposed to obscene phone calls: What they remember and tell. *Child Abuse & Neglect*, *29*, 871–888. doi:10.1016/j.chiabu.2004.12.012
- Leichtman, M. D., & Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology*, *31*, 768–778. doi:10.1037//0012-1649.31.5.758
- Leippe, M. R., & Romanczyk, A. (1989). Reactions to child (versus adult) eyewitness. *Law and Human Behavior*, *13*, 103–132.
- Levine, L. J., & Edelman, R. S. (2009). Emotion and memory narrowing: A review and goal-relevance approach. *Cognition & Emotion*, *23*, 833–875. doi:10.1080/02699930902738863
- Lindsay, D. S. (2002). Children's source monitoring. In H. L. Wescott, G. M. Davies, & R. H. C. Bull (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 83–98). New York, NY: Wiley.
- Lindsay, R. C. L., Pozzulo, J. D., Craig, W., Lee, K., & Corber, S. (1997). Simultaneous lineups, sequential lineups, and showups: Eyewitness identification decisions of adults and children. *Law and Human Behavior*, *21*, 391–404. doi:10.1023/A:1024807202926
- Liu, S., Quinn, P. C., Wheeler, A., Xiao, N., Ge, L., & Lee, K. (2011). Similarity and difference in the processing of same- and other-race faces as revealed by eye tracking in 4- to 9-month-olds. *Journal of Experimental Child Psychology*, *108*, 180–9. doi:10.1016/j.jecp.2010.06.008
- Loftus, E. F. (1975). Leading questions and the eyewitness report. *Memory & Cognition*, *7*, 368–374.
- Loftus, E. F. (1979). Reactions to blatantly contradictory information. *Memory & Cognition*, *7*, 368–374. doi:10.3758/BF03196941
- Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals*, *25*, 720–725.
- London, K., Bruck, M., Ceci, S. J., & Shuman, D. W. (2005). Disclosure of child sexual abuse: What does research tell us about the ways that children tell? *Psychology, Public Policy, and Law*, *11*, 194–226. doi:10.1037/1076-8971.11.1.194
- London, K., & Nuñez, N. (2002). Examining the efficacy of truth/lie discussions in predicting and increasing the veracity of children's reports. *Journal of Experimental Child Psychology*, *83*, 131–147.
- Lyon, T. D. (2002). Expert testimony on the suggestibility of children: Does it fit? In B. L. Bottoms, M. B. Kovera, & B. D. McAuliff (Eds.), *Children, social science, and the law* (pp. 378–411). New York, NY: Cambridge University Press.
- Lyon, T. D. (2011). Assessing the competency of child witnesses: Best practice informed by psychology and law. In M. E. Lamb, D. J. La Rooy, L. C. Malloy, & C. Katz (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 69–85). Sussex, UK: Wiley-Blackwell.
- Lyon, T. D., Ahern, E. C., Malloy, L. C., & Quas, J. A. (2010). Children's reasoning about disclosing adult transgressions: Effects of maltreatment, child age, and adult identity. *Child Development*, *81*(6), 1714–28. doi:10.1111/j.1467-8624.2010.01505.x
- Lyon, T. D., & Dorado, J. S. (2008). Truth induction in young maltreated children: The effects of oath-taking and reassurance on true and false disclosures. *Child Abuse & Neglect*, *32*(7), 738–48. doi:10.1016/j.chiabu.2007.08.008
- Lyon, T. D., & Saywitz, K. J. (2006). From post-mortem to preventive medicine: Next steps for research on child witnesses. *Journal of Social Issues*, *62*, 833–861.

- Malloy, L. C., Brubacher, S. P., & Lamb, M. E. (2011). Expected consequences of disclosure revealed in investigative interviews with suspected victims of child sexual abuse. *Applied Developmental Science, 15*, 8–19. doi:10.1080/10888691.2011.538616
- Malloy, L. C., Johnson, J. L., & Goodman, G. S. (2013). Children's memory and event reports: The current state of knowledge and best practice. *Journal of Forensic Social Work, 3*, 1–30.
- Malloy, L. C., Lyon, T. D., & Quas, J. A. (2007). Filial dependency and recantation of child sexual abuse allegations. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*(2), 162–70. doi:10.1097/01.chi.0000246067.77953.f7
- Malloy, L. C., Mitchell, E. B., Block, S. D., Quas, J. A., & Goodman, G. S. (2006). Children's eyewitness memory: Balancing children's needs and defendants' rights when seeking the truth. In M. P. Toglia, J. D. Read, D. F. Ross, & R. Lindsay (Eds.), *Handbook of eyewitness psychology: Volume 1: Memory for events* (pp. 545–574). Mahwah, NJ: Erlbaum.
- Malloy, L. C., & Quas, J. A. (2009). Children's suggestibility: Areas of consensus and controversy. In K. Kuehnle & M. Connell (Eds.), *The evaluation of child sexual abuse allegations* (pp. 267–297). Hoboken, NJ: Wiley.
- Malone, B. E., & DePaulo, B. M. (2001). Measuring sensitivity to deception. In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal sensitivity* (pp. 103–124). Mahwah, NJ: Erlbaum.
- Maryland v. Craig, 497 U.S. 836. (1990).
- Mather, M., & Sutherland, M. (2009). Disentangling the effects of arousal and valence on memory for intrinsic details. *Emotion Review, 1*, 118–119. doi:10.1177/1754073908100435
- McAuliff, B. D., & Kovera, M. B. (2012). Do jurors get what they expect? *Psychology, Crime & Law, 18*, 27–47. doi:10.1080/1068316X.2011.613391
- McCauley, M. R., & Parker, J. F. (2001). When will a child be believed? The impact of the victim's age and juror's gender on children's credibility and verdict in a sexual-abuse case. *Child Abuse & Neglect, 25*, 523–539. doi:10.1016/S0145-2134(01)00224-1
- McFarlane, F., Powell, M. B., & Dudgeon, P. (2002). An examination of the degree to which IQ, memory performance, socio-economic status and gender predict young children's suggestibility. *Legal and Criminological Psychology, 7*(2), 227–239. doi:10.1348/135532502760274729
- McWilliams, K., Bederian-Gardner, D., Hobbs, S. D., Bakanosky, S., & Goodman, G. S. (2012). Children's eyewitness memory and suggestibility: Revisiting Ceci and Bruck's (1993) review. In A. M. Slater & P. C. Quinn (Eds.), *Developmental psychology: Revisiting the classic studies* (pp. 101–117). Thousand Oaks, CA: Sage.
- McWilliams, K., Harris, L. S., & Goodman, G. S. (2012). *Child maltreatment and eyewitness memory*. Poster presented at the meeting of American Psychology-Law Society Conference, San Juan, Puerto Rico.
- McWilliams, K., Narr, R., Goodman, G. S., Ruiz, S., & Mendoza, M. (2013, January). Children's memory for their mother's murder: Accuracy, suggestibility, and resistance to suggestion. *Memory*. [Epub ahead of print]
- Melinder, A., Baugerud, G. A., Ovenstad, K., & Goodman, G. S. (2013). Children's memories of removal: A test of attachment theory. *Journal of Traumatic Stress, 26*(1), 125–133.
- Morris, G., & Baker-Ward, L. (2007). Fragile but real: Children's capacity to use newly acquired words to convey preverbal memories. *Child Development, 78*, 448–458. doi:10.1111/j.1467-8624.2007.01008.x
- Myers, J. E. B. (1993a). A call for forensically relevant research. *Child Abuse & Neglect, 17*, 573–579.

- Myers, J. E. B. (1993b). Expert testimony regarding child sexual abuse. *Child Abuse & Neglect*, *17*, 175–185.
- Myers, J. E. B., Redlich, A. D., Goodman, G. S., Prizmich, L. P., & Imwinkelried, E. (1999). Jurors' perceptions of hearsay in child sexual abuse cases. *Psychology, Public Policy, and Law*, *5*(2), 388–419. doi:10.1037//1076-8971.5.2.388
- Nairne, J. S., Thompson, S. R., & Pandeirada, J. N. S. (2007). Adaptive memory: Survival processing enhances retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 263–273. doi:10.1037/0278-7393.33.2.263
- Neisser, U., & Nicole, H. (1992). Phantom flashbulbs: False recollections of hearing the news about Challenger. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of "flashbulb" memories* (pp. 9–31). New York, NY: Cambridge University Press.
- Newman, B. S., Dannenfelser, P. L., & Pendleton, D. (2005). Child abuse investigations: Reasons for using child advocacy centers and suggestions for improvement. *Child and Adolescent Social Work Journal*, *22*, 165–180. doi:10.1007/s10560-005-3416-9
- Nightingale, N. N. (1993). Juror reactions to child victim witnesses: Factors affecting trial outcome. *Law and Human Behavior*, *17*, 679–694. doi:10.1007/BF01044689
- Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgments*. Englewood Cliffs, NJ: Prentice Hall.
- Norgate, S. H., & Littleton, K. S. (2011). Children's memories for events relating to treatment for eye cancer: Influence of age at loss of eye. *Infant Mental Health Journal*, *32*, 563–577. doi:10.1002/imhj.
- Núñez, N., Gray, J., & Buck, J. A. (2012). Educative expert testimony: A one-two punch can affect jurors' decisions. *Journal of Applied Social Psychology*, *42*, 535–559. doi:10.1111/j.1559-1816.2011.00782.x
- Nysse-Carris, K. L., Bottoms, B. L., & Salerno, J. M. (2011). Experts' and novices' abilities to detect children's high-stakes lies of omission. *Psychology, Public Policy, and Law*, *17*, 76–98. doi:10.1037/a0022136
- Orbach, Y., Lamb, M. E., La Rooy, D. J., & Pipe, M.-E. (2012). A case study of witness consistency and memory recovery across multiple investigative interviews. *Applied Cognitive Psychology*, *26*, 118–129. doi:10.1002/acp.1803
- Orcutt, H. K., Goodman, G. S., Tobey, A. E., Batterman-Faunce, J. M., & Thomas, S. F. (2001). Detecting deception in children's testimony: Factfinders' abilities to reach the truth in open court and closed-circuit trials. *Law and Human Behavior*, *25*, 339–372. doi:10.1023/A:1010603618330
- Otgaar, H., & Candel, I. (2011). Children's false memories: Different false memory paradigms reveal different results. *Psychology, Crime & Law*, *17*, 513–528. doi:10.1080/10683160903373392
- Otgaar, H., & Smeets, T. (2010). Adaptive memory: Survival processing increases both true and false memory in adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1010–1016. doi:10.1037/a0019402
- Paz-Alonso, P. M., Goodman, G. S., & Ibabe, I. (2013). Adult eyewitness memory and compliance: Effects of post-event misinformation on memory for a negative event. Manuscript submitted for publication.
- Paz-Alonso, P. M., Ogle, C. M., & Goodman, G. S. (2013). Children's memory and testimony in "scientific case studies" of child sexual abuse: A review. In M. Ternes, D. Griesel,

- & B. Cooper (Eds.), *Applied issues in investigative interviewing, eyewitness memory, and credibility assessment*. New York, NY: Springer.
- People v. Dutro, A132883 Cal. App. (2012).
- Peterson, C. (1999). Children's memory for medical emergencies: 2 years later. *Developmental Psychology*, *35*, 1493–1506. doi:10.1037//0012-1649.35.6.1493
- Peterson, C. (2002). Children's long-term memory for autobiographical events. *Developmental Review*, *22*, 370–402. doi:10.1016/S0273-2297(02)00007-2
- Peterson, C. (2011). Children's memory reports over time: Getting both better and worse. *Journal of Experimental Child Psychology*, *109*, 275–193. doi:10.1016/j.jecp.2011.01.009
- Peterson, C. (2012). Children's autobiographical memories across the years: Forensic implications of childhood amnesia and eyewitness memory for stressful events. *Developmental Review*, *32*, 287–306. doi:10.1016/j.dr.2012.06.002
- Peterson, C., Sales, J. M., Rees, M., & Fivush, R. (2007). Parent-child talk and children's memory for stressful events. *Applied Cognitive Psychology*, *1075*, 1057–1075. doi:10.1002/acp
- Pezdek, K., & Hodge, D. (1999). Planting false childhood memories in children: The role of event plausibility. *Child Development*, *70*, 887–895.
- Pezdek, K., & Roe, C. (1995). The effect of memory trace strength on suggestibility. *Journal of Experimental Child Psychology*, *60*, 116–128.
- Phelps, E. A. (2004). Human emotion and memory: Interactions of the amygdala and hippocampal complex. *Current Opinion in Neurobiology*, *14*, 198–202. doi:10.1016/j.conb.2004.03.015
- Pipe, M.-E., Lamb, M. E., Orbach, Y., & Esplin, P. W. (2004). Recent research on children's testimony about experienced and witnessed events. *Developmental Review*, *24*, 440–468. doi:10.1016/j.dr.2004.08.006
- Pipe, M.-E., & Wilson, J. C. (1994). Cues and secrets: Influences on children's event reports. *Developmental Psychology*, *30*, 515–525. doi:10.1037//0012-1649.30.4.515
- Poole, D. A., & Lindsay, D. S. (1995). Interviewing preschoolers: Effects of nonsuggestive techniques, parental coaching, and leading questions on reports of nonexperienced events. *Journal of Experimental Child Psychology*, *60*, 129–154.
- Pozzulo, J. D., & Dempsey, J. (2006). Biased lineup instructions: Examining the effect of pressure on children's and adults' eyewitness identification accuracy. *Journal of Applied Social Psychology*, *36*, 1381–1394. doi:10.1111/j.0021-9029.2006.00064.x
- Pozzulo, J. D., Dempsey, J., & Crescini, C. (2009). Preschoolers' person description and identification accuracy: A comparison of the simultaneous and elimination lineup procedures. *Journal of Applied Developmental Psychology*, *30*, 667–676. doi:10.1016/j.appdev.2009.01.004
- Pozzulo, J. D., & Lindsay, R. C. L. (1998). Identification accuracy of children versus adults: A meta-analysis. *Law and Human Behavior*, *22*(5), 549–570. doi:10.2307/1394573
- Pozzulo, J. D., & Lindsay, R. C. L. (1999). Elimination lineups: An improved identification procedure for child eyewitnesses. *Journal of Applied Psychology*, *84*, 167–176. doi:10.1037//0021-9010.84.2.167
- Qin, J., Ogle, C. M., & Goodman, G. S. (2008). Adults' memories of childhood: True and false reports. *Journal of Experimental Psychology: Applied*, *14*, 373–391. doi:10.1037/a0014309
- Quas, J. A., Bottoms, B. L., Haecerich, M., & Nysse-Carris, K. L. (2002). Effects of victim, defendant, and juror gender on decisions in child sexual assault cases. *Journal of Applied Social Psychology*, *32*, 1993–2021.

- Quas, J. A., Carrick, N., Alkon, A., Goldstein, L., & Boyce, W. T. (2006). Children's memory for a mild stressor: The role of sympathetic activation and parasympathetic withdrawal. *Developmental Psychobiology*, *48*, 686–702. doi:10.1002/dev
- Quas, J. A., & Goodman, G. S. (2012). Consequences of criminal court involvement for child victims. *Psychology, Public Policy, and Law*, *18*, 392–414. doi:10.1037/a0026146
- Quas, J. A., Goodman, G. S., Ghetti, S., Alexander, K. W., Edelstein, R. S., Redlich, A. D., Cordon, I. M., & Jones, D. P. (2005). Childhood sexual assault victims: Long-term outcomes after testifying in criminal court. *Monographs of the Society for Research in Child Development*, *70*, vii, 1–117. doi:10.1016/j.soncn.2011.11.001
- Quas, J. A., & Lench, H. C. (2007). Arousal at encoding, arousal at retrieval, interviewer support, and children's memory for a mild stressor. *Applied Cognitive Psychology*, *30*(5), 289–305. doi:10.1002/acp
- Quas, J. A., Malloy, L. C., Melinder, A., Goodman, G. S., D'Mello, M., & Schaaf, J. M. (2007). Developmental differences in the effects of repeated interviews and interviewer bias on young children's event memory and false reports. *Developmental Psychology*, *43*, 823–837. doi:10.1037/0012-1649.43.4.823
- Quas, J. A., Thompson, W. C., & Clarke-Stewart, K. A. (2005). Do jurors "know" what isn't so about child witnesses? *Law and Human Behavior*, *29*, 425–456. doi:10.1007/s10979-005-5523-8
- Quas, J. A., Yim, I. S., Edelstein, R. S., Cahill, L., & Rush, E. B. (2011). The role of cortisol reactivity in children's and adults' memory of a prior stressful experience. *Developmental Psychobiology*, *53*, 166–174. doi:10.1002/dev.20505
- Read, J. D., Connolly, D. A., & Welsh, A. (2006). An archival analysis of actual cases of historic child sexual abuse: A comparison of jury and bench trials. *Law and Human Behavior*, *30*, 2592–85. doi:10.1007/s10979-006-9010-7
- Regan, P. C., & Baker, S. J. (1998). The impact of child witness demeanor on perceived credibility and trial outcome in sexual abuse cases. *Journal of Family Violence*, *13*, 187–196.
- Reyna, V. F., & Brainerd, C. J. (2011). Dual processes in decision making and developmental neuroscience: A fuzzy-trace model. *Developmental Review*, *31*, 180–206. doi:10.1016/j.dr.2011.07.004
- Reyna, V. F., & Kiernan, B. (1994). Development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology*, *30*, 178–191. doi:10.1037//0012-1649.30.2.178
- Rodrigues dos Santos, B., & Batista Gonçalves, I. (2009). *Testimony without fear? Non-revictimizing cultures and practices*. São Paulo, Brazil: Childhood Brasil.
- Roebbers, C. M., & Schneider, W. (2001). Individual differences in children's eyewitness recall: The influence of intelligence and shyness. *Applied Developmental Science*, *5*, 9–20.
- Roediger, H. L., Meade, M. L., & Bergman, E. (2011). Social contagion of memory. *Psychonomic Bulletin & Review*, *2*(8), 365–371.
- Ross, D. F., Jurden, F. H., Lindsay, R. C. L., & Keeney, J. M. (2003). Replications and limitations of a two-factor model of child witness credibility. *Journal of Applied Social Psychology*, *33*, 418–431. doi:10.1111/j.1559-1816.2003.tb01903.x
- Ross, D. F., Lindsay, R. C. L., & Marsil, D. F. (1999). The impact of hearsay testimony on conviction rates in trials of child sexual abuse: Toward balancing the rights of defendants and child witnesses. *Psychology, Public Policy, and Law*, *5*, 439–455. doi:10.1037/1076-8971.5.2.439

- Ross, D. F., Miller, B. S., & Moran, P. B. (1987). The child in the eyes of the jury: Assessing mock jurors' perceptions of the child witness. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children's eyewitness memory* (pp. 142–154). New York, NY: Springer-Verlag.
- Rubin, D. C., Berntsen, D., & Bohni, M. K. (2008). A memory-based model of posttraumatic stress disorder: Evaluating basic assumptions underlying the PTSD diagnosis. *Psychological Review*, *115*, 985–1011. doi:10.1037/a0013397
- Rudy, L., & Goodman, G. S. (1991). Effects of participation on children's reports: Implications for children's testimony. *Developmental Psychology*, *27*, 527–538. doi:10.1037//0012-1649.27.4.527
- Rush, E. B., Quas, J. A., & Yim, I. S. (2011). Memory narrowing in children and adults. *Applied Cognitive Psychology*, *25*, 841–849. doi:10.1002/acp.1757
- Sayfan, L., Mitchell, E. B., Goodman, G. S., Eisen, M. L., & Qin, J. (2008). Children's expressed emotions when disclosing maltreatment. *Child Abuse & Neglect*, *32*, 1026–1036. doi:10.1016/j.chiabu.2008.03.004
- Saywitz, K. J., & Camparo, L. B. (2009). Contemporary child forensic interviewing. In B. L. Bottoms, C. Najdowski, & G. S. Goodman (Eds.), *Children as victims, witnesses, and offenders: Psychological Science and the Law* (pp. 102–127). New York, NY: Guilford Press.
- Saywitz, K. J., Goodman, G. S., Nicholas, E., & Moan, S. F. (1991). Children's memories of a physical examination involving genital touch: Implications for reports of child sexual abuse. *Journal of Consulting and Clinical Psychology*, *59*, 682–691. doi:10.1037//0022-006X.59.5.682
- Schaaf, J. M., Alexander, K. W., & Goodman, G. S. (2008). Children's false memory and true disclosure in the face of repeated questions. *Journal of Experimental Child Psychology*, *100*, 157–185. doi:10.1016/j.jecp.2007.09.002
- Scherf, K. S., & Scott, L. S. (2012). Connecting developmental trajectories: Biases in face processing from infancy to adulthood. *Developmental Psychobiology*, *54*, 643–663. doi:10.1002/dev.21013
- Schwartz-Kenney, B. M., & Goodman, G. S. (1999). Children's memory of a naturalistic event following misinformation. *Applied Developmental Science*, *3*, 34–46. doi:10.1207/s1532480xads0301_5
- Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law*, *7*, 36–97. doi:10.1037//1076-8971.7.1.36
- Talwar, V., Gordon, H. M., & Lee, K. (2007). Lying in the elementary school years: Verbal deception and its relation to second-order belief understanding. *Developmental Psychology*, *43*, 804–810. doi:10.1037/0012-1649.43.3.804
- Talwar, V., & Lee, K. (2008a). Little liars: Origins of verbal deception in children. In B. Ellis & D. Bjorklund (Eds.), *Origins of the social mind* (pp. 157–178). New York, NY: Springer. doi:10.1007/978-4-431-75179-3_8
- Talwar, V., & Lee, K. (2008b). Social and cognitive correlates of children's lying behavior. *Child Development*, *79*, 866–881.
- Talwar, V., Lee, K., Bala, N., & Lindsay, R. C. L. (2002). Children's conceptual knowledge of lying and its relation to their actual behaviors: Implications for court competence examinations. *Law and Human Behavior*, *26*, 395–415. doi:10.1023/A:1016379104959
- Talwar, V., Lee, K., Bala, N., & Lindsay, R. C. L. (2004). Children's lie-telling to conceal a parent's transgression: Legal implications. *Law and Human Behavior*, *28*, 411–435. doi:10.1023/B:LAHU.0000039333.51399.f6

- Teicher, M. H., Anderson, C. M., & Polcari, A. (2012). Childhood maltreatment is associated with reduced volume in the hippocampal subfields CA3, dentate gyrus, and subiculum. *Proceedings of the National Academy of Sciences, USA*, *109*, E563–72. doi:10.1073/pnas.1115396109
- Terr, L. C. (1983). Chowchilla revisited: The effects of psychic trauma four years after a school-bus kidnapping. *The American Journal of Psychiatry*, *140*(December), 1543–1550.
- Terr, L. C. (1988). Case study: What happens to early memories of trauma: A study of twenty children under age five at the time of documented traumatic events. *Journal of the American Academy of Child and Adolescent Psychiatry*, *27*, 96–104. doi:10.1097/00004583-198801000-00015
- Troxel, N. R., Ogle, C. M., Cordon, I. M., Lawler, M. J., & Goodman, G. S. (2009). Child witnesses in criminal court. In B. L. Bottoms, C. J. Najdowski, & G. S. Goodman (Eds.), *Children as victims, witnesses, and offenders: Psychological science and the law* (pp. 150–166). New York, NY: Guilford Press.
- Convention on the Rights of the Child, November, 20, 1989, United Nations Treaty Series, Vol. 1577, p. 3.
- Usher, J. A., & Neisser, U. (1993). Childhood amnesia and the beginnings of memory for four early life events. *Journal of Experimental Psychology: General*, *122*, 155–165. doi:10.1037/0096-3445.122.2.155
- Valentino, K., Toth, S. L., & Cicchetti, D. (2009). Autobiographical memory functioning among abused, neglected, and nonmaltreated children: The overgeneral memory effect. *Journal of Child Psychology and Psychiatry*, *50*, 1029–1038. doi:10.1111/j.1469-7610.2009.02072.x
- Walsh, W. A., Lippert, T., Cross, T. P., Maurice, D. M., & Davison, K. S. (2008). How long to prosecute child sexual abuse for a community using a Children's Advocacy Center and two comparison communities? *Child Maltreatment*, *13*, 3–13. doi:10.1177/1077559507307839
- Warren, A. R., Nuñez, N., Keeney, J. M., Buck, J. A., & Smith, B. (2002). The believability of children and their interviewers' hearsay testimony: When less is more. *Journal of Applied Psychology*, *87*, 846–857. doi:10.1037/0021-9010.87.5.846
- Warren, K. L., Dodd, E., Raynor, G., & Peterson, C. (2012). Detecting children's lies: Comparing true accounts about highly stressful injuries with unprepared, prepared, and coached lies. *Behavioral Sciences and the Law*, *30*, 329–341. doi:10.1002/bsl
- Wells, G. L., & Loftus, E. F. (2003). Eyewitness memory for people and events. In A. M. Goldstein (Ed.), *Forensic psychology* (pp. 149–160). Vol. 11 in I. B. Weiner (Ed.-in-Chief), *Handbook of psychology*. Hoboken, NJ: Wiley.
- Wells, G. L., & Quinlivan, D. S. (2009). Suggestive eyewitness identification procedures and the Supreme Court's reliability test in light of eyewitness science: 30 years later. *Law and Human Behavior*, *33*, 1–24. doi:10.1007/s10979-008-9130-3
- Wessel, E., & Melinder, A. (2012). Credible girls ought to cry: Emotions displayed during the disclosure of abuse and their effects on judgments of credibility. Manuscript submitted for publication.
- Williams, A. J. (2008). The veiled truth: Can the credibility of testimony given by a niqab-wearing witness be judged without the assistance of facial expressions? *University of Detroit Mercy Law Review*, *85*, 273–292.

- Williams, J. M. G. (1996). Depression and the specificity of autobiographical memory. In D. C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 244–267). New York, NY: Cambridge University Press.
- Windmann, S., & Krüger, T. (1998). Subconscious detection of threat as reflected by an enhanced response bias. *Consciousness and cognition*, 7(4), 603–633. doi:10.1006/ccog.1998.0337
- Wood, B., Orsak, C., Murphey, M., & Cross, H. J. (1996). Semistructured child sexual abuse interviews: Interview and child characteristics related to credibility of disclosure. *Child Abuse & Neglect*, 20, 81–92.
- Wright, D. B., Hanoteau, F., Parkinson, C., & Tatham, A. (2010). Perceptions about memory reliability and honesty for children of 3 to 18 years old. *Legal and Criminological Psychology*, 15, 195–207. doi:10.1348/135532508X400347
- Young, K., Powell, M. B., & Dudgeon, P. (2003). Individual differences in children's suggestibility: A comparison between intellectually disabled and mainstream samples. *Personality and Individual Differences*, 35, 31–49. doi:10.1016/S0191-8869(02)00138-1

Employing Polygraph Assessment

WILLIAM G. IACONO AND CHRISTOPHER J. PATRICK

ALTHOUGH the exact number is not known, it is a safe bet that tens of thousands of polygraph tests are administered in the United States every year. Most of these tests are administered by federal agencies as part of the government's national security screening program, and some are given by law enforcement to screen the integrity of potential new recruits. A substantial fraction are forensic polygraphs administered by law enforcement as an investigative tool to assist the resolution of criminal cases. Some come from criminal defendants who hire examiners in private practice with the hope of obtaining exculpatory outcomes. Others arise from civil cases involving parental custody/fitness, sex offender commitment, and employee rights. In any case, it is unlikely that a forensic psychologist has administered a polygraph. Instead, polygraphs are administered by polygraphers who work in a profession that is largely disconnected from psychology and informed little by psychological science. Our aim in this chapter is to bridge this gap between applied polygraphy and forensic psychology by providing the information needed to critically evaluate polygraph practice. In addition to examining the current state of polygraph testing, we also review future possible applications of deception detection techniques.

CURRENT APPLICATIONS

Conventional polygraph tests typically are used when the question at hand cannot be easily resolved by the available evidence. When the investigation reaches an evidentiary dead end, police may rely on a polygraph test of a known suspect as the means of last resort to resolve the case. Sometimes those who fail these tests, pressured to own up to their misdeeds, confess, thereby providing the police with incriminating evidence they otherwise would not have. In the absence of a confession, a failed test may lead the police to cease the investigation, believing the suspect at hand is guilty even if the evidence is insufficient for

successful prosecution. By contrast, a passed test provides incentive to continue the investigation and look for new suspects.

Polygraph tests are relied on by psychologists in a number of ways:

- In sex offender treatment programs to ensure that offenders are fully disclosing their offenses and fantasies
- By insurance agencies to verify the claims of those insured
- In family court to help resolve charges of misbehavior parents level at each other in their effort to obtain custody of their children
- By the police to verify victims' charges
- By controversial people in the public eye who wish to sway public opinion in their favor by advertising the fact that they passed a "lie detector"
- By the government to protect national security by requiring those with access to classified information to pass tests confirming that they are not spies
- Even by those running fishing contests to verify that winners actually followed contest rules rather than purchasing their lunger from the local supermarket

The Employee Polygraph Protection Act (EPPA; 1988) eliminated much of the most widespread application of polygraph testing, the periodic screening of employees to verify their good behavior and the pre-employment screening of potential hires to see if they possess the qualities desired by the employer. Ironically, the government exempted itself from coverage by this law and has been expanding polygraph testing programs in light of concerns about terrorism and national security. For instance, since the passage of the EPPA, Public Law 106-65, passed as part of the National Defense Authorization Act (2000), requires scientists at nuclear weapons laboratories to submit to polygraph tests to maintain their security clearance. Besides many state and local law enforcement agencies and polygraphers in private practice, over two dozen federal agencies routinely use polygraph tests, including those that are part of the Departments of Defense, Energy, Homeland Security, Justice, and Treasury.

THE POLYGRAPH AND THE POLYGRAPH EXAMINER

Traditional polygraphs are briefcase-size instruments that use moving chart paper to record the autonomic responses elicited by the subject's answers to test questions. Although these devices are still in use, portable computers that digitally record autonomic activity, displaying and storing it in a manner that mimics the appearance of paper chart recordings, are now in common use. Expandable pneumatic belts positioned around the upper thorax and abdomen provide two separate recordings of the chest movements associated with inspiration and expiration. Changes in palmar sweating (skin conductance, aka the galvanic skin response [GSR]) are detected by electrodes attached to the fingertips. For the "cardio" channel, a partly inflated blood pressure cuff attached to the arm reflects relative changes in blood

pressure and provides an index of pulse. Occasionally a fifth channel monitoring blood flow to the fingertip is included. Although this instrumentation is relatively simple, it produces valid records of physiological reactivity that are comparable to those obtained by sophisticated laboratory equipment (Patrick & Iacono, 1991a).

Training in polygraphy is provided by free-standing polygraph schools, most of which are accredited by the American Polygraph Association. The most prestigious of these is at the National Center for Credibility Assessment (NCCA; formerly the Department of Defense Polygraph Institute) located at Fort Jackson, South Carolina. This school offers a one-semester, intensive, hands-on course in polygraphy that covers ethics, law, the physiology and psychology of deception detection, and the various techniques and interview practices employed by examiners. Graduates of the program typically are apprenticed to practicing examiners before becoming fully certified to administer tests on their own. NCCA offers training for many state and city police departments and most federal government agencies, including the military police, the Federal Bureau of Investigation, the Internal Revenue Service, and all of the government security agencies. NCCA also has an in-house research program staffed by doctoral-level psychologists, some of whom share in the teaching of students with polygraph examiners and law enforcement agents.

NCCA, which requires a college degree and two years of law enforcement experience for program admission, represents the best training the profession of polygraphy has to offer. Most accredited schools do not offer as rigorous a program; not all practicing polygraph examiners are graduates of approved schools; and, because polygraphy is not regulated in most states, polygraphers are not necessarily licensed to practice their trade.

POLYGRAPH TECHNIQUES

The polygraph instrument is not capable of detecting lies, and no pattern of physiological response is unique to lying. Consequently, all polygraph techniques involve asking different types of questions, with differential responding to those pertinent to the issue at hand determining outcome. The techniques, all of which have multiple variants, fall into two categories involving either specific incident or personnel screening applications.

SPECIFIC INCIDENT INVESTIGATIONS

There are three types of specific incident polygraph tests. These procedures are applied when polygraph examiners are aware that an event has occurred but are uncertain what role the examinee played in the incident.

Control (Comparison) Question Technique. The so-called control or comparison question technique (CQT) remains the procedure of choice for specific incident investigations like those concerned with known criminal acts. The CQT typically

consists of about 10 questions. The two types of question that are important to the determination of guilt or innocence are referred to as relevant and control questions. The relevant questions deal directly with the incident under investigation (e.g., Did you shoot Bill Birditsman on the night of March 18?). Control items are paired with relevant questions and cover past behaviors that one might associate with “the kind of person” who is capable of killing (e.g., Before the age of 24, did you ever deliberately hurt someone you were close to?). It is assumed that guilty suspects will be more concerned with the relevant questions than with the control questions. The reverse pattern is expected with innocent people.

The typical CQT has three parts: (1) a pretest interview (lasting between 30 minutes and 2 hours) during which the question list is formulated, (2) the presentation of the question list (usually repeated three times with the question order varied for each of the three “charts”) while physiological responses are recorded, and (3) a posttest interrogation.

The pretest interrogation is designed to determine if the examinee is suitable for testing—for example, if he or she slept the night before and is in reasonably good health. It also provides an opportunity for the examinee to provide an account of the facts in dispute, information that is used in combination with the background material provided the examiner about the case to develop the test questions.

The pretest phase of the CQT is critical to the successful administration of the test. It is during this interview that the polygrapher attempts to create circumstances that lead the innocent person to be more disturbed by the possibly trivial issues raised by the control items than by the relevant questions that have to do with the matter under investigation. A common criticism of the CQT is that it is biased against truthful persons, because the relevant questions may be just as arousing to innocent suspects, who may view their freedom or livelihood as dependent on their physiological response to these items, as they are to the guilty (Lykken, 1974). To reduce the likelihood of this occurrence, polygraphers use the pretest interview to focus the subject’s “psychological set” on the control questions if the examinee is innocent or on the relevant questions if she or he is guilty. Two tactics are used to accomplish this objective.

The first is to convince the subject that lies will be detected. One way to achieve this goal is to demonstrate that the polygraph can detect a known lie. In a typical scenario, the examiner connects the subject to the polygraph and says, “I’m going to ask you to pick a number from 1 to 10, write it down, and then show it to me. Both of us will know which number you’ve picked. After that, I will say a number and ask you if it is yours. I want you to answer ‘no’ to each number I say, including the one you picked.” The examiner then records the subject’s responses to each number and tells him or her afterward that the largest reaction occurred when the person lied; if this was indeed the case, the examiner may point it out to the subject on the chart. If it was not the case, the examiner may imply that it was anyway (“I can see from the results that I will be able to tell when you are lying or telling the truth”) or alter the subject’s response to the target number to create

the impression that it elicited a clearly detectable reaction. Some examiners achieve the desired result by having the subject pick a card from a stacked deck and then rely on the physiological record to “determine” which one he or she picked. Most polygraphers routinely use some variant of this type of demonstration procedure, often called a stim or acquaintance test.

A second tactic for establishing the correct psychological set is to continually emphasize the importance of always being truthful. No distinction is made between the relevant and the control questions regarding the burden of truthfulness. Consequently, innocent individuals are led to believe that lying to control questions will lead to a failed test outcome. How it is that they should reach this conclusion is explained for a case of theft by one of polygraphy’s leading proponents, David Raskin (1989), as follows:

Since this is a matter of a theft, I need to ask you some general questions about yourself in order to assess your basic honesty and trustworthiness. I need to make sure that you have never done anything of a similar nature in the past and that you are not the type of person who would do something like stealing that ring and then would lie about it. . . . So if I ask you, “Before the age of 23, did you ever lie to get out of trouble. . . ?” you could answer that no, couldn’t you? Most subjects initially answer no to the control questions. If the subject answers yes, the examiner asks for an explanation. . . [and] leads the subject to believe that admissions will cause the examiner to form the opinion that the subject is dishonest and therefore guilty. This discourages admissions and maximizes the likelihood that the negative answer is untruthful. However, the manner of introducing and explaining the control questions also causes the subject to believe that deceptive answers to them will result in strong physiological reactions during the test and will lead the examiner to conclude that the subject was deceptive with respect to the relevant issues concerning the theft. In fact, the converse is true. Stronger reactions to the control questions will be interpreted as indicating that the subject’s denials to the relevant questions are truthful. (pp. 254–255)

Charts are scored using one or a combination of three approaches. With global scoring, all the information available to the examiner is used to make the determination of truthfulness. Hence, in addition to inspection of the physiological data, the plausibility of the subject’s account of the facts during the pretest interview, his or her demeanor during the examination, and information from the investigative file may all figure into the evaluation.

With now widely employed numerical scoring, the examiner derives a score from the physiological recordings. The magnitude of the response to pairs of control and relevant questions is estimated for each separate physiological channel. In the most commonly employed of several popular methods, a score from +1 to +3 is assigned if the response to the control item is larger, with the magnitude of the score determined by how large a difference is observed. Likewise, a score from –1 to –3 is assigned if the relevant item of the question pair elicited the stronger response. A total score is obtained by summing these values over all

channels and charts, with a negative score less than -5 prompting a deceptive verdict, a positive score exceeding $+5$ a truthful verdict, and scores between -5 and $+5$ considered inconclusive and therefore warranting further testing. In our experience with government examiners, about 10% of CQTs end with inconclusive outcomes.

Both global and numerical chart evaluation have high interscorer reliability. Studies in which examiners blind to case facts evaluate the original examiners' charts typically report reliabilities around .90 (e.g., Honts, 1996; Horvath, 1977; Patrick & Iacono, 1991a, 1991b). The retest reliability of polygraph testing has not been evaluated. The absence of such data is unfortunate, because often questions about the possible increment in validity gained by retesting a defendant arise in legal hearings regarding the possible admissibility of polygraph results. In addition, the CQT, a collection of different techniques, is not a standardized test, so in the absence of retest data, it is not known to what degree examiners, all of whom have their own way of administering the CQT, are likely to obtain the same result when testing the same individual.

The third approach to chart scoring derives from computerized recording systems. Typically the computer provides a verdict in the form of a probability statement as to the likelihood the person was truthful when responding to the questions. Because these systems are marketed commercially, the algorithms and data used to justify the probability statements are proprietary. Although computer scoring is reliable, little is known about the validity of the outputted probability statements, and few polygraphers rely exclusively on computer scoring of charts, especially in forensic evaluations.

Once the charts are scored, the posttest phase of the CQT is launched. Those individuals who are believed to have been untruthful are interrogated during this phase. The point of the interrogation is to leverage the polygraph test outcome to obtain incriminating admissions or an outright confession. During this phase, skillful interrogators may resolve a case that otherwise would never have been resolved. It is this hoped-for outcome, which speaks to the utility and not the validity of the CQT, that keeps the CQT in widespread use despite its general inadmissibility as evidence in legal proceedings.

Directed Lie Technique. The directed lie technique (DLT) is considered a subtype of the CQT. The chief difference lies in the nature of the control questions. For a DLT, the "probable lie" control questions of the CQT are replaced with "directed lie" questions. Directed lies are statements that the subject admits involve a lie before the test begins. In fact, the polygrapher specifically instructs the subject to answer the question deceptively and to think of a particular time when he or she has done whatever the directed lie question covers. Examples of directed lies are "Have you ever done something that hurt or upset someone?" or "Have you ever made even one mistake?" As with the CQT, guilty subjects are expected to respond

more strongly to the relevant questions, and innocent subjects should react more strongly to the directed lies.

Guilty Knowledge or Concealed Information Test. An alternative to the CQT for specific incident investigations is the guilty knowledge test (GKT; Lykken, 1959, 1960), sometimes referred to as a concealed information or knowledge test. Rather than asking directly whether the examinee was responsible for the crime under investigation, the GKT probes for knowledge indicative of guilt—details regarding a crime or incident that only the person who did it would know about. The GKT consists of a series of questions about the crime posed in multiple-choice format. Each question asks about one specific detail of the crime and is followed by a series of alternative answers, including the correct answer as well as other plausible but incorrect options. The following is an example of a GKT question concerning one detail of a homicide: “If you were the one who beat Donna Fisbee to death, then you will know what was used to kill her. Was she beaten with: (a) a brick? (b) a crowbar? (c) a pipe? (d) a baseball bat? (e) a hammer?” When presented with a question of this type, the true culprit would be expected to emit a larger physiological reaction to the correct alternative than an innocent person who knows nothing about the incident and would respond at random.

The simple premise underlying the GKT is that a person will exhibit larger orienting reactions to key information only if he or she recognizes it as distinctive or important. The GKT tests for knowledge of information rather than for deceptiveness, and the irrelevant alternatives are true controls rather than pseudocontrols. In the CQT, deceptiveness is inferred from a pattern of enhanced reactions to relevant questions, but the possibility that “innocent concern” rather than deception is responsible for this outcome can never be ruled out. A pattern of consistent reactions to critical items on a GKT can (within a small, estimable probability) mean only that the examinee possesses guilty knowledge. On a GKT question with five alternative answers, the odds that an innocent person with no knowledge of the crime would react most intensely to the key (relevant) alternative are 1 in 5. On a GKT that included 10 such questions, the odds are vanishingly small (<1 in 10,000,000) that an innocent person would react differentially to the key alternative on each and every test question.

The first study of the GKT (Lykken, 1959) and most others conducted since have utilized peripheral response measures, most commonly skin resistance or skin conductance, as indices of stimulus orienting. More recently, brain potentials recorded from the electroencephalogram have been utilized to detect deception within a GKT format. Measuring how reaction times differ to GKT key and irrelevant multiple choice alternatives has provided another method for identifying those with guilty knowledge (Seymour & Fraynt, 2009). The “attentional blink” paradigm has also been adapted to the GKT (Ganis & Patnaik, 2009). This paradigm makes use of

the fact that when two stimuli are presented in close temporal proximity, attention to the first stimulus in the pair (which may or may not convey guilty knowledge) makes identifying the second stimulus difficult (causing a “blink” in attention).

PERSONNEL SCREENING

Modern screening tests differ from specific incident tests in that it is not known whether any particular transgression has taken place. Consequently, the relevant questions typically cover extended periods of time and many topics, leaving ambiguous what form an adequate “control” question should have. Whereas there are many different types of screening tests, these procedures are historically linked to the relevant/irrelevant technique (RIT), a polygraphic interrogation method that preceded the development of the CQT and was used originally in criminal investigations.

Relevant/Irrelevant Technique. In the original RIT, relevant questions (like those used on the CQT) were each preceded and followed by an irrelevant question (e.g., “Is your name Ralph?” or “Is today Tuesday?”). Consistently greater reactions to the relevant items of the test were interpreted as evidence of deceptiveness. However, because of the obvious confound posed by the differential potency of the two categories of questions, the traditional RIT has been roundly criticized and thus is used only occasionally today. For purposes of employment screening, polygraph examiners now commonly use a variant of the RIT procedure that might more appropriately be called the relevant/relevant technique, because interpretation of test outcome depends on the pattern of responses across all of the relevant questions.

In contrast to specific incident tests, screening examinations contain relevant questions of the form “Have you ever . . . ?” or “During the period in question, did you . . . ?” These questions, which may tap themes related to drug use, trustworthiness, and rule violations, are alternated with innocuous or irrelevant questions (also called norms). Law enforcement and security agencies use these types of tests both with prospective and current employees. Although government secrecy makes it difficult to determine how these two types of subjects fare on these tests, it is clear that prospective employees are much more likely to fail such tests (perhaps a third or more do, depending on the government agency) than those already screened, trained, and employed (where failure rates hovering around 1%–2% are seen).

In a screening test of this type, typically three or more question sequences are presented covering the same topics, but with the form of the questions and their order varied. The irrelevant items are included mainly to provide a rest period or return to baseline rather than a norm for comparison purposes. The RIT is a polygraph-assisted interview in which the development of questions is guided both by the polygrapher’s impressions of the examinee’s truthfulness as well as the comparative reactions to the various relevant items: “The cardinal rule in chart interpretation is, any change from normal requires an explanation” (Ferguson, 1966,

p. 161). If the subject shows persistently strong reactions to one or more content areas in relation to the rest, the examiner concludes that the subject lied or was particularly sensitive about these issues for some hidden reason. In this case, the examiner will probe the examinee for an explanation of what might have provoked these responses and will administer additional question sequences focusing on these specific issues. Examinees who are adept at explaining away their reactions are thus likely to avoid incrimination. Thurber (1981) reported that, among applicants for a police training academy, those who scored highest on a questionnaire measure of impression management were most likely to pass a polygraph screening test.

National security organizations use both periodic and aperiodic screening tests. Periodic screening tests are conducted at regular intervals to determine whether existing employees have been honest in their work and remain loyal to the agency. Aperiodic screenings are conducted less frequently and with minimal advance warning. Besides being more economical, this practice is thought to produce a more powerful deterrent to malfeasance. The knowledge that they may be asked to submit to a polygraph test at any time is believed to dissuade existing employees from engaging in misconduct. In effect, the polygraph establishes a climate of fear in which employees presumably are less inclined to be dishonest because they fear detection (National Research Council [NRC], 2003; Samuels, 1983).

Test for Espionage and Sabotage. In addition to RIT-derived tests, national security agencies have introduced a type of directed lie test as part of their counterintelligence program called the Test for Espionage and Sabotage (TES; or test for espionage, sabotage, and terrorism, TEST), a procedure that has been used extensively with scientists at nuclear weapons laboratories. With the TES, questions such as "Have you given classified information to any unauthorized person?" are paired with directed lies such as "Did you ever violate a traffic law?" Unlike other types of screening tests, the TES can be scored using the same procedures followed for the CQT.

DETERMINING VALIDITY

Hundreds of papers discuss the validity of polygraph testing. Much of this work is unpublished, and much that is published appears in poor-quality or trade journals. Because so many studies touch on the accuracy issue, and because much of the research conducted in this field is not carried out by scientists or published in scientific, peer-review journals, we preface our evaluation of the literature with a summary of the important methodological issues that a serious investigation of polygraph validity must address.

EVALUATION OF POLYGRAPH CHARTS

Although currently semi-objective numerical scoring is the preferred technique for chart evaluation among professional polygraphers, the global approach to

chart interpretation still is used occasionally. For CQTs conducted using either procedure, the field examiner is exposed to extrapolygraphic cues, such as the case facts, the behavior of the suspect during the examination, and sometimes inculpatory admissions from the examinee. For a validity study to provide a meaningful estimate of the accuracy of the psychophysiological test, the original examiner's charts must be reinterpreted by blind evaluators who have no knowledge of the suspect or case facts. Even though those trained in numerical scoring are specifically taught to ignore extrapolygraphic cues, Patrick and Iacono (1991b), in their field study of Royal Canadian Mounted Police (RCMP) polygraph practices, showed that even these elite examiners nevertheless attend to them. In 21% of the 279 examinations investigated, the original examiners contradicted the conclusions dictated by their own numerical scores by offering written verdicts that were not supported by the charts. We also found that original examiner opinions were likely to be more accurate than their numerical scores, indicating that examiners improved their accuracy when they relied on case facts and other extraneous information. Although one may be tempted to use such data to argue that blind chart scoring underestimates the accuracy of polygraph verdicts (e.g., see Honts, Raskin, & Kircher, 2002), the probative value of the CQT derives from the possibility that the psychophysiological measurements provide a scientifically valid method for detecting liars. No court of law would accept as evidence the opinion of a human "truth verifier," a skilled interviewer who can use the available evidence to reach a correct judgment. The fact that our RCMP data showed that original examiners were more accurate when they overrode the charts speaks to the invalidity of the psychophysiological test when used to determine truthfulness.

FIELD VERSUS LABORATORY INVESTIGATIONS

Field studies, like our study with the RCMP just discussed, involve real-life cases and circumstances. The subjects are actual criminal suspects. Laboratory studies require naive volunteers to simulate criminal behavior by enacting a mock crime. The latter approach provides unambiguous criteria for establishing ground truth but cannot be used to establish the real-life error rate, because the motivational and emotional concerns of the suspects are too dissimilar from those involved in real-life examinations. Unlike those faced with an actual criminal investigation, guilty subjects in the laboratory have little incentive to try and no time to research how to "beat" the test, guilty subjects are following instructions to lie rather than lying out of self-interest, and both guilty and innocent subjects have little to fear if they are classified as deceptive. Administering the CQT to laboratory subjects is especially likely to lead to overestimates of accuracy for the innocent. Innocent subjects can reasonably be expected to respond more strongly to the potentially embarrassing control questions concerning their personal integrity and honesty than to the relevant questions dealing with a simulated crime they carried out only to satisfy experimental requirements. However, laboratory research does permit

efficient investigation of the influence of factors that may affect test outcome (e.g., effects of CMs or personality traits).

Laboratory studies of the GKT are also likely to overestimate its accuracy, more so for guilty than innocent individuals. Well-designed laboratory experiments construct a scenario in which guilty participants must attend to details of the “crime” that the examiner expects perpetrators to know and that can be used to construct the GKT. In real life, a criminal may not attend to the aspects of a crime that an investigator views as salient, and many details may be forgotten. For example, there is evidence that psychopathic individuals are less able to process incidental details when focusing on a primary task (Kosson, 1988), and such individuals may thus be less detectable using the GKT (Verschuere, Crombez, De Clercq, & Koster, 2005; Waid, Orne, & Wilson, 1979). If a person does remember the details of a real-life crime, however, his or her recognition should evoke greater physiological reactions, thereby making it easier to detect the guilty.

Although the GKT is used in Israel and exclusively in Japan, there are two reasons why it is seldom used in real-life investigations in North America. First, there is a prevailing belief among field examiners that the CQT is virtually infallible (Patrick & Iacono, 1991b). Thus, there is no need to develop an alternative procedure, especially one that is more complicated to administer than the CQT. Second, to construct a valid GKT, there must be salient details of the crime known only to the perpetrator. Not all crimes meet this criterion, in part because often pertinent facts are generally known (e.g., through media reports). Rape provides a crime well suited for GKT development when the victim can provide pertinent crime details for test construction. Alleged sexual assaults in which the question of force versus consent is the only issue to be resolved would not be amenable to a GKT. However, DNA and fingerprint evidence are not available or necessarily relevant for many crimes, but this has not diminished their evidentiary value for those crimes where such evidence exists.

The problems with laboratory studies dictate that real-life applications must be used to evaluate polygraph tests. Although the CQT has been subjected to field research, there are no field studies of personnel screening tests and only two of the GKT, facts that limit efforts to evaluate these techniques.

PROBLEMS ESTABLISHING GROUND TRUTH

The advantage of field investigations—that they are based on actual crimes—is also a significant drawback, because *prima facie* evidence of innocence or guilt is often lacking. Proponents of polygraphy have argued that confessions provide the best method for operationalizing ground truth. Confessions identify the culpable and clear the innocent. Although occasionally confessions are false, and those who confess may differ in important ways from those who do not, the major problem with this strategy concerns the likelihood that the confession is not independent of the original polygraph examiner’s assessment. For reasons that are unrelated

to test accuracy, confessions are obtained during posttest interrogations and are associated almost exclusively with charts that indicate a deceptive outcome. When this occurs, the verified cases selected for a validity study will be biased in favor of demonstrating high accuracy for the technique.

To clarify this point, consider the following example. Ten women are suspects in a criminal investigation. A polygrapher tests them one by one until a deceptive outcome is obtained, say on the sixth suspect tested. (Under these circumstances, the remaining four women typically would not be tested, unless the crime was believed to involve more than one perpetrator.) According to usual practice, the examiner then attempts to extract a confession from the sixth suspect. If the examinee fails to confess, her guilt or innocence cannot be confirmed. It is possible that the polygrapher committed two errors in testing these six cases: The person with the deceptive chart may have been innocent, and one of those tested before her could have been guilty. In the absence of confession-backed verification, however, the polygraph records from these six cases will never be included as part of a sample in a validity study. If the sixth suspect does confess, however, these six charts, all of which confirm the original examiner's assessment, will be included. The resulting sample of cases would consist entirely of charts the original examiner judged correctly and would never include cases in which an error was made. As Iacono (1991) has shown, if polygraph testing actually had no better than chance accuracy, by basing validity studies on confession-verified charts selected in this manner, a researcher could misleadingly conclude that the technique was virtually infallible. Given how cases are selected in confession studies of validity, it should not be surprising that field validity studies typically report that the original examiner was 100% correct (or nearly so; see Honts et al., 2002) for the cases chosen for study. The case selection method assures this result.

Polygraph proponents have asserted that, because it is the original examiners who testify in court, it is the "accuracy" of the original examiners in these field confession studies that constitutes the "the true figure of merit" to determine how accurate polygraph tests would be in legal proceedings (Honts et al., 2002). Despite the fact that the hit rate of the original examiner in these studies is entirely misleading, given how cases are selected for study inclusion, this argument also ignores the contribution of extrapolygraph information to the original examiner's opinion and the resulting necessity of blind chart scoring to determine how useful the psychophysiological data are for deciding guilt.

WHAT CAN BE CONCLUDED ABOUT POLYGRAPH VALIDITY?

Different conclusions apply to the validity of each of the different types of polygraph procedures. Serious questions have been raised about the accuracy of each of the procedures that polygraph examiners commonly use. Ironically, the one procedure they seldom use, the GKT, has high potential validity.

CONTROL QUESTION TECHNIQUE

The literature relevant to the validity of CQT polygraph testing has been reviewed repeatedly, including in the three prior editions of this text (Iacono & Patrick, 1987, 1999, 2006) as well as in other more recent publications (Iacono, 2007, 2008b, 2010; Iacono & Lykken, 2009; Meijer, Verschuere, Merckelbach, & Crombez, 2008; Vrij, 2008). Despite the importance of determining CQT accuracy and the inability to do so relying on studies contaminated by the confession-verification confound, only one study to date has tackled directly the confession-bias problem that characterizes field research (Patrick & Iacono, 1991b), and we thus focus on the results of that investigation here. In that RCMP field study involving over 400 cases, we attempted to circumvent the confession-bias confound by reviewing police files for evidence of ground truth that was collected outside of the context of the polygraph examination (e.g., a confession by someone who did not take a polygraph test, a statement that no crime was committed because items believed stolen actually were misplaced). Independent evidence of ground truth was uncovered for one criterion-guilty and 24 criterion-innocent suspects. The fact that it was easier to come by independent evidence of the innocence rather than the guilt of someone taking a CQT stemmed from how the police use polygraph tests to assist their investigations. Polygraph tests typically are administered in cases where the evidence is ambiguous and the police have exhaustively explored available leads to no avail. When a case reaches this point, the investigating officer is hoping that polygraph testing will help resolve the case. Ideally, the suspect will fail and confess, thus giving the investigating officer incriminating evidence that can be used to prosecute the suspect. However, if the suspect merely fails, with no new evidentiary leads to follow, the case is effectively closed, with the police concluding that the individual who failed is guilty. If the suspect passes, the case is often left open, and the search for new suspects and evidence continues.

For those independently confirmed as innocent, the blind rescoring of their polygraph charts produced a hit rate of 57%. Because chance accuracy is 50%, this result indicates the CQT has little better than chance accuracy with the innocent. It also indicates that innocent people are indeed often more disturbed by relevant than control questions. Because only one criterion-guilty person was identified in this investigation, it was not possible to estimate the accuracy of the CQT with persons independently confirmed as guilty.

Despite Patrick and Iacono (1991b) laying out how confession studies bias CQT accuracy estimates and the many subsequent reviews that have echoed this concern about field studies (Fiedler, Schmod, & Stahl, 2002; NRC, 2003), a field study was recently published in a peer-reviewed scientific journal that claimed 100% accuracy for the CQT (Mangan, Armitage, & Adams, 2008). This study failed to cite the relevant literature regarding this confession bias problem, and it represents a flawed report that one published commentary characterized as a failure of the peer review system (Verschuere, Meijer, & Merckelbach, 2008; see also Iacono, 2008a).

Although there are no scientifically credible data regarding the accuracy of the CQT with guilty people, there is reason to doubt the validity of truthful polygraph verdicts. Honts, Raskin, and Kircher (1994) showed that with less than a half hour of instruction regarding CQT theory and how to recognize control and relevant questions, guilty subjects in a mock crime study could learn to escape detection by augmenting their autonomic responses to control questions. They were able to do this using both physical and mental CMs, such as biting the tongue or subtracting 7 serially from a number over 200 when the control question was asked. Moreover, experienced examiners were unable to identify those subjects who employed CMs successfully. The information contained in the instructions given to those escaping detection in this study is widely available in various publications (including in Honts et al., 1994, as well as Lykken, 1998) and on the Web (e.g., www.polygraph.com, <https://antipolygraph.org/>), making it relatively easy for those so motivated to learn both how the CQT works and how to augment responses to control questions. Subsequent studies by Honts and colleagues (reviewed in Honts & Amato, 2002; see also Honts & Alloway, 2007) have explored how easy it is for naive volunteers to determine on their own how to use CMs and have concluded that uninformed individuals resort to CM strategies that are often ineffective. However, in these studies, the guilty volunteers typically are given little incentive to use CMs effectively, thus leaving their generalizability to real life settings questionable.

DIRECTED LIE TECHNIQUE

Little is known about the validity of the DLT. Although one field study involving the DLT has been published (Honts & Raskin, 1988), this study was also subject to the confession-bias problem. In addition, only a single directed lie question was used, and this question was embedded in a conventional CQT, making it difficult to determine how the test would have fared had directed lie controls been used exclusively. The DLT appears especially susceptible to CMs. When the examiner introduces the directed lies to the subject, they are explained as questions designed to elicit a response pattern indicative of lying. Hence, their purpose is made transparent to subjects, who may understand that an exaggerated response to these questions will help them pass test items on which they lie and presumably offer a less significant response. In addition, the examiner has no idea what issues are covered by the directed lies and how strong an emotional response they are capable of eliciting. For instance, if the subject is directed to answer no to the question "Have you ever done something that you later regretted?" and the subject had an abortion or killed someone in a drunk driving incident, might not the emotions elicited by the directed lie elicit stronger autonomic responses than the material covered by a question concerned with less significant matters, such as theft or fraud?

GUILTY KNOWLEDGE TEST

Of the three classes of polygraph tests considered in this review, only the GKT is spurned by practicing polygraphers. Because of this, few data available from real-life GKT applications can be used to evaluate validity. There are many laboratory simulations of the GKT, and Lykken (1998) has outlined the criteria that define a well-conceived GKT and also reviewed studies that use GKTs meeting these criteria. For instance, Lykken noted that a good test might have 10 items, each with five alternatives, and the person taking the test would be asked to repeat each alternative rather than merely responding no to each, to ensure the examinee was paying attention. The alternatives for each item should be distinctly different from each other, so the examinee can readily recognize the guilty alternative. Lykken's review of eight studies with well-constructed GKTs found accuracy rates of 88% and 97% for guilty and innocent study subjects, respectively.

A meta-analysis of 22 investigations by MacLaren (2001) that used less selective criteria for study inclusion reported somewhat lower accuracies (76% for guilty and 83% for innocent subjects). In a comprehensive meta-analytic review, Ben-Shakhar and Elaad (2003) examined 80 studies and included moderator analyses that pointed to several factors that enhanced validity. Studies that employed mock crime simulations, motivational incentives to succeed, verbal responses to item alternatives, and five or more questions produced better hit rates than those without these features. The authors concluded that "the GKT may turn out to be one of the most valid applications [of a test based on] psychological principles" (p. 145). Another study by Ben-Shakhar and Elaad (2002) showed that a GKT composed of many questions that focus on numerous aspects of the event at issue has better detection efficiency than a test of identical length that focuses on only one or two aspects of the event. This finding is important because, in field applications, it is often difficult to develop questions, so it is easier to generate a test composed of one or a few items presented repeatedly than a test composed of many different items.

The GKT, as represented in the studies reviewed previously, relies on the measurement of autonomic nervous system measures, most typically the electrodermal response. However, measures of other functions may work as well as or better than autonomic measures. For instance, GKT studies in which brain event-related potentials (ERP) have served as the dependent measure have been similarly impressive in their classification accuracy. Farwell and Donchin (1991) reported perfect classification of "guilty" and "innocent" subjects based on a comparison of their P300 reactions to relevant and irrelevant items of information. A more detailed review of brain-based techniques for assessing deception, including variants of the GKT that have utilized P300 response, is provided below (see the "Alternative Methods" section).

Because the test is virtually never used in North America, no field studies of the GKT have been conducted here. However, the GKT is routinely used in

Japan (Nakayama, 2002), and two studies have been reported by investigators in Israel. Eiaad (1990) and Eiaad, Ginton, and Jungman (1992) examined the GKT records of 178 criminal suspects tested by examiners from the Israel Police Scientific Interrogation Unit, whose criterion status had been established via confessions. In all but one instance, the GKT was administered following a CQT and included from one to six questions repeated from two to four times, a procedure that, as noted (Ben-Shakhar & Eiaad, 2002), diminishes the effectiveness of the GKT. Excluding inconclusive outcomes, innocent examinees were identified with high accuracy (error rate of 2%–3%). Guilty people were less accurately identified, with hit rates varying from 42% to 75% depending on the choice of scoring criteria.

PERSONNEL SCREENING

Because almost everyone recognizes that the RIT is biased against the innocent (e.g., Horowitz, Kircher, Honts, & Raskin, 1997), it has been replaced by the CQT for specific incident investigations. However, despite their lack of empirical foundation, RIT variants and the TES are nevertheless commonly used by the government for employee screening.

Although personnel screening tests that require responses of consistently similar magnitude across many relevant questions to identify truthfulness may appear more credible than the traditional RIT, their premises and applications also have been challenged. Heightened reactions to certain specific questions may occur for reasons other than deceptiveness, such as indignation about being asked the question, exposure to some related issue through the media, or knowledge of someone else who has engaged in the sort of activity covered by that question. Moreover, there is no reason to assume that enhanced reactions to an evocative question will subside once the examinee has offered an explanation for those enhanced reactions to the examiner. In fact, the CQT rests on the opposing (also unproven) assumption that truthful subjects will remain worried about control questions even after these items have been modified to accommodate their admissions. These criticisms give rise to the concern that personnel screening is likely to be associated with a high false positive error rate. In fact, however, as applied by government agents, the false negative error rate seems to be a much more substantial concern, because out of the thousands of personnel screening tests administered every year, only a handful of individuals fail (NRC, 2003).

Research conducted at the NCCA offers some insight into why few individuals fail polygraph screening tests. In their unpublished government report, Barland, Honts, and Barger (1989) described the results of a large analog study that was designed to assess the validity of periodic espionage screening tests administered by experienced government examiners from multiple federal agencies. The 207 study participants were government military and civilian employees. “Guilty” subjects went through complex simulations in which they met with an agent purportedly engaged in espionage who recruited them to collaborate in this activity. Consequent to their

recruitment, these “spies” committed acts of mock espionage in which they copied or stole classified documents—just the type of activities that periodic screening tests were designed to detect.

The results of this study indicated a high rate of correct classification for innocent participants (94%) but a low hit rate for guilty participants (34%). The high false negative rate could be related to several factors, but the one most likely is related to the fact that the examiners in this study, who were unaware of the base rate of guilt (about 50%), were following the established field practice of passing almost everyone who took the test. Because periodic screening in real life is in a sense a fishing expedition in which the base rate of spying is presumably negligible, and because examiners are likely to be discouraged from falsely accusing innocent people, many of whom are high-ranking, well-educated, and trained government officials with many years of government service, testing and decision-making practices in the screening context are likely to be biased toward finding few examinees deceptive (Barland et al., 1989; Honts et al., 1994).

Apparently in part because of findings like these, the TES was developed and subjected to two laboratory studies (Research Division Staff, Department of Defense Polygraph Institute, 1997, 1998) that reported relatively low rates of both false positive (12.5%) and false negative (17%) error. As noted previously, classification rates observed in analog studies cannot be expected to generalize to the field, where one could expect many innocent government employees with top-secret security clearances to be more bothered by loyalty-challenging questions about espionage and sabotage than directed lie questions about traffic violations. Moreover, because even a 12.5% false positive rate among highly trained weapons lab scientists would wreak havoc on the ability of the United States to carry out its nuclear weapons program, field examiners adjust the threshold for failing the TES so virtually no one fails. The NRC analysis of the TES (NRC, 2003), which included additional unpublished government studies not available to the public, reached the conclusion that “these studies do not provide strong evidence for the validity or utility of polygraph screening” (p. 133).

NRC’S CONCLUSION REGARDING LIE DETECTION ACCURACY

A report from the National Academy of Sciences (NRC, 2003) provides the most comprehensive review of the evidence for polygraph test accuracy ever undertaken. This review, which was requested by the Department of Energy, was launched in part because of concerns regarding the desirability of expanding the government’s personnel screening program to include scientists working in the Department of Energy’s weapons laboratories. However, the review covered polygraph testing in its entirety, focusing on specific incident polygraph tests because, as we have noted, there are no scientifically peer reviewed, published studies on the validity of screening tests. The review was carried out by a panel of 14 distinguished scientists, with no connection to polygraphy, who represented a variety of disciplines

and types of scientific expertise. These scientists had the training, education, and stature to provide a competent and unbiased professional evaluation of the polygraph literature.

Their critique, spread throughout a 398-page volume, was overwhelmingly negative. The panel members did not attempt to estimate precisely polygraph accuracy, nor did they distinguish among types of tests (e.g., CQT versus GKT) or how hit rates may vary for guilty and innocent subjects. Instead, they identified a set of 57 specific incident studies that met “minimal criteria” (NRC, 2003, p. 107) for consideration, noting that the selected studies “do not generally reach the high levels of research quality desired in science” (p. 108). Using the data from these studies, they plotted receiver operating curves (ROCs), borrowing a method from signal detection theory. The primary statistic derived from this analysis was an “accuracy index (*A*)” corresponding to the area under the ROC curve. *A* takes on a value between .5 and 1.00 and, although similar to percentage correct, does not translate directly to the types of percentage estimates reported in the studies analyzed or to those typically reported in reviews of this literature, in part because the ROC analysis takes into account inconclusive outcomes as well as the differences across studies in the rules followed to determine how the outcome of a polygraph test was classified. Because none of the analyzed studies showed the polygraph to have accuracy at or below chance and because these studies indicate well below perfect accuracy, the panel concluded that, for naive examinees untrained in CMs, specific incident polygraph tests have hit rates “well above chance, though well below perfection” (p. 214).

ALTERNATIVE APPROACHES TO DETECTING DECEPTION

A growing area of interest concerns alternatives to conventional polygraph techniques, including reliance on brain ERPs, functional magnetic resonance imaging (fMRI), and thermal imaging techniques. Interest in these methods has been spurred in part by the desire to develop new methods for lie detection that, unlike the CQT, are likely to meet legal standards for what constitutes scientific evidence. This section considers each of these alternative methods in turn.

ERP-BASED DETECTION METHODS

The most extensively researched alternative approach to detection of deception has utilized components of the brain ERP, in particular the P300 component, which occurs in response to significant, infrequent (often referred to as oddball) stimuli. In a P300-based GKT procedure, the crime-relevant keys comprise the rare, meaningful stimuli. When interspersed with the crime-irrelevant multiple-choice alternatives, none of these key alternatives appears “odd” to the person without guilty knowledge, so they elicit minimal P300 response. For the guilty person, the crime-relevant keys are far fewer in number than irrelevant alternatives and are recognized as special, and thus they elicit enhanced P300 reactions.

A real-life example of the use of this approach was in the case of *Harrington v. State of Iowa* (1997). Here, a P300-based GKT was admitted as evidence in the appeal of Terry Harrington, a man who consistently maintained his innocence despite being convicted of murder more than 20 years earlier. Using the procedures outlined in Farwell and Donchin (1991), Harrington was found to have passed a brain ERP-GKT related to his knowledge of the crime scene by showing no enhanced-P300 recognition response to stimuli involving crime details that were identified by Farwell. Moreover, Harrington showed a brain recognition response to stimuli involving his alleged alibi that were developed independently by Farwell without the knowledge or participation of Harrington. Harrington's conviction was ultimately overturned. Farwell has used the term *brain fingerprinting* to refer to this ERP-GKT and formed a company to market its application. Because this methodology, unlike conventional lie detector methods, is based on the strong scientific foundation afforded by decades of research on the GKT (Verschuere, Ben-Shakhar, & Meijer, 2011), we devote special attention to research on the ERP-GKT in this section.

Initial Published Studies. The first published research report of P300 in the detection of guilty knowledge was by Rosenfeld, Nasman, Whalen, Cantwell, and Mazzeri (1987).¹ In this study, participants were shown a box containing nine items (e.g., camera, film, coins), identified the item they would most want to keep, and wrote a 100-word essay describing reasons for this choice. Next, participants viewed a series of words on a monitor, each repeated several times, with instructions to attend carefully to all words. For "guilty" participants ($n = 10$), one of the words (the key) corresponded to the chosen item, with the rest consisting of words for novel items of commensurate value (e.g., radio, cassette, medal). For "innocent" participants ($n = 6$), all of the words consisted of labels for novel items; one of these was arbitrarily designated the key. Statistical analysis of ERP amplitude within a 400- to 700-ms window following word onset revealed significantly larger P300 for the key versus the irrelevant words in the guilty group; statistics were not presented for the innocent group. A practical limitation of this study was that no criteria were presented for classifying individuals as guilty versus innocent. Nevertheless, based on a visual inspection of the waveforms for each individual, the authors concluded that all but one of the guilty participants showed distinct P300 differentiation between key and irrelevant words.

A further limitation of this study was that participants were explicitly instructed to attempt deception by thinking no whenever the key word appeared, which may have contributed to enhanced P300 responses. This feature of the procedure also

1. A conference abstract summary of a study utilizing a P300-based ERP approach to detection of deception—subsequently reported as Study 2 of an article by Farwell and Donchin (1991)—appeared a year earlier (Farwell & Donchin, 1986), at which time the Rosenfeld et al. (1987) report was under editorial review (cf. Rosenfeld, 2011).

limits external validity, insofar as real-life guilty suspects could not reasonably be expected to comply with such an instruction. Rosenfeld et al. (1988) addressed this issue with a revised protocol in which attention to test words was ensured by instructing participants to look for and count occurrences of one of the novel irrelevant words whenever it appeared on the screen. Results paralleled those of the initial study. Participants in the guilty group ($n = 7$) showed significantly larger P300 responses to the key nontarget word than to irrelevant nontarget words, and for all individuals, responses to the key word exceeded those to irrelevant nontargets (i.e., in no case did amplitude of response to the seven irrelevant nontargets exceed 75% of the amplitude for the key word). A procedural limitation in terms of realism was that participants, as in the study conducted by Rosenfeld et al. (1987), were required to compose an essay regarding the chosen item prior to testing. Other limitations were (a) statistics were not presented for innocent participants ($n = 5$), (b) no quantitative criteria were provided for categorizing participants as guilty versus innocent based on their test responses, and (c) data from three additional guilty participants were excluded from the report due to excessive eye movements or P300 nonresponding.

Two follow-up studies by Rosenfeld and colleagues evaluated the use of ERP measures in more conventional polygraph testing formats. Rosenfeld, Angell, Johnson, and Qian (1991) examined the accuracy of P300 as an index of deception in a procedure analogous to the standard control question test. Rather than testing for knowledge of specific crime details, the test included "Did you do it?" questions pertaining to a specific offense under investigation, along with control questions pertaining to other accusations. Based on a complex, four-step classification algorithm, hit rates for guilty and innocent participants in this study were 92% and 86.6%, respectively. M. M. Johnson and Rosenfeld (1992) evaluated the utility of P300 for detecting deception in a variant of a pre-employment screening test. P300 response was recorded to phrases describing various antisocial acts, presented sequentially on a computer monitor, interspersed with a target phrase to which participants responded with a button press. Upon completion of the test, ground truth was evaluated by having participants complete a checklist under ostensibly anonymous conditions, on which they indicated whether they had committed any of the antisocial acts listed in the ERP test. Hit rates for guilty and innocent participants, based on a three-step classification algorithm, were 100% and 76%, respectively. Although these results appear fairly impressive, the studies themselves are subject to the same sorts of criticisms described earlier with regard to other laboratory investigations of the control question and employee screening tests.

Another influential early article on the use of P300 to detect guilty knowledge was authored by Farwell and Donchin (1991). The two experiments described in this report were innovative in several respects. First, the crime scenarios were quite realistic. In Experiment 1, participants underwent one of two espionage role-plays involving the exchange of information with a "foreign agent," in which they were exposed to six critical details included as probes on the guilty knowledge test.

In Experiment 2, participants were tested about details of minor offenses they had committed in real life. In both experiments, guilt versus innocence was manipulated within subjects (i.e., in Experiment 1, each individual was tested concerning details of the role-play in which he or she participated [guilty condition] as well as the other scenario [innocent condition]; in Experiment 2, each participant was tested regarding the offense he or she had committed [guilty] along with details of another offense committed by a different study participant [innocent]). Another notable feature of these experiments was that the GKT protocol, which paralleled that described by Farwell and Donchin (1986), required participants to respond to all test stimuli: Irrelevant targets (one sixth of trials) prompted a left button press, and irrelevant nontargets (two thirds of trials) and crime-relevant nontargets (probes; one sixth of trials) prompted a right button press. This ensured that participants attended to all stimuli and classified them in a manner that optimized P300 responses.

A further innovation of this study was that it introduced a statistical criterion for classifying participants as innocent or guilty based on comparative P300 responses to irrelevant nontargets and crime-relevant probes. The technique, known as bootstrapping (Efron, 1979), yields an estimate of the sampling distribution for a parameter under circumstances of limited data, by randomly and iteratively sampling from available scores and computing values of the parameter for each subsample. In the Farwell and Donchin (1991) study, bootstrapping was used to estimate, for each participant, cross-correlations (i.e., reflecting the degree of relationship between corresponding points of one ERP waveform and another across time) between (a) the average P300 response to probes and the average response to irrelevant non-targets, and (b) the average response to probes and the average response to irrelevant targets. If the estimated correlation between probe and target values significantly exceeded that between probe and nontarget values, it was concluded that the participant had recognized the probes as rare and distinctive compared with nontargets and that “guilty knowledge” was present. Conversely, if the correlation between values for probe and nontarget trials exceeded that between probe and target trials, it was concluded that guilty knowledge was not present.

Results were impressive. In Experiment 1, 18 of 20 participants were classified correctly in the guilty condition, with 2 cases inconclusive (i.e., above-mentioned correlations did not differ significantly), and 17 of 20 were correctly classified in the innocent condition, with 3 inconclusives. In Experiment 2, all 4 participants were classified correctly in the guilty condition, and 3 of 4 were correctly classified in the innocent condition, with 1 inconclusive. Thus, in cases for which the bootstrap classification analysis yielded a conclusive outcome, 100% accuracy was achieved.

Nonetheless, there were some notable limitations in this study. Sample sizes were small, particularly in Experiment 2. The accuracy of the test in the guilty conditions was almost certainly enhanced by the fact that, in both experiments, participants explicitly reviewed the crime-relevant details (probe items) prior to taking the test—in contrast to real life, where crime-relevant details are encoded ad

hoc and unlikely to be rehearsed prior to testing. Also, no adverse consequences were contingent on test performance, unlike real-life circumstances. Although the presence of threat could augment reactions to critical items among suspects with guilty knowledge, it is also possible that high negative affect might impair memory retrieval and brain response differentiation. A further point is that a simple reaction time (RT) measure (i.e., latency to press the designated button following the stimulus) also differentiated clearly between criterion conditions in Experiment 1: Participants in the guilty condition showed reliably longer RTs to probes versus irrelevant nontargets, whereas in the innocent condition they did not. Although the authors dismissed RT as a viable index of guilt status on the grounds that it can easily be manipulated, findings from subsequent studies have demonstrated that simple CMs can in fact be used to alter P300 responses in an ERP-based GKT (Rosenfeld, Soskins, Bosh, & Ryan, 2004; see "The Impact of Countermeasures," further on) and that an RT-based GKT actually might be more resistant to CMs (Seymour, Seifert, Shafto, & Mosmann, 2000).

One other early published study that served as a foundation for subsequent work in this area was conducted by Allen, Iacono, and Danielson (1992). Although framed more as a study of memory than deception, this study nonetheless employed a test protocol similar to that of Farwell and Donchin (1991) to assess for the presence of guilty knowledge. Findings were reported for three experiments involving a common protocol. Participants learned two lists of category words, one at the beginning of the experimental session (delayed list) and the other just prior to the P300-based memory test (immediate list), after completing a series of intervening tasks. In the memory test, participants pressed a "yes" button whenever they saw a word from the immediate category list (1/7 of trials), and a different "no" button whenever they saw a word from either the delayed list (1/7 of trials) or from 1 of 5 nonlearned category lists (5/7 of trials). Thus, on the test, participants had to inhibit a tendency to respond to previously learned words in the same way as words they had just learned. Recognition of words from the delayed list was predicted to yield enhanced P300 response in comparison with nonlearned words.

A key feature of this study was that it relied on a novel statistical technique for classifying individual participants as knowledgeable or not with regard to specific word lists, a Bayesian classification strategy. This involved selecting various parameters of the ERP waveform that differentiated learned from unlearned words (e.g., P300 amplitude; area under the curve within 200 ms on either side of the P300 peak) and then using information about the discriminability of these parameters and the relative frequencies of learned and unlearned trials to compute a probability for each participant that an ERP average for a given word list reflected one or the other condition. The discrimination parameters and Bayesian classification algorithm were developed using data from 20 participants in Experiment 1 and then cross-validated on two new samples of 20 participants each in Experiments 2 and 3. Procedures were identical across experiments except that (a) instructions differed slightly in Experiments 2 and 3 (i.e., participants were told to press yes for

words they had learned and no for words they had not, but to deliberately lie about words from the initial learned list by pressing the “no” button to these words, and (b) participants in Experiment 3 were promised \$5 if they could control their brain responses so as to prevent detection of words they had lied to.

Using the Bayesian algorithm and cutpoints developed in Experiment 1, the sensitivities (probability of correctly classifying a learned list as learned; cf. true positive rate) in Experiment 2 and 3 were .925 and .95, respectively, compared with .95 in Experiment 1. The specificities (probability of correctly classifying an unlearned list as unlearned; cf. true negative rate) were .94 and .98, respectively, compared with .96 in Experiment 1. In a reanalysis of data from this study, Allen and Iacono (1997) found that the use of Farwell and Donchin’s (1991) bootstrapping method to classify lists as learned versus unlearned yielded no incorrect classifications; however, it yielded inconclusive results for learned lists in 13% of cases and for unlearned lists in 28% of cases. Allen et al. (1992) also examined the accuracy of classifications based on two indicators of behavioral response to words from each list (RT, response errors); sensitivities were .95 and .95 in Experiments 2 and 3, respectively, and specificities were .95 and .98 (versus .975 and 1.0, respectively, in Experiment 1). Thus, classification accuracies based on behavioral response indices were commensurate with those based on ERP parameters (cf. Farwell & Donchin, 1991). They were also in line with the findings of Seymour et al. (2000), who found that RTs to probe stimuli could be used to separate guilty from innocent individuals in an RT-based GKT, even when subjects were instructed to modify their responses to escape detection.

These results indicated that concealed information could be detected with very high accuracy in individual cases using a probabilistic analysis of ERP response parameters. However, some limitations of the Allen et al. (1992) study are important to consider in relation to detection of deception in real-life cases. In particular, the word-learning task has limited external validity vis-à-vis a real-world crime situation. Simple category words are obviously very different from crime-relevant details. Also, as in other work cited, participants in this study explicitly learned the relevant words as opposed to encountering them incidentally in a dynamic real-world context. Furthermore, the Bayesian classification algorithm developed in Experiment 1 capitalized on information that may not readily be available in real-life cases—namely, the ground-truth status of previously learned lists. ERP parameters were selected in part because they discriminated words on these “concealed” lists from words on the unlearned lists. With real-world suspects, the status of information as concealed or not is normally indeterminate. Although a parallel algorithm could be developed using ERP data from real-life cases in which a solid ground-truth criterion (e.g., a corroborated confession; DNA evidence) became available after testing, the generalizability of this algorithm to cases different from those included in the development sample (e.g., in terms of type of crime, latency since commission, suspect characteristics, etc.) would be open to question. With regard to these points, it should be reiterated that the Allen et al. study was framed

as an investigation of memory rather than of deception. Nevertheless, issues such as these are important to consider in applying the findings of this study to the problem of detecting deception.

Subsequent Studies Building on Initial Published Work. The most active researcher in this area over the past two decades in terms of published studies in peer-reviewed journals has been Peter Rosenfeld of Northwestern University. Many of the studies reported by Rosenfeld and his colleagues through the early 2000s, following the approach of Allen et al. (1992), focused on P300 as an index of dissimulated (“malingered”) amnesia for simple types of learned material, such as words, numbers, and basic autobiographical facts, rather than details of an enacted “crime” (for reviews of this work, see Rosenfeld, 2002; Rosenfeld & Ellwanger, 1999). Other investigations of this type were published during this period by Allen and colleagues (Allen, Iacono, Laravuso, & Dunn, 1995; Allen & Movius, 2000; van Hoof, Brunia, & Allen, 1996; for reviews of this work, see Allen, 2002; Allen & Iacono, 2001).

Building on this basic work investigating P300-based detection of generic learned information, research over the past decade has focused on further evaluating the effectiveness of ERP methods for detecting crime-relevant knowledge in investigative contexts. One series of studies by Rosenfeld and colleagues, on the impact of CMs on detectability using P300-based methods, is discussed in the next subsection. Another line of work, by Lawrence Farwell and colleagues, has focused on a scoring method termed MERMER (memory and encoding related multifaceted electroencephalographic response) that entails quantification of multiple features of the ERP response to test stimuli, including the P300 along with other parameters. In an initial full-length report of this quantification method by Farwell and Smith (2001), six participants were tested, three of them regarding known biographical details from their own lives and the other two regarding unfamiliar biographical details. The test protocol, like that of Farwell and Donchin (1991), was a response task that included irrelevant target stimuli (calling for a left button press) along with irrelevant nontargets and crime-relevant nontargets (each calling for a right button press). Hit rates for both conditions in this study (“guilty”-informed, “innocent”-uninformed) were reported as 3/3 (100%). Subsequent studies of this method have evaluated its accuracy in mock crime (Farwell, Hernandez, & Richardson, 2006) and actual or simulated field contexts (Farwell, 2008; Farwell et al., 2006; Farwell, Richardson, & Richardson, 2011), but these studies have been reported only in conference abstract form. In a recent review of studies using the MERMER scoring approach, Farwell (2012) characterized this technique as yielding 100% accurate classifications in all research studies to date, with no “indeterminate” (inconclusive) outcomes. The lack of indeterminate outcomes was cited as an advantage of the MERMER scoring approach over the more standard P300-focused scoring approach.

These reported findings for the MERMER method have been criticized on several grounds. As noted by Rosenfeld (2005), a serious limitation of this work from the standpoint of scientific evaluation is that the quantification parameters for MERMER are not described in sufficient detail in any published report to permit replication, because they are patented and considered proprietary. Although Farwell and Smith (2001) stated that MERMER scoring entails quantification of the parietal P300 and a subsequent negative-polarity component, maximal at frontal sites, along with “phasic changes in the frequency and structure of the [ERP] signal” (p. 137), the nature of these latter “phasic changes” was not specified in this article or in subsequent reports by Farwell and colleagues. Farwell’s (2012) review article does clarify that the bootstrap cross-correlation approach of Farwell and Donchin (1991) serves as the basis in MERMER for evaluating similarity of ERP components across differing stimulus conditions but fails to specify how (a) evaluations for P300 and late-negative components are combined, or (b) phasic signal changes are incorporated into the waveform morphology comparisons.

Rosenfeld (2005) also raised other concerns regarding the MERMER scoring technique. He questioned Farwell et al.’s characterization of the MERMER approach as yielding 100% accuracy of classifications based on only a single published journal article, when studies published by other investigative groups using P300-based approaches had reported accuracies below this level. In addition, Rosenfeld challenged the scientific status of the additional late-negative component and “phasic change” parameters utilized in the MERMER approach. Whereas an extensive literature shows that P300 is sensitive to the salience/recognizability of presented information, the functional significance of the other MERMER parameters is unclear. Citing Soskins, Rosenfeld, and Niendam (2001), Rosenfeld (2005) pointed out that, although the late-negative component in part reflects recovery to baseline of the preceding P300 response, a parameter that may contribute incrementally to detectability of known versus unknown information, it also may contain some nonspecific, artifact-related variance (i.e., associated with capacitive rebound of the signal at filter settings used for recording of P3).

Regarding the “phasic change” parameters used in MERMER, Rosenfeld (2005) pointed out: “The meaning of these other claimed independent (but undocumented) frequency phenomena, which, according to Farwell himself, are not found in all persons, is another matter. . . . *The supportive data—e.g., power spectra illustrating these claimed frequency effects—have never been shown anywhere*” (p. 27 [emphasis in original]). Some of these concerns raised by Rosenfeld were echoed in a more recent critique of the MERMER technique by Meijer, Ben-Shakhar, Verschuere, and Donchin (2012). Additionally, these authors challenged the assertion made by Farwell (2012) that his brain fingerprinting method had been evaluated in studies involving “over 200 test cases”; they pointed out that the set of studies cited by Farwell overlapped substantially in terms of participant samples and included many unpublished datasets, such that peer-reviewed findings pertaining to the method’s

validity are in fact limited to results from a total of only 30 participants across two peer-reviewed journal articles (Farwell & Donchin, 1991; Farwell & Smith, 2001).

With regard to evaluation of ERP-based methods in applied contexts, a further study that warrants mention is one by Mertens and Allen (2008), which employed a virtual reality crime procedure to evaluate the accuracy of a P300-based detection of deception test. Participants in the study logged on to a computer in an unoccupied office and navigated through a highly realistic virtual environment depicting the interior of a multi-room apartment. Innocent participants were instructed simply to explore the virtual apartment for a designated period of time. Guilty participants entered the virtual apartment for a similar period of time under instructions to “steal” specified items from the apartment through use of a computer mouse. Following exposure to the virtual environment, participants underwent a P300-based detection test akin to that of Farwell and Donchin (1991), including probe (crime-relevant), target (learned irrelevant), and distracter (nonlearned irrelevant) items. Guilty participants completed the detection test either without instruction regarding how to defeat the test (subgroup 1 = no CMs) or under instructions to perform specific types of CMs (subgroup 2 = mental CMs to target stimuli; subgroup 3 = physical CMs to target stimuli; subgroup 4 = alternating physical and mental CMs to distracter stimuli).

An additional feature of the study was that classification accuracy was compared for three different scoring methods: bootstrapped cross-correlation (Farwell & Donchin, 1991), bootstrapped peak-to-peak amplitude difference (Rosenfeld et al., 2004; Soskins et al., 2001), and Bayesian probability analysis (Allen et al., 1992). For innocent participants, the cross-correlation method produced a very high rate of indeterminate outcomes (56%), with the remainder of cases (44%) correctly classified. By contrast, the two other scoring methods yielded conclusive classifications for all innocent participants, with accuracy for the peak-to-peak method (100%) slightly exceeding that for the Bayesian method (96%). In the case of uninstructed (non-CM) guilty participants, the indeterminate rate for the cross-correlation method was again very high (60%), with 27% of cases correctly classified as guilty and 13% incorrectly classified as innocent. For these same guilty participants, the peak-to-peak and Bayesian methods each produced 47% correct (“guilty”) classifications and 53% incorrect (“innocent”) classifications, with no indeterminate outcomes. Results for the guilty CM groups are discussed in the next subsection.

Based on these results, Mertens and Allen (2008) concluded that the accuracy of P300-based detection tests with guilty suspects may be appreciably lower in field contexts involving memory for real-life crime details as compared to lab contexts involving learned lists of probe items. At the same time, these investigators noted that the P300-based detection method—in contrast with the conventional control question procedure used by North American polygraph examiners—is advantageous in terms of yielding very low rates of false positives (i.e., innocent cases mistakenly classified as guilty). In sum, these authors concluded that guilty/deceptive outcomes of ERP-based detection tests are likely to be of

substantial value for investigative decision making in real-life cases (i.e., because such outcomes are strongly diagnostic of the presence of crime-relevant knowledge), whereas innocent/truthful outcomes are likely to be less accurate and thus of questionable utility.

Impact of Countermeasures on P300-Based Information Detection. A major development over the past few years has been the emergence of systematic research on the effects of CMs on the accuracy of ERP-based detection tests and on approaches to detecting and reducing their effectiveness. Whereas an initial study by Sasaki, Hira, and Matsuda (2001) focusing on use of a simple mental CM manipulation (counting backward by sevens) reported no impact on P300-based detection, a subsequent article by Rosenfeld et al. (2004) reported significant effects on P300 responsivity and detectability as a function of physical and affective-imaginal CMs. This article reported findings from two experiments testing for guilty knowledge, the first entailing crime scenarios akin to those used in Experiment 1 of Farwell and Donchin (1991) and the other a simpler procedure entailing detection of personal knowledge (i.e., participant's birth date).

Because procedures aside from the "crime" manipulation (including scoring methods) were similar, only the first experiment is detailed here. In Rosenfeld et al.'s (2004) Experiment 1, guilty participants ($n = 22$) participated in one of two mock crime scenarios, one involving the theft of a ring from a desk drawer and the other involving the removal of an official university grade roster from its posted location. Innocent participants ($n = 11$) did not participate in either scenario. To ensure encoding of crime details, guilty participants rehearsed the instructions for their scenario several times and were required to pass a test pertaining to the critical details before undergoing the ERP-based GKT procedure. The test stimuli were a series of single words from various categories, some representing crime-relevant details (probes; 17% of trials) and others irrelevant category words, with some of these designated targets (17% of trials) and the remainder nontargets (66% of trials). As in Farwell and Donchin (1991), participants were instructed to press one button whenever a target word appeared and a different button whenever a nontarget word (whether crime-relevant or irrelevant) appeared. Half of the guilty participants received training in CM use prior to the GKT, and the others did not. CMs entailed specific covert responses (e.g., pressing a finger against a leg; wiggling a big toe; imagining being slapped) following the appearance of irrelevant nontarget words in order to enhance brain responses to these stimuli.

Two bootstrap techniques were used to classify participants as guilty or innocent on the basis of ERP responses. One of these involved comparing the estimated correlation between probe and target response with that between probe and nontarget response (cf. Farwell & Donchin, 1991). The other entailed comparing the difference in mean amplitude of response to probes versus irrelevant nontargets for each individual participant against a distribution of difference scores formed

by iterative resampling of the available data. Within each analysis, participants were classified as innocent if a significant difference in the direction indicating guilt was not obtained (differences were evaluated in terms of both base-to-peak and peak-to-peak amplitude; results for the more effective, peak-to-peak score analysis are described here).

Using the correlation-difference method, 10 of 11 innocent participants (90.9%) were correctly classified, but only 6 of 11 in each of the simple-guilty and guilty-CMs groups (54.5%) were correctly classified. For the amplitude-difference method, 10 of 11 innocent participants (90.9%) and 8 of 11 simple-guilty participants (72.7%) were correctly classified, but only 2 of 11 guilty-CMs participants (18.2%) were correctly classified. In Experiment 2, hit rates for guilty-CMs participants were: correlation difference method, 3/12 (25%); amplitude-difference method, 6/12 (50%). Comparative hit rates without CMs for these same participants were 62.9% and 92.3% when tested prior to instruction in and use of CMs and 25% and 58.3% when tested again after instruction/use of CMs.

Some interpretive difficulties are evident in this study. No inconclusive category was employed in classifications, making it difficult to compare these findings with those of Farwell and Donchin (1991). In addition, the hit rate in Rosenfeld et al.'s (2004) Experiment 1 for simple-guilty participants based on the correlation-difference method ($6/11 = 55\%$) was substantially and inexplicably lower than the rate for guilty participants in Experiment 1 of the Farwell and Donchin study; even with inconclusives considered as incorrect, the hit rate across the two experiments in this earlier study was $22/24 = 91.7\%$. This comparatively unimpressive hit rate for non-CM participants in this experiment clouds interpretation of the low hit rate for CM participants. Interpretation of CM effects in Experiment 2 was likewise complicated by differences in non-CM hit rates across for two separate comparison sessions as well as by the artificiality and narrowness of the test procedure (i.e., focus on a single biographic detail). Notwithstanding these limitations, the Rosenthal et al. study was important in raising concerns about deliberate CMs being used to beat an ERP-based detection test and inspiring further research on this topic.

The next published investigation of the impact of CMs was conducted by Mertens and Allen (2008), whose results from the no-CM guilty condition were summarized in the preceding section. This study was notable for its highly realistic virtual reality theft scenario that served as the crime manipulation, inclusion of multiple CM conditions (mental CMs to target stimuli, physical CMs to target stimuli, and alternating physical and mental CMs to distracter stimuli), and comparison of differing approaches to the scoring of test data. Whereas the hit rate for non-CM guilty participants based on the optimal method of scoring (either peak-to-peak or Bayesian) was 47%, the maximum hit rate for any of the CM conditions achieved by any method of scoring was only 27%.

A third, bootstrapped cross-correlation scoring method produced very high rates of indeterminate decisions (56%–93%) in all study conditions, including the innocent condition. As with the Rosenfeld et al. (2004) study, the modest detection

rate for non-CM guilty participants reported by Mertens and Allen (2008; which, by implication, casts further doubt on the 100% across-the-board accuracy rate for MERMER-based detection claimed by Farwell [2012]) complicates interpretation of the low detection rate reported for CM participants. Nonetheless, the hit rates for CM groups in this study were significantly lower than rates for the non-CM group, corroborating Rosenfeld et al.'s (2004) conclusion that CMs can be effective in reducing the accuracy of ERP-based detection.

In response to this emerging evidence for the effectiveness of CMs, Rosenfeld et al. (2008) sought to develop an alternative, CM-resistant ERP test protocol. In this procedure, termed the complex trial protocol (CTP), two stimuli are presented in sequence on each test trial, separated by a varying interstimulus interval of ~1 to 2 seconds. The first (S1) consists of either a rare probe stimulus (20% probability), relevant to the matter under investigation, or a frequent irrelevant stimulus (80%). To this initial stimulus, the participant responds with a standard designated button press, regardless of stimulus type, to signify registration of the S1. The second stimulus (S2) appears within 1.2 to 1.8 seconds after offset of the first and comprises either a target or non-target stimulus calling for differential button press responses. The intent of the CTP procedure is to separate the processing of relevant and irrelevant test stimuli from the discriminative response task required to generate a referent against which to compare responses to test stimuli, in the form of target P3 response. Although delayed, the target (S2) part of the task serves to maintain attention and ensure compliance with the task.

Rosenfeld et al. (2008) reported this task to be highly accurate in the detection of personal (birth date) information, yielding correct decisions for 12 of 12 non-CM participants in two separate studies when an individualized ("flexible") search window was used to define P300 and correct decisions for 11 of 12 participants trained to use the same types of CMs employed by Rosenfeld et al. (2004). These authors found that the use of CMs led to enhancement of P300 responses to both probe and irrelevant S1 stimuli, which they attributed to increased attentional processing required to decide whether to initiate CMs upon presentation of each stimulus. To test the hypothesis that the observed P300 differentiation between probes and targets for CM participants in this task might be attributable in part to omission of CMs selectively for probe stimuli, Meixner and Rosenfeld (2010) instructed participants to selectively omit CMs for one of five S1 stimuli in a CTP while employing CMs with the others and demonstrated the occurrence of an enhanced P300 response to the omit-CM stimulus relative to the employ-CM stimuli. Further, they found that, if the omit-CM stimulus was personally meaningful in some way, the degree of response augmentation for this stimulus was markedly larger. In subsequent work, Rosenfeld and Labkovsky (2010) showed that, although use of CMs for some but not all irrelevant stimuli in a CTP (i.e., for two of four irrelevants, as opposed to four of four) resulted in elimination of this omission-enhancement effect, the detection rate for CM guilty participants under these instructional conditions was nonetheless very high (100%).

A potential weakness of the CTP (noted by Farwell, 2012) is that participants are not required to differentiate behaviorally between probe and irrelevant stimuli presented as S1s (i.e., a common button press response is made in each case). Consequently, it is conceivable that participants motivated to defeat the test could avoid processing S1 stimuli in the test beyond the level of detecting visible changes in the foreground display associated with their occurrence and responding accordingly (i.e., without registering the distinct features of individual stimuli).

A contrasting perspective on this issue comes from work by Lui and Rosenfeld (2009) demonstrating enhancement of P300 response to a dishonestly answered probe stimulus when preceded by a subliminal presentation of the stimulus (i.e., very brief occurrence, in the context of a surrounding visual mask). The authors reported an overall accuracy rate of 86% for this method across guilty and innocent participants in this study. Although this study did not examine the impact of CMs on detectability using this method, it will be interesting in future work to examine whether a subliminal priming manipulation can be incorporated into a CTP procedure in a manner designed to protect against deliberate inattentiveness to S1 stimuli. More broadly, it will be important to evaluate the effectiveness of methods such as CTP and subliminal priming in more highly realistic experimental scenarios such as that used by Mertens and Allen (2008).

Other ERP Response Components. In addition to P300, other components of brain potential response have been applied to the detection of deception. One is the N400 response that reliably occurs in response to semantic incongruity (i.e., words that complete a sentence in an unexpected fashion; Kutas & Hillyard, 1980). Boaz, Perry, Raney, Fischler, and Shuman (1991) developed an N400-based GKT procedure in which participants, after viewing either a tape of an enacted burglary or a noncrime control tape, were presented with crime-relevant phrases that concluded with either true or false endings. Hit rates in this study (73.2% overall in cross-validation samples) were markedly lower than in most P300-based GKT studies to date.

Subsequently, Fang, Liu, and Shen (2003) explored the use of contingent negative variation (CNV) in detection of deception. The CNV is a slow negative shift in electroencephalogram potential that develops during anticipation of a target stimulus following presentation of a warning cue. Fang et al. examined CNV in a task in which participants were first presented with face stimuli and then, upon presentation of a follow-up signal, indicated whether the face was familiar to them or not. These authors reported significantly enhanced CNV on trials in which participants prepared to enact a false response compared with trials on which they responded truthfully. The comparative promise of this method for detecting deception is difficult to evaluate, because no effort was made to classify participants as truthful or deceptive on the basis of brain response. Further work with this type of procedure is needed to evaluate its effectiveness at the level of individuals.

Two other trends in the use of brain response measures to detect deception are noteworthy. One consists of studies designed to link differing components of the ERP to specific cognitive *processes* underlying deception (cf. Furedy, Davis, & Gurevich, 1988). Along this line, R. Johnson, Barnhardt, and Zhu (2003) reported evidence for two distinct components of the ERP connected with the act of deception, one reflecting inhibition of the prepotent (truthful) response and the other reflecting monitoring of past truthful and deceptive responses (see also R. Johnson, Barnhardt, & Zhu, 2004). Subsequently, R. Johnson, Barnhardt, and Zhu (2005) reported differential effects of practice (i.e., trial repetitions) on reaction time and ERP parameters of response to test questions—indicating that cognitively mediated response conflicts underlying deceptive behaviors are resistant to practice effect, in a manner similar to perceptually driven response conflicts.

In contrast with the ERP studies just reviewed, the focus of work of this kind is on gaining insights into the dynamics of neurocognitive processing associated with deception rather than on classifying individuals as truthful or deceptive based on ERP parameters that discriminate these conditions empirically. Notably, this has been a prominent focus to date in neuroimaging studies of deception, reviewed in the next subsection.

NEUROIMAGING-BASED DETECTION METHODS

A major development over the past decade has been the growing use of neuroimaging measurement in research on deception. Studies of this kind have utilized the technique of fMRI, in which changes in blood flow within specific regions of the brain are indexed by perturbations in a magnetic field surrounding the head, or in some cases (e.g., Abe, Suzuki, Mori, Itoh, & Fujii, 2007; Abe et al., 2006) positron emission tomography (PET), an imaging technique in which neural activity in specific brain regions is indexed through measurement of subatomic particles emitted by a radioactive isotope injected into the brain.

As in the ERP work of Johnson and colleagues (Johnson, 2003, 2004), many studies of this type have focused on processes associated with deception (and affiliated brain regions) rather than on classifying participants as deceptive or truthful. The first such study was by Spence et al. (2001), which reported enhanced activation in the ventrolateral prefrontal cortex (Brodmann area 47) bilaterally when participants lied about activities they had performed earlier in the day. This activation was interpreted as reflecting an inhibitory process associated with the effort to withhold the truth. Two subsequent studies reported increased activity in a wider array of brain regions (including frontal/prefrontal, parietal, and temporal cortices) when participants lied to critical items on a GKT (Langleben et al., 2002) or a GKT-like memory test (Lee et al., 2002).

In another early study (Ganis, Kosslyn, Stose, Thompson, & Yurgelun-Todd, 2003), researchers examined activations associated with two distinct parameters

of a lie: (1) whether it is spontaneous or rehearsed and (2) whether it is isolated or part of a broader story the participant is telling. Well-rehearsed lies connected to a broader narrative evoked greater activation in right anterior frontal cortex than spontaneous isolated lies, whereas spontaneous isolated lies elicited greater activation in anterior cingulate and posterior visual cortices. Lies of both types evoked greater activation (versus truth-telling) in right and left anterior prefrontal cortex and parahippocampal gyrus, right precuneus, and left cerebellum. These findings indicate that different brain regions are recruited in the context-different forms of lying activity.

In the decade or so since publication of these initial studies, the neuroimaging literature on deception has expanded rapidly. A pervasive finding has been increased activity in regions of the prefrontal cortex—including bilateral dorsolateral and anterior regions (middle and superior frontal gyri) and inferior frontal regions—during deception or concealment of information (Abe, 2009). Other brain regions that have been implicated in deception include the angular gyrus, caudate nucleus, and supplementary motor area. A major question arising from this work has to do with the specificity of the role that these brain regions play in deception. Prominent investigators in this area (e.g., Kozel, Padgett, & George, 2004; Langleben et al., 2005) have argued that brain activations indexed by neuroimaging are more revealing of the basic cognitive processes underlying deceptive responding than peripheral response measures, which index more generic bodily activation. Alternatively, it is conceivable that activations reliably reported in neuroimaging studies of deception reflect engagement of brain systems that play a supportive role in many contexts calling for concentrated attention and cognitive control—as opposed to systems that mediate deception or information concealment per se (cf. Gamer, 2011).

In addition to studies focusing on identifying brain regions associated with the act of deception, a number of studies have examined the effectiveness of fMRI-based assessment for classifying individuals as deceptive versus truthful. An initial study along this line by Kozel et al. (2004) examined the consistency with which particular brain regions were activated across participants during lying as compared to truth telling. Some degree of consistency was evident, in anterior brain regions in particular, encouraging further work.

In a follow-up study involving more than 60 participants, Kozel et al. (2005) evaluated the accuracy of fMRI-based testing for classifying individuals as truthful or deceptive in relation to the commission of a mock theft. Participants stole a watch or a ring under instruction and then underwent a test protocol that resembled a control question polygraph test sequence. The test included crime-relevant questions, neutral questions dealing with facts and personal preferences, and control questions dealing with illegal or rule-breaking behaviors of differing types. Data for half the sample were used to identify patterns of brain activation that differentiated deception from truthful responding. This resulted in three anatomic clusters (centered around the right orbitofrontal/inferior frontal cortex, right middle

frontal gyrus, and anterior cingulate cortex) being selected as discriminating. These clusters were then used as regions of interest for classifying participants in the remainder of the sample.

More specifically, the number of activated voxels for each of these areas in the deceptive versus truthful condition was calculated for each participant in the second half-sample with reference to an a priori statistical threshold. The resulting difference score was then used to classify each participant as deceptive or truthful with respect to one or the other mock theft based on whether the difference was significant in a positive or a negative direction. Based on this approach, 28 of 31 participants (93%) were classified correctly with respect to the theft they had committed (watch or ring). Corresponding rates in three subsequent replication samples (Kozel, Johnson, et al., 2009; Kozel, Laken, et al., 2009) were 71%, 93%, and 86%, for an average rate of 85.8% across all four cross-validation samples, including that of Kozel et al. (2005).

Findings such as these have engendered considerable enthusiasm around the possibilities of “direct” brain-based detection of deception. Alongside the growing cadre of studies in this area, commercial enterprises have surfaced in the United States that offer neuroimaging-based detection of deception services (e.g., Cephos Corporation, www.cephoscorp.com; No Lie MRI, Inc., www.noliemri.com). Stephen Laken, who was a coauthor on the Kozel reports and founder of Cephos, unsuccessfully testified in court to have exculpatory fMRI results admitted in *United States v. Semrau* (2012; see Shen & Jones, 2011, for analysis of this decision). In *State of Maryland v. Gary Smith* (2012, see Shen & Jones, 2011), scientists associated with No Lie MRI attempted to achieve the same outcome for the defendant in state court but were similarly unsuccessful. Considering that most of the available published research on the use of fMRI-based testing for classifying individuals as deceptive versus truthful has utilized basic “Did you do it?” question formats analogous to conventional RIT or CQT test procedures, this trend toward the use of a discredited questioning format is cause for some concern. As noted, it is quite conceivable that the brain activations that differentiate deceptive and truthful conditions in lab studies are indicative of more general cognitive processes such as focused attention, working memory, and cognitive control. It is also unclear to what extent activation of similar regions might occur in innocent individuals responding to “Did you do it?” questions under the conditions of uncertainty and anxiousness that tend to characterize real-life detection tests. The one individual-classification study that included a no-crime innocent group (Kozel, Johnson, et al., 2009) reported an accuracy rate of only 8 out of 21 (38%) for a CQT-type test in this group. This result casts doubt on the effectiveness of standard detection test protocols with innocent participants, even when based on functional neuroimaging methodology. Others have also expressed considerable skepticism regarding the evidentiary value of fMRI findings in court (Bizzi et al., 2009; Greeley & Illes, 2007; Wagner, 2010).

Some more recent evidence, however, does indicate that neuroimaging-based detection can achieve higher rates of classification of innocent subjects through

the use of a GKT-like (concealed information test) format. The first study to examine brain activations associated with deception in a GKT-type test was one by Gamer, Bauermann, Stoeter, and Vossel (2007). This study focused on average condition effects rather than classification of individuals. However, a subsequent study by Nose, Murai, and Taira (2009) that focused on classification of individual participants in a GKT-type test procedure reported accuracy rates of 84% for both guilty/deceptive and innocent/truthful participants. Further research is needed to evaluate whether this more costly and technically demanding approach to GKT testing carries any advantage relative to ERP-based or more conventional autonomic-response based testing.

THERMAL IMAGING

Thermal imaging has also been used to detect deception by employing a high-speed motion picture camera sensitive to rapid changes in facial regional blood flow. For example, in a mock crime study by Pavlidis, Eberhardt, and Levine (2002), 6 of 8 guilty and 11 of 12 innocent subjects were correctly identified based on an undescribed "thermal signature," apparently involving changes in blood flow around the eyes, in relation to an incident involving theft of \$20.

This method is intriguing, because it may be possible to use it without the subject's knowledge, potentially under conditions of remote testing (e.g., through a computer-video interface). A high-profile example of an application of this kind came to the attention of the public in 2011 when it was announced that thermal imaging technology would be tested in an undisclosed U.K. airport as a security screening device. That same year, an empirical study was published that evaluated the accuracy of thermal imaging for this specific purpose (Warmelink et al., 2011). In this study, 51 passengers in an international airport either lied or told the truth about a forthcoming trip in an on-site interview conducted by study experimenters. Skin temperature was recorded using a thermal imaging camera. On the basis of increases in facial skin temperature during the interview, 69% of deceptive participants and 64% of truthful participants were classified correctly. The authors noted that judgments of veracity made by interviewers after interacting with participants achieved higher rates of accuracy (77% and 72%, respectively) than the thermal imaging-based classifications. The authors concluded that thermal imaging is unlikely to have incremental validity over standard questioning methods for purposes of airport screening.

More research is needed to evaluate this technique in other contexts and to determine whether it might be vulnerable to many of the same criticisms leveled at conventional polygraph tests. In particular, it seems likely that, in real-life testing situations, a considerable portion of falsely accused innocent people would show heightened facial blood flow when asked a threatening question they answer honestly. The results of the Warmelink et al. (2011) study, which yielded a hit rate of only 64% for innocent participants, appear consistent with this possibility.

VOICE STRESS ANALYSIS

One technique unlikely to be of any value in the detection of deception is voice stress analysis. Recent heightened concerns about security have led to an increase in interest in this technique, which involves analyzing a sample of human speech for effects presumed due to inaudible microtremors of the vocal muscles reflective of the stress of lying. The advantage of voice stress analysis is that it can be used with recorded or broadcast speech without the subject's knowledge. The disadvantage is that, despite 30 years of research, there is virtually no evidence for its scientific basis or that it accurately detects lying (NRC, 2003).

THE POLYGRAPH IN COURT

Two important considerations in courtroom presentation of polygraph findings are the admission of polygraph testimony into evidence and how juries evaluate this evidence.

ADMISSION OF POLYGRAPH TESTIMONY

Polygraph tests often find their way into criminal court through one of two routes. One involves the stipulated test in which polygraph examinations are administered with the prior agreement of prosecutor and defense attorney. Often the prosecution will agree to a stipulated test when the case against the defendant is weak. In these circumstances, if the suspect passes the test, the charges are dismissed. If the suspect fails the test, the prosecution reserves the right to submit the polygraph findings to the court. About half of U.S. states endorse the use of stipulated tests, but Canadian courts refuse them.

Another way that polygraph results may enter a courtroom is over the objection of the prosecution in cases where it is argued that polygraph results constitute valid scientific evidence. This practice is allowed by law in New Mexico (*Lee v. Martinez*, 2004), provided the polygraph test administration satisfies certain standards. It is also a strategy increasingly adopted by defense attorneys who wish to determine if current circumstances favor the admission of a polygraph test that they have arranged for a client who subsequently passed. Often a hearing is requested before a judge who is asked to determine if polygraph tests satisfy standards for scientific evidence in light of new laws and rulings and/or in light of recent developments in the field (e.g., computerization) that may indicate polygraphy has been improved significantly since the last time the court considered admitting such evidence.

In 1923, in *Frye v. United States*, the U.S. Supreme Court established the rules for determining the admissibility of testimony based on novel scientific techniques in federal proceedings. In this case, James Frye was denied the opportunity to have considered as evidence the results of a polygraph test administered by psychologist William Marston, the "father" of modern polygraphy. Although the *Frye* ruling no longer controls federal proceedings, it is still influential to the laws of some states

that followed the *Frye* precedent of requiring “general acceptance” of a technique by the relevant scientific community before testimony based on the technique can be admitted. In federal courts as well as in many state courts, the standards that control are those laid out by the U.S. Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals* (1993). These standards direct that judges are to consider the admissibility of testimony based on newly developed scientific procedures after consideration of a number of factors including, but not limited to, whether the procedure (a) is supported by scientific theory, (b) has been subjected to peer review, (c) has a known error rate, (d) is governed by uniform standards, and (e) is generally accepted in the relevant scientific community.

Hence, following motions submitted by defense attorneys, many courts hold hearings based on principles outlined in *Frye* or *Daubert* to determine if testimony based on a defendant’s passed polygraph test should be admitted as evidence (see Faigman, Saks, Sanders, & Cheng, 2009, for a more thorough review of the legal status of polygraph testing in the United States). Such hearings are likely to be influenced by the Supreme Court’s decision in *United States v. Scheffer* (1998), in which it ruled that defendants in military court martial proceedings do not have a right to admit as evidence the results of exculpatory polygraph tests, based on the justices’ ruling that there is no consensus in the scientific community that polygraph evidence is valid.

When a defense attorney arranges for a client to take a polygraph test, the results of the test are protected by attorney–client privilege. If the defendant fails the test, the results would not be divulged, because doing so would only serve to undermine the defendant’s credibility. A test administered under these circumstances is considered to be “friendly.” Such a test stands in contrast to an “adversarial” test administered by the law enforcement personnel, the results of which would be made known to the prosecution and defense. Because fear of the consequences of being detected is considered to be important to the valid outcome of a test, and there appears to be less to lose and therefore less to fear with a friendly test, it seems likely that friendly tests would be easier to pass than adversarial tests. Moreover, because the defendant is paying the polygrapher with the hope of passing the test, the examiner is being pressured, at least by the defendant, to produce the desired outcome. In a procedure that is as subjective and unstandardized as the CQT, it is easy to imagine how subtle adjustments to the procedure could increase the likelihood of friendly tests being passed. Unfortunately, there are no empirical studies attesting to the validity of friendly tests. All the existing field studies deal with adversarial tests.

HOW JURIES EVALUATE POLYGRAPH EVIDENCE

An important issue surrounding the use of polygraph evidence in court is the weight that is likely to be attached to this evidence by juries. This concern derives in part from Rule 403 of the Federal Rules of Evidence (and its state court equivalents; see Daniels, 2002), which allows courts to exclude evidence if its probative value is

substantially outweighed by the prejudicial impact it may have on the jury. Unlike other types of evidence a jury may hear, polygraph evidence has the potential to usurp the jury's constitutionally mandated task of deciding guilt. Thus, courts have also excluded polygraph testimony on the grounds that the scientific and technical aura that surrounds the practice of polygraph testing may lead juries to assign excessive probative weight to this evidence (see, e.g., *United States v. Alexander*, 1975).

Since our review of how juries consider polygraph evidence in the last edition of this text, one new study has appeared (Myers, Latter, & Abdollahi-Arena, 2006). This study and those that have preceded it suggest that mock juries are skeptical of polygraph test results. However, just as field studies are needed to estimate the accuracy of polygraph tests, it would be worthwhile to have data from polled jurors following trials in which polygraph testimony was offered to determine how jurors weighed this evidence in actual legal proceedings.

SCIENTIFIC OPINION

The opinions of scientists regarding polygraphy are obviously important. Conventional polygraph tests have a weak conceptual foundation. Moreover, serious methodological problems that are unlikely to be easily overcome make it unlikely that any line of research will yield findings that resolve concerns about accuracy. Given this state of affairs, there is considerable value in the broad-based sampling of the opinions of scientists with the background and expertise to evaluate polygraph tests. In addition, the *Frye* and *Daubert* decisions make clear that the views of the scientific community about the general acceptance of a technique are important to considering admissibility of testimony based on the technique.

Only one investigation of scientific opinion regarding polygraph techniques has been published in a scientific peer-reviewed journal (Iacono & Lykken, 1997). This study polled members of the Society for Psychophysiological Research and fellows in Division 1 (General Psychology) of the American Psychological Association (APA). High response rates (>74%) were obtained from those in both organizations, and there was remarkable agreement across groups regarding CQT polygraphy. These scientists expressed a high level of skepticism regarding the claims of polygraph proponents. They did not find the theory of the CQT to be scientifically sound or the accuracy claims of polygraph proponents to be credible. In addition, they expressed opinions indicating that friendly tests have little value and CMs pose a significant threat to the validity of passed tests. Members of neither group would recommend that testimony based on the results of CQTs be admitted in court.

Only APA members were asked about directed lie tests, and they did not agree that these tests are scientifically sound. In contrast to these negative opinions about conventional specific incident tests, those polled had favorable opinions about the GKT. The contrast in the scientific credibility of the CQT and GKT is important, because it indicates that respondents were not generally skeptical about

detection of deception techniques but have doubts that are specific to the CQT. The results of these surveys parallel the opinions of the NRC (2003) committee that reviewed polygraph test validity as well as those of many other scientists and professional societies at arm's length from the polygraph profession have conducted critical appraisals of the field (APA, 2004; Ben-Shakhar, 2002; British Psychological Association, 2004; Fiedler et al., 2002; Oksol & O'Donohue, 2003; Verschuere et al., 2008; Vrij, 2008).

CONCLUSION

Despite this scientific skepticism, the use of polygraph tests continues unabated, presumably reflecting beliefs among law enforcement and national security policy makers that their utility benefits outweigh concerns regarding costs associated with their misuse. There appears to be little dispute about the utility of polygraph testing, although only anecdotal, not scientific evidence, exists to support this contention (NRC, 2003). Nevertheless, many criminal suspects confess following failed tests, providing a means to resolve criminal investigations that otherwise would go unprosecuted.

In employee screening, the admissions employees make about their alcohol use, sex lives, and colleagues' suspect behavior provide the government with what is considered to be valuable information that would be virtually impossible to obtain via any other (legal) means. Likewise, those administering sex offender treatment programs have come to rely on polygraph tests to encourage offenders to divulge fully their past sexual misdeeds, so much so that the use of polygraph tests in these programs is now widespread. When used in such contexts, the polygraph is little more than a prop intended to encourage socially undesirable self-disclosure among those who believe it genuinely works, a phenomenon established over 40 years ago as the "bogus pipeline" effect (Jones & Sigall, 1971). However, as the NRC panel noted, in the long run, evidence that a technique lacks validity will eventually undercut its utility.

For many decades, polygraph testing has been part of the fabric of our institutions for law enforcement and national security. Consequently, reliance on polygraphy as an investigative tool is unlikely to diminish in the future. Although it remains possible that the CQT will become accepted as credible scientific evidence, courts have not shown a readiness to embrace the admission of specific incident tests in the first 15 years following *Daubert* (Faigman et al., 2009). As our review indicates, there is little evidence to support their admission, and what evidence does exist, coupled with the obvious weaknesses in CQT theory, indicates that the CQT has little more than chance accuracy with innocent people and can be easily defeated by guilty people who learn to augment their responses to control questions.

Although the GKT and the ERP-GKT appear to offer promising alternatives to the CQT (Ben-Shakhar, 2012; Ben-Shakhar, Bar-Hillel, & Kremnitzer, 2002; Iacono, 2011), research with these procedures has not focused on how to adapt them

successfully to field applications (Iacono, 2008b, 2011). fMRI and other methods that are not based on the GKT have produced a body of research that is vulnerable to the same criticisms that have been leveled against the CQT. As scientists who have worked in this area for over 30 years, we are struck by the fact that this literature has focused on pushing the technological prowess of fMRI while neglecting the importance of the strong research designs that a half century of CQT research has taught us are needed to credibly anchor validity claims for lie detection methods.

Despite a lack of adequate field study and standardized test protocols, the GKT is based on sound theory, and it is possible for a jury to weigh evidence regarding the adequacy of a GKT. Consider, for instance, how much weight might be assigned a properly conducted GKT indicating the presence of guilty knowledge. Assume suspect John Fisbee is asked to preapprove the questions on a 12-item GKT by indicating whether he knows the answer to any of the questions, and he claims no knowledge. In addition, after the test is administered, he is asked if he can guess the answers to any of the items, and the two items he “guesses” the correct answer to are eliminated from further consideration. The test is given by an examiner who is unaware of the correct answers. On the GKT, Fisbee shows the strongest physiological response to all of the guilty alternatives for the remaining questions. When the same test was given to 10 individuals, none of whom could be involved in the crime, they responded to the guilty alternatives at chance levels. Because it is difficult to understand how such an outcome could come about in the absence of Fisbee’s guilty knowledge, such a test result provides relatively strong *prima facie* evidence of guilt. One can alter aspects of this hypothetical scenario in various ways (e.g., Fisbee fails 8 of the remaining 10 items), but with each alteration, it is possible to make a scientifically informed appraisal regarding the level of confidence one can have in the outcome. By contrast, passing a GKT is much more difficult to interpret because the field research needed to determine what those who commit crimes are likely to remember has not been conducted. Until this work is carried out, passed GKTs will remain suspect.

REFERENCES

- Abe, N. (2009). The neurobiology of deception: Evidence from neuroimaging and loss-of-function studies. *Current Opinion in Neurology*, *22*, 594–600.
- Abe, N., Suzuki, M., Mori, E., Itoh, M., & Fujii, T. (2007). Deceiving others: Distinct neural responses of the prefrontal cortex and amygdala in simple fabrication and deception with social interactions. *Journal of Cognitive Neuroscience*, *19*, 287–295.
- Abe, N., Suzuki, M., Tsukiura, T., Mori, E., Yamaguchi, K., & Itoh, M. (2006). Dissociable roles of prefrontal and anterior cingulate cortices in deception. *Cerebral Cortex*, *16*, 192–199.
- Allen, J. J. B. (2002). The role of psychophysiology in clinical assessment: ERPs in the evaluation of memory. *Psychophysiology*, *39*, 261–280.
- Allen, J. J. B., & Iacono, W. G. (1997). A comparison of methods for the analysis of event-related potentials in deception detection. *Psychophysiology*, *34*, 234–240.

- Allen, J. J. B., & Iacono, W. G. (2001). Assessing the validity of amnesia in dissociative identity disorder: A dilemma for the DSM and the courts. *Psychology, Public Policy, and Law*, 7(2), 311–344.
- Allen, J. J. B., Iacono, W. G., & Danielson, K. D. (1992). The identification of concealed memories using the event-related potential and implicit behavioral measures: A methodology for prediction in the face of individual differences. *Psychophysiology*, 29, 504–522.
- Allen, J. J. B., & Movius, H. L. (2000). The objective assessment of amnesia in dissociative identity disorder using event-related potentials. *International Journal of Psychophysiology*, 38, 21–41.
- Allen, J. J., Iacono, W. G., Laravuso, J. J., & Dunn, L. A. (1995). An event-related potential investigation of posthypnotic recognition amnesia. *Journal of Abnormal Psychology*, 104(3), 421–430.
- American Psychological Association. (2004, August 5). The truth about lie detectors (aka polygraph tests). Retrieved from <http://www.apa.org/research/action/polygraph.aspx>
- Barland, G. H., Honts, C. R., & Barger, S. D. (1989). Studies of the accuracy of security screening polygraph examinations. Fort McClellan, AL: Department of Defense Polygraph Institute.
- Ben-Shakhar, G. (2002). A critical review of the control questions test (CQT). In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 103–126). San Diego, CA: Academic Press.
- Ben-Shakhar, G. (2012). Current research and potential applications of the concealed information test: An overview. *Frontiers in Psychology*, 3, 342. doi:10.3389/fpsyg.2012.00342
- Ben-Shakhar, G., Bar-Hillel, M., & Kremnitzer, M. (2002). Trial by polygraph: Reconsidering the use of the guilty knowledge technique in court. *Law and Human Behavior*, 26(5), 527–541.
- Ben-Shakhar, G., & Elaad, E. (2002). Effects of questions' repetition and variation on the efficiency of the guilty knowledge test: A reexamination. *Journal of Applied Psychology*, 87(5), 972–977.
- Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology*, 88, 131–151.
- Bizzi, E., Hyman, S. E., Raichle, M. E., Kanwisher, N., Phelps, E. A., Morse, S. J., . . . Greely, H. T. (2009). *Using imaging to identify deceit: Scientific and ethical questions*. Cambridge, MA: American Academy of Arts & Sciences.
- Boaz, T. L., Perry, N. W., Raney, G., Fischler, I. S., & Shuman, D. (1991). Detection of guilty knowledge with event related potentials. *Journal of Applied Psychology*, 76, 788–795.
- British Psychological Association. (2004, October 6). *A review of the current scientific status and fields of application of polygraphic deception detection*. Leicester, England: Author.
- Daniels, C. W. (2002). Legal aspects of polygraph admissibility in the United States. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 327–338). San Diego, CA: Academic Press.
- Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993).
- Department of Defense Polygraph Institute Research Division (1997). A comparison of psychophysiological detection of deception accuracy rates obtained using the counter-intelligence scope polygraph and the test for espionage and sabotage. *Polygraph*, 26, 79–106.

- Department of Defense Polygraph Institute Research Division. (1998). *Psychophysiological detection of deception accuracy rates obtained using the test for espionage and sabotage*. Ft. McClellan, AL: Department of Defense Polygraph Institute.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7, 1–26.
- Elaad, E. (1990). Detection of guilty knowledge in real-life criminal investigation. *Journal of Applied Psychology*, 75, 521–529.
- Elaad, E., Ginton, A., & Jungman, N. (1992). Detection measures in real-life criminal guilty knowledge tests. *Journal of Applied Psychology*, 77, 757–767.
- Employee Polygraph Protection Act, Pub. L. 100–347 (1988).
- Faigman, D. L., Saks, M. J., Sanders, J., & Cheng, E. K. (2009). The legal relevance of scientific research on polygraph tests. In D. L. Faigman, M. J. Saks, J. Sanders, & E. K. Cheng (Eds.), *Modern scientific evidence: The law and science of expert testimony* (Vol. 5, pp. 266–297). Eagan, MN: Thomson Reuters/West.
- Fang, F., Liu, Y., & Shen, Z. (2003). Lie detection with contingent negative variation. *International Journal of Psychophysiology*, 50, 247–255.
- Farwell, L. A. (2008). Brain fingerprinting detects real crimes in the field despite one-hundred-thousand-dollar reward for beating it. *Psychophysiology*, 45, S104.
- Farwell, L. A. (2012). Brain fingerprinting: A comprehensive tutorial review of detection of concealed information with event-related brain potentials. *Cognitive Neurodynamics*, 6, 115–154.
- Farwell, L. A., & Donchin, E. (1986). The “Brain Detector”: P300 in the detection of deception. *SPR Abstracts*, 23, 434.
- Farwell, L. A., & Donchin, E. (1991). The truth will out: Interrogative polygraphy (“lie detection”) with event related brain potentials. *Psychophysiology*, 28, 531–547.
- Farwell, L. A., Hernandez, R. S., & Richardson, D. C. (2006). Brain fingerprinting in laboratory conditions. *Psychophysiology*, 43, S38.
- Farwell, L. A., Richardson, D. C., & Hernandez, R. S. (2006b). Brain fingerprinting in field conditions. *Psychophysiology*, 43, S38.
- Farwell, L. A., Richardson, D. C., & Richardson, G. (2011). Brain fingerprinting field studies comparing P300-MERMER and P300 ERPs in the detection of concealed information. *Psychophysiology*, 48, S95–S96.
- Farwell, L. A., & Smith, S. S. (2001). Using brain MERMER testing to detect knowledge despite efforts to conceal. *Journal of Forensic Sciences*, 46, 1–9.
- Ferguson, R. J. (1966). *The polygraph in private industry*. Springfield, IL: Charles C Thomas.
- Fiedler, K., Schmod, J., & Stahl, T. (2002). What is the current truth about polygraph lie detection? *Basic & Applied Social Psychology*, 24, 313–324.
- Frye v. United States, 293 F. 1013 (D.C. Cir. 1923).
- Furedy, J. J., Davis, C., & Gurevich, M. (1988). Differentiation of deception as a psychological process: A psychophysiological approach. *Psychophysiology*, 25, 683–688.
- Gamer, M. (2011). Detecting of deception and concealed information using neuroimaging techniques. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Theory and application of the Concealed Information Test* (pp. 90–113). New York, NY: Cambridge University Press.
- Gamer, M., Bauermann, T., Stoeter, P., & Vossel, G. (2007). Covariations among fMRI, skin conductance, and behavioral data during processing of concealed information. *Human Brain Mapping*, 28, 1287–1301.

- Ganis, G., Kosslyn, S. M., Stose, S., Thompson, W. L., & Yurgelun-Todd, D. A. (2003). Neural correlates of different types of deception: An fMRI investigation. *Cerebral Cortex*, *13*, 830–836.
- Ganis, G., & Patnaik, P. (2009). Detecting concealed knowledge using a novel attentional blink paradigm. *Applied Psychophysiology & Biofeedback*, *34*(3), 189–196. doi: 10.1007/s10484-009-9094-1
- Greeley, H. T., & Illes, J. (2007). Neuroscience-based lie detection: The urgent need for regulation. *American Journal of Law & Medicine*, *33*, 377–431.
- Harrington v. State of Iowa, 109 F. 3d 1275; Court of Appeals, 8th Cir. (1997).
- Honts, C. R. (1996). Criterion development and validity of the CQT in field application. *Journal of General Psychology*, *123*, 309–324.
- Honts, C. R., & Alloway, W. R. (2007). Information does not affect the validity of a comparison question test. *Legal & Criminal Psychology*, *12*, 311–320.
- Honts, C. R., & Amato, S. (2002). Countermeasures. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 251–264). London, England: Academic Press.
- Honts, C. R., & Raskin, D. C. (1988). A field study of the validity of the directed lie control question. *Journal of Police Science & Administration*, *16*, 56–61.
- Honts, C. R., Raskin, D., & Kircher, J. (1994). Mental and physical countermeasures reduce the accuracy of polygraph tests. *Journal of Applied Psychology*, *79*, 252–259.
- Honts, C. R., Raskin, D., & Kircher, J. (2002). The scientific status of research on polygraph techniques: The case for polygraph tests. In D. L. Faigman, D. H. Kaye, M. J. Saks, & J. Sanders (Eds.), *Modern scientific evidence: The law and science of expert testimony* (Vol. 2, pp. 446–483). St. Paul, MN: West.
- Horowitz, S. W., Kircher, J. C., Honts, C. R., & Raskin, D. C. (1997). The role of comparison questions in physiological detection of deception. *Psychophysiology*, *34*, 108–115.
- Horvath, F. (1977). The effect of selected variables on the interpretation of polygraph records. *Journal of Applied Psychology*, *62*, 127–136.
- Iacono, W. G. (1991). Can we determine the accuracy of polygraph tests? In J. R. Jennings, P. K. Ackles, & M. G. H. Coles (Eds.), *Advances in psychophysiology* (pp. 201–207). London, England: Jessica Kingsley.
- Iacono, W. G. (2007). Polygraph testing. In E. Borgida & S. T. Fiske (Eds.), *Beyond common sense: Psychological science in the courtroom* (pp. 219–235). Oxon, England: Blackwell.
- Iacono, W. G. (2008a). Accuracy of polygraph techniques: Problems using confessions to determine ground truth. *Physiology & Behavior*, *95*, 24–26.
- Iacono, W. G. (2008b). Effective policing: Understanding how polygraph tests work and are used. *Criminal Justice & Behavior*, *35*, 1295–1308.
- Iacono, W. G. (2008c). The forensic application of “brain fingerprinting”: Why scientists should encourage the use of P300 memory detection methods. *American Journal of Bioethics*, *8*, 30–32.
- Iacono, W. G. (2010). Lie detection. In I. B. Weiner & W. E. Craighead (Eds.), *The Corsini encyclopedia of psychology* (Vol. 2, pp. 928–930). Hoboken, NJ: Wiley.
- Iacono, W. G. (2011). Encouraging the use of the guilty knowledge test (GKT): What the GKT has to offer law enforcement. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the concealed information test* (pp. 12–23). New York, NY: Cambridge University Press.
- Iacono, W. G., & Lykken, D. T. (1997). The validity of the lie detector: Two surveys of scientific opinion. *Journal of Applied Psychology*, *82*, 426–433.

- Iacono, W. G., & Lykken, D. T. (2009). The case against polygraph tests. In D. L. Faigman, M. J. Saks, J. Sanders, & E. K. Cheng (Eds.), *Modern scientific evidence: The law and science of expert testimony* (Vol. 5, pp. 342–406). Eagan, MN: Thomson Reuters/West.
- Iacono, W. G., & Patrick, C. J. (1987). What psychologists should know about lie detection. In I. B. Weiner & A. K. Hess (Eds.), *The handbook of forensic psychology* (pp. 460–489). New York, NY: Wiley.
- Iacono, W. G., & Patrick, C. J. (1999). Polygraph (“lie detector”) testing: The state of the art. In A. K. Hess & I. B. Weiner (Eds.), *The handbook of forensic psychology* (2nd ed., pp. 440–473). New York, NY: Wiley.
- Iacono, W. G., & Patrick, C. J. (2006). Polygraph (“lie detector”) testing: Current status and emerging trends. In I. B. Weiner & A. K. Hess (Eds.), *The handbook of forensic psychology* (3rd ed., pp. 552–588). Hoboken, NJ: Wiley.
- Johnson, M. M., & Rosenfeld, J. P. (1992). Oddball-evoked P300-based method of deception detection in the laboratory II: Utilization of non-selective activation of relevant knowledge. *International Journal of Psychophysiology*, *12*, 289–306.
- Johnson, R., Barnhardt, J., & Zhu, J. (2003). The deceptive response: Effects of response conflict and strategic monitoring on the late positive component and episodic memory-related brain activity. *Biological Psychology*, *64*, 217–253.
- Johnson, R., Barnhardt, J., & Zhu, J. (2004). The contribution of executive processes to deceptive responding. *Neuropsychologia*, *42*, 878–901.
- Johnson, R., Barnhardt, J., & Zhu, J. (2005). Differential effects of practice on the executive processes used for truthful and deceptive responses: An event-related brain potential study. *Cognitive Brain Research*, *24*, 386–404.
- Jones, E. E., & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, *76*, 349–364.
- Kosson, D. S. (1988). Psychopathy and dual-task performance under focusing conditions. *Journal of Abnormal Psychology*, *105*, 391–400.
- Kozel, F. A., Johnson, K. A., Grenesko, E. L., Laken, S. J., Kose, S., & Lu, X. (2009). Functional MRI detection of deception after committing a mock sabotage crime. *Journal of Forensic Sciences*, *54*, 220–231.
- Kozel, F. A., Johnson, K. A., Mu, Q., Grenesko, E. L., Laken, S. J., & George, M. S. (2005). Detecting deception using functional magnetic resonance imaging. *Biological Psychiatry*, *58*, 605–613.
- Kozel, F. A., Laken, S. J., Johnson, K. A., Boren, B., Mapes, K. S., & Morgan, P. S. (2009). Replication of functional MRI detection of deception. *Open Forensic Science Journal*, *2*, 6–11.
- Kozel, F. A., Padgett, T. M., & George, M. S. (2004). A replication study of the neural correlates of deception. *Behavioral Neuroscience*, *118*, 852–856.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205.
- Langleben, D. D., Loughhead, J. W., Bilker, W. B., Ruparel, K., Childress, A. R., Busch, S. I., & Gur, R. C. (2005). Telling truth from lie in individual subjects with fast event-related fMRI. *Human Brain Mapping*, *26*, 262–272.
- Langleben, D. D., Schroeder, L., Maldjian, J. A., Gur, R. C., McDonald, S., Ragland, J. D., . . . Childress, A. R. (2002). Brain activity during simulated deception: An event-related functional magnetic resonance study. *Neuroimage*, *15*, 727–732.

- Lee, T. M., Liu, H. L., Tan, L. H., Chan, C. C., Mahankali, S., Feng, C. M., . . . Gao, J. H. (2002). Lie detection by functional magnetic resonance imaging. *Human Brain Mapping, 15*, 157–164.
- Lee v. Martinez, Supreme Court of New Mexico, 96 P.3d 291 (2004).
- Lui, M., & Rosenfeld, J. P. (2009). The application of subliminal priming in lie detection: Scenario for identification of members of a terrorist ring. *Psychophysiology, 46*, 889–903.
- Lykken, D. T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology, 43*, 385–388.
- Lykken, D. T. (1960). The validity of the guilty knowledge technique: The effects of faking. *Journal of Applied Psychology, 44*, 258–262.
- Lykken, D. T. (1974). Psychology and the lie detector industry. *American Psychologist, 29*, 725–739.
- Lykken, D. T. (1998). *A tremor in the blood: Uses and abuses of the lie detector* (2nd ed.). New York, NY: Plenum Press.
- MacLaren, V. V. (2001). A quantitative review of the guilty knowledge test. *Journal of Applied Psychology, 86*(4), 674–683.
- Mangan, D. J., Armitage, T. E., & Adams, G. C. (2008). A field study on the validity of the Quadri-Track Zone Comparison Technique. *Physiology & Behavior, 95*, 17–23.
- Meijer, E. H., Ben-Shakhar, G., Verschuere, B., & Donchin, E. (2013). A comment on Farwell (2012): Brain fingerprinting: A comprehensive tutorial review of detection of concealed information with event-related brain potentials. *Cognitive Neurodynamics, 7*, 155–158.
- Meijer, E. H., Verschuere, B., Merckelbach, H. L., & Crombez, G. (2008). Sex offender management using the polygraph: A critical review. *International Journal of Law and Psychiatry, 31*(5), 423–429.
- Meixner, J. B., & Rosenfeld, J. P. (2010). Countermeasure mechanisms in a P300-based concealed information test. *Psychophysiology, 47*, 57–65.
- Mertens, R., & Allen, J.J.B. (2008). The role of psychophysiology in forensic assessments: Deception detection. *Psychophysiology, 45*, 286–298.
- Myers, B., Latter, R., & Abdollahi-Arena, M. K. (2006). The court of public opinion: Lay perceptions of polygraph testing. *Law and Human Behavior, 30*(4), 509–523. doi:10.1007/s10979-006-9041-0
- Nakayama, M. (2002). Practical use of the concealed information test from criminal investigation in Japan. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 49–86). San Diego, CA: Academic Press.
- National Defense Authorization Act, Pub. L. 106–65 (2000).
- National Research Council. (2003). *The polygraph and lie detection*. Washington, DC: National Academies Press.
- Nose, I., Murai, J., & Taira, M. (2009). Disclosing concealed information on the basis of cortical activations. *Neuroimage, 44*, 1380–1386.
- Oksol, E. M., & O'Donohue, W. T. (2003). A critical analysis of the polygraph. In W. T. O'Donohue & E. R. Levensky (Eds.), *Handbook of forensic psychology: Resource for mental health and legal professionals* (pp. 601–634). San Diego, CA: Academic Press.
- Patrick, C. J., & Iacono, W. G. (1991a). A comparison of field and laboratory polygraphs in the detection of deception. *Psychophysiology, 28*(6), 632–638.
- Patrick, C. J., & Iacono, W. G. (1991b). Validity of the control question polygraph test: The problem of sampling bias. *Journal of Applied Psychology, 76*, 229–238.

- Pavlidis, I., Eberhardt, N. L., & Levine, J. A. (2002). Seeing through the face of deception: Thermal imaging offers a promising hands-off approach to mass security screening. *Nature*, *415*, 435.
- Raskin, D. (1989). Polygraph techniques for the detection of deception. In D. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 247–296). New York, NY: Springer.
- Research Division Staff, Department of Defense Polygraph Institute (1997). A comparison of psychophysiological detection of deception accuracy rates obtained using the counterintelligence scope polygraph and the test for espionage and sabotage. *Polygraph*, *26*, 79–106.
- Research Division Staff, Department of Defense Polygraph Institute (1998). Psychophysiological detection of deception accuracy rates using the test for espionage and sabotage. *Polygraph*, *27*, 68–73.
- Rosenfeld, J. P. (2002). Event-related potentials in the detection of deception, malingering, and false memories. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 265–286). New York, NY: Academic Press.
- Rosenfeld, J. P. (2005). “Brain fingerprinting”: A critical analysis. *Scientific Review of Mental Health Practice*, *4*, 20–37.
- Rosenfeld, J. P. (2011). P300 in detecting concealed information. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the concealed information test* (pp. 63–89). Cambridge, UK: Cambridge University Press.
- Rosenfeld, J. P., Angell, A., Johnson, M., & Qian, J. (1991). An ERP-based, control-question lie detector analog: Algorithms for discriminating effects within individuals’ average waveforms. *Psychophysiology*, *38*, 319–335.
- Rosenfeld, J. P., Cantwell, B., Nasman, V. T., Wodjdac, V., Ivanov, S., & Mazzeri, L. (1988). A modified, event-related potential-based guilty knowledge test. *International Journal of Neuroscience*, *42*, 157–161.
- Rosenfeld, J. P., & Ellwanger, J. W. (1999). Cognitive psychophysiology in detection of malingered cognitive deficit. In J. J. Sweet (Ed.), *Forensic neuropsychology: Fundamentals and practice* (pp. 287–312). Lisse, the Netherlands: Swets & Zeitlinger.
- Rosenfeld, J. P., & Labkovsky, E. (2010). New P300-based protocol to detect concealed information: Resistance to mental countermeasures against only half the irrelevant stimuli and a possible ERP indicator of countermeasures. *Psychophysiology*, *47*, 1002–1010.
- Rosenfeld, J. P., Labkovsky, E., Winograd, M., Lui, M. A., Vandenboom, C., & Chedid, E. (2008). The Complex Trial Protocol (CTP): A new, countermeasure-resistant, accurate, P300-based method for detection of concealed information. *Psychophysiology*, *45*, 906–919.
- Rosenfeld, J. P., Nasman, V. T., Whalen, R., Cantwell, B., & Mazzeri, L. (1987). Late vertex positivity in event-related potentials as a guilty knowledge indicator: A new method of lie detection. *Polygraph*, *16*(4), 258–263.
- Rosenfeld, J. P., Soskins, M., Bosh, G., & Ryan, A. (2004). Simple, effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology*, *41*, 205–219.
- Rule 403 Federal Rules of Evidence, Pub. L. 93–595, §1, January 2, 1975, 88 Stat. 1932; April 26, 2011, effective December 1, 2011.
- Samuels, D. J. (1983). What if the lie detector lies? *Nation*, *237*, 566–567.
- Sasaki, M., Hira, S., & Matsuda, T. (2001). Effects of a mental countermeasure on the physiological detection of deception using the event-related brain potentials. *Japanese Journal of Psychology*, *72*, 322–328.

- Seymour, T. L., & Fraynt, B. R. (2009). Time and encoding effects in the concealed knowledge test. *Applied Psychophysiology and Biofeedback, 34*(3), 177–187. doi:10.1007/s10484-009-9092-3
- Seymour, T. L., Seifert, C. M., Shafto, M. G., & Mosmann, A. L. (2000). Using response time measures to assess “guilty knowledge.” *Journal of Applied Psychology, 85*, 30–37.
- Shen, F. X., & Jones, O. D. (2011). Brain scans as evidence: Truths, proofs, lies, and lessons. *Mercer Law Review, 62*. 861–883.
- Soskins, M., Rosenfeld, J. P., & Niendam, T. (2001). Peak-to-peak measurement of P300 recorded at 0.3 Hz high pass filter settings in intraindividual diagnosis: Complex vs. simple paradigms. *International Journal of Psychophysiology, 40*, 173–180.
- Spence, S. A., Farrow, T. F. D., Herford, A. E., Wilkinson, I. D., Zheng, Y., & Woodruff, P. W. R. (2001). Behavioural and functional anatomical correlates of deception in humans. *NeuroReport, 12*, 2433–2438.
- State of Maryland v. Gary Smith, (2012). September Term, No. 10.
- Thurber, S. (1981). CPI variables in relation to the polygraph performance of police officer candidates. *Journal of Social Psychology, 113*, 145–146.
- United States v. Alexander, 526 F.2d 161. 168 (1975 (8th Cir.)).
- United States v. Scheffer, WL 141151 U.S. (1998).
- United States v. Semrau, No. 1:07-cr-10074-1. United States Court of Appeals, 6th Circuit (2012).
- Van Hoof, J. C., Brunia, C. H. M., & Allen, C. J. (1996). Event-related potentials as indirect measures of recognition memory. *International Journal of Psychophysiology, 21*, 15–31.
- Verschuere, B., Ben-Shakhar, G., & Meijer, E. (2011). *Memory detection: Theory and application of the concealed information test*. Cambridge, England: Cambridge University Press.
- Verschuere, B., Crombez, G., De Clercq, A., & Koster, E. H. (2005). Psychopathic traits and autonomic responding to concealed information in a prison sample. *Psychophysiology, 42*(2), 239–245. doi:10.1111/j.1469-8986.2005.00279.x
- Verschuere, B., Meijer, E., & Merkelbach, H. (2008). The quadri-track zone comparison technique: It’s just not science. A critique to Mangan, Armitage, and Adams (2008). *Physiology & Behavior, 95*, 27–28.
- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). West Sussex, England: Wiley.
- Wagner, A. D. (2010). Can neuroscience identify lies? *A judge’s guide to neuroscience: A concise introduction* (pp. 13–25). Santa Barbara: University of California.
- Waid, W. M., Orne, M. T., & Wilson, S. K. (1979). Effects of level of socialization on electrodermal detection of deception. *Psychophysiology, 16*, 15–22.
- Warmelink, L., Vrij, A., Mann, S., Leal, S., Forrester, D., & Fisher, R. P. (2011). Thermal imaging as a lie detection tool at airports. *Law and Human Behavior, 35*, 40–48.

Applying Hypnosis in Forensic Contexts

ALAN W. SCHEFLIN

FOR more than 160 years, American courts have been asked to resolve hypnosis issues (Gravitz, 1995). Until the 1970s, however, hypnosis was rarely a topic for judicial consideration. Prior to 1970, only 38 appellate cases involved hypnosis (Schefflin & Shapiro, 1989). Since 1970, over 1,000 appellate decisions have addressed five legal issues that arise when courts and legislatures consider the use of hypnosis in medical or forensic settings.

First are the fundamental questions of *who* may practice hypnosis and *how* hypnosis may be practiced. These issues involve the *regulation of hypnosis*, and they generally concern articulating the dividing line between licensed professionals and lay practitioners. The second area of forensic interest involves *hypnosis and antisocial conduct*. In this category are the fascinating questions about the abuse and misuse of hypnosis for the purpose of seduction or other criminal conduct (Laurence & Perry, 1988; Schefflin & Opton, 1978). The focus here is on the legal responsibility of hypnotizing influencers and of their hypnotized subjects who commit criminal acts.

The third and most heavily litigated intersection between law and hypnosis involves *hypnosis for memory recall* or, more accurately, whether a person who has been hypnotized is permitted to testify as a witness in court. The law on this topic is relatively recent. Before 1969, only two appellate cases involved hypnosis with memory. By contrast, since 1970, almost all of the approximately 1,000 cases involving hypnosis have dealt with the issue of the admissibility of hypnotically refreshed recollection (Schefflin, 1994a, 1994b). Although Kihlstrom (1985) barely mentioned forensic issues in his review of trends in hypnosis research, a decade later Sheehan (1996) observed that forensic hypnosis had become one of the fastest-growing areas of specialization for hypnosis professionals who were being retained by attorneys to use hypnosis to assist witnesses and victims of crime, or parties in civil cases, to recall vital information relevant to court proceedings.

Court rulings involving the admissibility of hypnotically refreshed recollection have affected how therapists use hypnosis with patients (Schefflin, 1993). How licensed mental health professionals use hypnosis in therapy is the fourth area of intersection: *hypnosis and the legal standard of mental health care* (Schefflin, 1997a). For the past two decades, lawyers for patients who have sued their therapists have attempted to persuade courts, legislatures, and insurance carriers that hypnosis is a potentially dangerous experimental treatment modality. The basic aspect of this argument is that patients should be provided with information identifying hypnosis as “experimental and dangerous” before consenting to such treatment. The more severe argument is that hypnosis should never be used in therapy because of the danger of creating false memories (Lynn, 2001).

The fifth area where the law takes cognizance of hypnosis is the use of hypnotic techniques for *courtroom advocacy* (Schefflin, 1998). May lawyers who have some training in hypnosis use the skills they learned to influence judges, jurors, and witnesses?

Forensic applications of hypnosis deal with issues of central importance to the law: memory, free will, choice, voluntariness, and responsibility. In deciding cases, courts have rarely found it necessary to define hypnosis, which is itself a formidable task, especially considering the diversity of opinion in the scientific community (Lynn & Rhue, 1991). The American Society of Clinical Hypnosis (ASCH) Task Force (Hammond et al., 1995) considers hypnosis to be a congruence of three components: dissociation, absorption, and suggestibility. In addition, situational or contextual factors play an important part in determining whether a person is hypnotized.

Hypnosis is a complex alteration in consciousness that can be understood as attentive, receptive concentration characterized by parallel, or dissociated, awareness. This shift in concentration may result in intense absorbing perceptual and sensory experiences, similar to that of being so engrossed in a good novel, movie, or play that one temporarily suspends awareness of the surrounding circumstances. The interaction between focal attention and peripheral awareness is a constant theme in human consciousness, but with hypnosis there appears to be a relative diminution of peripheral awareness to facilitate the enhancement of focal concentration. Although at no time does peripheral awareness disappear entirely, its dimming allows for the relative suspension of critical judgment, which is often observed in highly hypnotizable individuals. People in a trance can experience profound sensory alterations, such as tingling, lightness, or heaviness in extremities and alterations in motor control (e.g., letting an arm float up in the air with a feeling that they cannot control it, although, in fact, they can). They also experience changes in temporal orientation, such as reliving the past as if it were the present, and dissociation (e.g., feeling a part of the body or a part of their awareness as being separate from the rest).

As the ASCH Task Force (Hammond et al., 1995) has pointed out:

The courts have leaned toward defining hypnosis in terms of its antecedents, i.e., whether or not a hypnotic induction ceremony was administered. . . . Unfortunately, this has downplayed the importance of defining hypnosis by its consequences (i.e., hypnotizability). . . . We believe that it should be demonstrated that both a hypnotic induction was administered, and that the subject was responsive to such a procedure (e.g., through the elicitation of phenomena either informally, or formally through the administration of a hypnotizability scale). (p. 3)

Some people cannot be hypnotized, a few are extremely hypnotizable, and the majority of the population has some moderate capacity to experience hypnosis. Court decisions on forensic hypnosis, however, almost never discuss the hypnotizability of the witness, and lawyers in these cases rarely have the witness assessed for responsiveness to hypnosis or suggestion. Each of the five areas where hypnosis and law intersect is examined in the pages that follow.

REGULATION OF HYPNOSIS

Hypnosis did not receive professional recognition until the 1950s, when the British Medical Association (1955), the American Medical Association (AMA; 1958), and the American Psychiatric Association (1961) officially approved hypnosis as a therapeutic modality. Today, statutes in 31 states include hypnosis or hypnotherapy within the definition of psychology or counseling, and two states list hypnosis within the definition of the practice of medicine. Only a few states, however, such as California, Connecticut, and Florida, have specific regulations concerning who may practice hypnosis.

Two appellate judicial decisions have dealt with regulating the conduct of lay hypnotists who provide treatment to their clients. In *Masters v. State* (1960), an appellate court upheld the conviction of a nonphysician hypnotist who advertised his hypnosis services for a variety of ailments. An inspector employed by the Better Business Bureau posed as a client and was told during a hypnosis session that his problem stemmed from his hatred of his father. At trial, a physician testified that the use of hypnosis required knowledge of medicine and specialized training. He added that "he never used hypnosis until he had made a complete physical examination of the patient and that he considered such an absolute necessity and that it was not safe for anyone to use hypnosis in an effort to cure unless such person had a background of medicine because by the improper use of hypnosis a patient might be made worse off than he had been before and might resort to suicide" (p. 474), Masters did not testify on his own behalf. After the jury convicted him for the unauthorized practice of medicine, the appellate court upheld the conviction.

In *People v. Cantor* (1961), Cantor had been convicted on two counts of practicing medicine without a license. He advertised himself to be the director of the National Hypnosis Institute of Los Angeles, and he guaranteed results in cases involving weight loss. The appellate court, noting that the question “whether practicing hypnotism is practicing medicine” (p. 849) was one of first impression, upheld the conviction. According to the judges:

It is our considered opinion that . . . the practice of hypnotism as a curative measure or mode of procedure by one not licensed to practice medicine, amounts to the unlawful practice of medicine. . . . To the extent that [Cantor] employed or attempted to practice his hypnotic powers, he was practicing medicine within the meaning of [the statute]. (p. 850)

California has a unique piece of legislation concerning lay hypnotists. Business & Professions Code §2908 provides that lay practitioners who use hypnosis do not violate the unauthorized practice laws if they have received a referral from a person licensed to practice medicine, dentistry, or psychology or if they use hypnosis to aid a person’s “avocational or vocational self-improvement” rather than for treatment. Each of the exemptions from the general licensing requirement raises as yet unanswered questions because this law, enacted in 1967, has never been interpreted by the courts. The first exemption appears to permit a lay hypnotist to practice psychology if the client is referred by a duly licensed health-care professional. It is not clear why a licensed professional would refer a client to a lay hypnotist for treatment. Indeed, it might be malpractice to do so. Because many professional societies have regulations against working with or training lay practitioners, referring a patient to a lay practitioner would violate the ethical code of these associations.

The second exemption leaves the crucial phrase *avocational or vocational self-improvement* undefined. Does use of hypnosis for weight loss or smoking cessation fall within this language, or are these medical and/or psychological problems that only licensed health-care professionals can treat? There is no definitive ruling on this question by any California court, but it is well known that lay hypnotists routinely provide services for both types of problems.

Even when the law permits the lay practice of hypnosis under certain circumstances, serious unresolved questions arise concerning the appropriate standard of care (*Johnson v. Gerrish*, 1986), especially because bad outcomes have been reported when lay hypnotists treat patients (Haberman, 1987). Furthermore, clients of non-licensed hypnotists are not afforded the protection of confidentiality and privilege, and they have no recourse to a disciplinary mechanism whereby improper conduct can result in a suspension of a right to practice.

In addition to the legal issue of the unauthorized practice of psychology or medicine, lay hypnotists may face liability if harm occurs as a result of their intervention. Dangers with stage hypnosis have frequently been reported (Echterling & Emmerling, 1987; Eimer, 2012; Erickson, 1962; Finkelstein, 1989; Gruzelier, 2000;

Harding, 1978; Kleinhaus, Dreyfuss, Beran, Goldberg, & Azikri, 1979). In *Hohe v. San Diego Unified School District* (1990), a high school junior (Hohe) who was injured during a hypnotism show at the school brought an action to recover damages. She had been selected by the hypnotist as a volunteer to participate in his "Magic of the Mind Show." Hohe had seen the prior year's hypnotism show and told her father about a "stunt where a subject was suspended between two objects while another person stood on the subject's stomach." During the show, Hohe slid from her chair and fell to the floor several times. The appellate court held that a fact issue existed as to whether the wording of the release signed by Hohe and her father barred them from recovering for her personal injuries, thereby allowing the case to go to trial.

A more tragic scenario occurred in Florida, where a high school principal hypnotized 75 students, staff, and parents on school grounds. One student committed suicide the day after the session, and two other students died a few months after being hypnotized. According to Newcomb (2012), "George Kenney, 51, was known to hypnotize students to help them achieve better test scores and peak athletic performance, despite being warned by his superiors to discontinue the practice." He pleaded no contest to practicing hypnosis without a license and was sentenced to a 1-year term of probation.

Whether hypnosis was a cause of the deaths, or was merely coincidental to them, will be a matter of controversy should the issue reach the civil courts, because, until recently, the subject of hypnosis and death has not been a topic of scholarly concern (Ewin, 2008; Frischholz & Schefflin, 2009).

HYPNOSIS AND ANTISOCIAL CONDUCT

The alleged power of hypnosis to override a person's will has been the subject of many works of fiction and numerous films. The image of Svengali, with his absolute power to bend his victims to his will, still looms large in the public perception of hypnosis (DuMaurier, 1894). Judges have rarely been asked to consider whether hypnosis dilutes criminal or civil responsibility. Cases involving the antisocial aspects of hypnosis fall into two categories: crimes committed *on* hypnotized subjects, usually sexual seduction or undue influence for economic gain, and crimes committed *by* subjects claiming to be in trance.

CRIMES COMMITTED ON HYPNOTIZED SUBJECTS

The power of hypnosis to override a person's will was a source of great concern to the public and the courts in the late 19th century (Laurence & Perry, 1988). Since the mid-1800s, appellate and federal district courts in the United States have dealt with hypnotic seduction in fewer than 25 cases. Nevertheless, allegations of hypnotic seduction continue to appear against professional and lay hypnotists, but not in great numbers (Venn, 1988). These cases are difficult to track because many of them

are never appealed, or they involve disciplinary actions that might not be publicly reported. In general, courts have been receptive to the claim that criminal charges may be brought, and convictions upheld, against a defendant accused of hypnotic seduction (*McIlwain v. State*, 1981; *Mirowitz v. State*, 1969; *People v. Sorscher*, 1986; *State v. Donovan*, 1905).

CRIMES COMMITTED BY HYPNOTIZED SUBJECTS

In cases of crimes committed by hypnotized subjects, defendants charged with criminal conduct argue that they should not be found culpable, because they were acting involuntarily under hypnotic command. The claim that “the hypnotist made me do it” became quite prevalent in the latter part of the 19th century. Several sensational trials in Europe in the 1880s and 1890s about the possibility of using hypnosis to induce criminal conduct, and the widespread publicity given to these cases, had an immediate influence on defense lawyers. As noted by Brodie-Innes (1891):

Recently the public mind has been startled by accounts of strange new powers, with mysterious and unknown possibilities, and by alarming hints of crimes of an entirely new class, more obscure, more terrible, and more difficult of detection than any yet known to medical jurisprudence. (p. 51)

An article in a British legal journal (“Hypnotism in Criminal Defence,” 1894) observed that while insanity was a favored plea of criminal suspects, “today hypnotism is the fashionable defense” (p. 249).

Expert opinion on the ability of hypnotists to control their subjects was far from consistent. Some experts contended that the uncontrolled and unscrupulous use of hypnosis could threaten “the national defense and civil society” (Harris, 1985, p. 209). Other experts, as Harris (1985) points out, claimed that a highly hypnotizable subject

obeyed all the commands of the magnetizer and would execute acts upon waking without any conscious awareness or subsequent memory. [However] deep inside this human marionette a consciousness of “self” continued to subsist, so that a truly pure hypnotic subject would fail to realize commands that were repugnant to his or her inner nature. (p. 207).

The hypnosis defense faded from the courts shortly after the turn of the 20th century and rarely appears in cases today. Eight appellate cases discuss hypnosis as a criminal defense. In *People v. Worthington* (1894), the first opinion dealing with the topic, a woman convicted of murdering her lover claimed she committed the act under her husband’s hypnotic power. The California Supreme Court (1894), in rejecting her argument, said “there was no evidence [that the] defendant was subject to the disease, if it be such. Merely showing that she was told to kill the deceased and that she did it does not prove hypnotism, or, at least, does not tend

to establish a defense to a charge of murder” (p. 172). In *Denis v. Commonwealth* (1926), a Roman Catholic priest who was convicted of forgery and uttering false financial instruments claimed that he was under the influence of a frail woman who was, in fact, insane. However, the priest admitted that he could not say that he was hypnotized, only that “I was under her influence to the extent that I did a lot of foolish things; whether that is hypnosis, I cannot tell” (p. 576). Not surprisingly, his conviction was affirmed. In *People v. Marsh* (1959), after Marsh escaped from prison and was recaptured, he claimed he had left prison under the influence of a posthypnotic suggestion given to him by his friend, an amateur hypnotist, who told him, while inducing a trance, to “go back where he . . . was having a good time” (p. 285). Marsh said he took the words literally and, at the first opportunity, escaped and went home. The court-appointed psychiatrist did not believe the story, the jury did not believe the story, and the court of appeals found no reason to reverse the conviction.

Other cases also strain one’s sense of credulity. Indeed, these cases appear to utilize hypnosis as a defense of last resort. In *People v. Baldi* (1974), the defendant’s counsel claimed his client committed murder in a self-induced trance brought about by his fixation on pictures of women with a “prominent bust” (p. 122). *Barfield v. State* (1974) involved a claim that the male victim had instructed the female defendant to kill him after he hypnotized her using “Vishanti” as a trigger word. The legitimacy of her defense was put into question when she simultaneously argued that she did not commit the crime and could prove she was 20 miles away and that she did commit the crime, but only because she was under the hypnotic influence of the deceased. In *United States v. Phillips* (1981), the trial judge commented that the case “was one of the most spellbinding that this writer has ever seen enacted on the forensic stage” (p. 758). The defendant claimed that she was a hypnotic slave of her Svengali-like husband. Interestingly, her husband bragged about his influence over his wife and testified that he had been hypnotizing her 10 or 15 times a day ever since she was 15 years old. He told her he was “her mother and her father, and her Lord and God” (p. 760), that he implanted memories of his having held her immediately after she was born and of his saving her from drowning when she was 9 years old. Despite the intrigue of the case, there was no direct ruling by the court on the hypnosis issue. Other cases raising the defense (*Tyrone v. State*, 1915; *United States v. McCollum*, 1984) discuss hypnosis only briefly.

Indeed, no reported case has upheld the validity of the hypnosis defense. A defendant who seeks to convince a court that he or she committed a crime while in a trance, or under a posthypnotic suggestion, will face an uphill struggle (Bonnema, 1993).

Despite the absence of a successful plea of hypnotic coercion in the courts, many states recognize that actions under hypnosis are not voluntary (Dressler, 2012; *People v. Dunigan*, 1981; *Rogers v. State*, 2003). Montana Crimes Code §45–2–101(33)(c) defines an “involuntary act” to include “conduct during hypnosis or resulting from hypnotic suggestion.” The influential Model Penal Code §2.01(2)(c) lists “conduct

during hypnosis or resulting from hypnotic suggestion" as involuntary (American Law Institute, 1985).

The possibility that subjects may be hypnotized into committing antisocial acts has been much debated in the hypnosis scientific community (Deyoub, 1984; Perry, 1979; Watkins, 1972). The voluntariness of the hypnotized subject and the ability to resist hypnotic suggestions remain the subject of substantial disagreement (Vingoe, 1997). Those who consider hypnosis a special mental state assert that executive control actually is altered during trance. According to Hilgard (1977), "Hypnosis is a condition in which the normal functioning of the executive ego is temporarily modified so that executive control is divided between the hypnotist and the person being hypnotized" (p. 229). Sociocognitive theorists, by contrast, assert that executive control does not *actually* diminish; rather, the hypnotic subject comes to perceive the hypnotic situation *as if* there is a loss of voluntary control. According to Lynn and Rhue (1991), there have been at least three different schools of thought regarding hypnotic involuntariness:

The term "involuntary" can be defined in at least three different ways.... An action can be termed involuntary if it is beyond one's control, so that one cannot act otherwise even if one wishes to. Since the so-called "golden age" of hypnotism (the 1880s and 1890s), the view of the hypnotized subject as a passive automaton under the sway of a powerful hypnotist had faded in popularity....

A second meaning of the term "involuntary" can be that the suggested response occurs automatically, without effort or activity to make it occur, even if the subject is able to prevent it from occurring if he or she so desires....

A third sense in which a response can be classified as "involuntary" is that the subject simply has the *experience* of an action as occurring without direct volitional effort....

Whatever their theoretical persuasion, workers in the field are in agreement that the experience of involuntariness frequently accompanies hypnotic responses. (pp. 606–610)

That hypnotic crime could also include programmed assassins was raised in Richard Condon's novel *The Manchurian Candidate* (1959). Although a work of fiction, it finds support in documents detailing the Central Intelligence Agency (CIA) experiments with hypnosis (Marks, 1978; Schefflin & Opton, 1978). The CIA efforts to develop hypnotically programmed agents who could be induced to violate their moral codes without remorse and with amnesia for the circumstances predated Condon's novel, and there is some evidence in those documents that their efforts met with some success.

HYPNOSIS FOR MEMORY RECALL

Although some aspects of hypnosis can trace their history back to the most ancient of civilizations (Kroger, 1977), there is little evidence that hypnosis was used before the 1880s to refresh the recollection of victims, witnesses, or culprits in criminal

or civil cases. The first recorded use of hypnosis to solve a crime appeared in 1845 (Gravitz, 1983). A local clairvoyant in a mesmeric sleep identified a teenager as the person who had stolen money from a shopkeeper. When confronted with this accusation, the teenager confessed. The first recorded admission in court of hypnotically facilitated memory occurred in 1848 (Gravitz, 1995). A witness in a murder case was hypnotized by the victim's husband to assist her recollection. At trial, the defense called Amariah Bingham, a pioneer in American psychiatry, to offer expert testimony that the witness was a hysteric. Hysterical women, he said, often create stories that are false although they believe them to be true. The jury may have accepted this testimony, because the defendant was acquitted.

Medical professionals first began systematically examining the relationship between hypnosis and memory for forensic purposes in the closing two decades of the 19th century (Ellenberger, 1970). French and German hypnosis specialists were aware of potential problems with hypnotically refreshed recollection. Albert Moll (1889/1958) noted that *retroactive hallucinations*, his name for *false memories*, "are of great importance in law. They can be used to falsify testimony. People can be made to believe that they have witnessed certain scenes, or even crimes" (pp. 345–346).

The great French hypnosis pioneer, Hippolyte Bernheim (1891/1980), penned similar concerns: "I have shown how a false memory can cause *false testimony given in good faith*, and how examining magistrates can unwittingly cause false testimony by suggestion" (p. 92 [emphasis in original]). Bernheim provided a dramatic example of his point by suggesting to a subject in trance that he had been awakened in the middle of the night by a raucous neighbor's singing and coughing. After the hypnosis, the subject not only reported the implanted incident of the loud neighbor but also supplied details of the event not suggested by Bernheim, thereby adding confabulation to the false report.

Despite this recognition of the use of hypnosis to refresh recollection, and the potential dangers, there is little evidence that hypnosis was used regularly by law enforcement officials or others for memory refreshment of victims or witnesses of crimes. When police were tempted to use hypnosis, they sought to obtain confessions from criminal defendants. In 1893, Dutch police captured a vicious serial killer and sought to discover the buried bodies of some of his victims. A statute prohibited hypnotically refreshed recollection from being introduced into evidence, but the police were concerned only with obtaining information about the crimes. When word of the police plans to hypnotize the killer became public, a multinational outcry forced them to back down ("Hypnotism and the Law," 1893). Although several European countries in the late 1800s and early 1900s conducted major trials involving hypnosis, none of these trials appears to have concerned hypnosis and memory (Harris, 1989).

The first appellate case to consider hypnosis is *People v. Ebanks* (1897). The hypnosis issue, which was discussed in a single paragraph, involved a criminal defendant's attempt to have an expert who had hypnotized him testify that, while in trance, he had made a statement professing his innocence and that the expert

believed he was telling the truth. The trial judge refused to permit this testimony: "The law of the United States does not recognize hypnotism. It would be an illegal defense, and I cannot admit it" (p. 665). The California Supreme Court quoted this statement and held that the trial judge was correct in his ruling.

Although *Ebanks* has been cited by some testifying experts for the proposition that hypnotically refreshed recollection is inadmissible, the case cannot be used to support this view. Hypnosis was not the central concern in *Ebanks*. The issue involved was *not* the admissibility of hypnotically refreshed memory but rather whether an *expert* could express an opinion on *Ebanks's* innocence based on his denial of guilt while in trance. Courts continue to recognize that testimony by an expert as to the truthfulness of a witness invades the province of the jury to decide the credibility of witnesses.

Also, an expert may not testify about what a person said in trance if the purpose of that testimony, as it was in *Ebanks*, is to prove the truth of the statements (*People v. Smith*, 1983; *State v. Harris*, 1965). In addition, permitting an expert to testify about whether the defendant committed the crime, or what the defendant said in trance, would essentially allow the defendant to offer testimony while avoiding taking the witness stand and being cross-examined (*United States v. Mest*, 1986). As noted by the Ninth Circuit Court of Appeals in *United States v. McCollum* (1984): "The attempt to introduce [a tape recording of the defendant] essentially amounted to an effort to put the defendant's testimony directly before the jury without subjecting him to the cross-examination and impeachment that would have followed had he taken the witness stand" (p. 1423). *People v. Ebanks* simply did not involve memory at all.

JUDICIAL RULINGS ON HYPNOTICALLY REFRESHED RECOLLECTION

After hypnosis received official professional approval as a therapeutic procedure in the 1950s, police departments began to express a renewed interest in using it to help solve crimes. Lay hypnotists began training police officials in the 1950s (Arons, 1967), as did some licensed mental health professionals (Bryan, 1962). By the 1970s, police departments were using hypnosis with increasing fervor (Block, 1976; Diggett & Mulligan, 1982; Hibbard & Worring, 1981; Kuhns, 1981; Monaghan, 1980; Reiser, 1980; Schefflin & Shapiro, 1989).

Police use of hypnosis to solve crimes sent hundreds of cases into court, raising the legitimacy of this method of assisting memory. Not surprisingly, the hypnosis community, prosecutors, and defense lawyers began to take interest in this development. The modern era of judicial interest in hypnotically refreshed recollection began in 1968 (Schefflin & Frischholz, 1999).

Also in 1968, the Federal Bureau of Investigation (FBI) started using hypnosis with crime witnesses and victims. Under the protocol adopted for conducting such hypnotic interviews, permission must first be obtained from a Justice Department assistant attorney general, and only certain crimes fit the profile for utilizing hypnotic memory refreshing techniques. These cases include bank robberies where

force was used or a large amount of money was stolen, kidnapping, extortion, crimes of violence, and terrorism (Ault, 1979, 1980).

The FBI, and other federal government agencies, developed the Federal Model, which “involves a team approach that relies on a qualified mental health professional to manage the hypnosis, and a law enforcement investigator (hypnosis coordinator) who is prepared to provide instruction on what information may be sought and whether the subject’s recollections are purely for the development of investigative leads, or whether the subject might later be asked to testify in court” (Hibler & Schefflin, 2012, p. 36). Care is taken by the licensed professional to give top priority to the mental and emotional well-being of the person to be hypnotized. Informed consent is first obtained, and a prehypnosis record is made of what the witness or victim can freely recall at that time.

Despite judicial concerns about the reliability of hypnotically refreshed recollection, federal hypnosis specialists, following careful guidelines for hypnosis with memory, have enjoyed success in solving difficult crimes. A fascinating glimpse into the FBI’s team approach is presented by Wester and Hammond (2011), who provide summary reports of 10 cases in which FBI agents and licensed professional hypnosis experts were able to access significant additional memories leading to the capture of a serial rapist, a murderer, church bombers, bank robbers, hate crime perpetrators, and other violent offenders.

OPEN ADMISSIBILITY RULE

In 1968, the first appellate opinion on hypnosis used to facilitate recall was decided. In *Harding v. State* (1968), the court held that a person who had been hypnotized to remember the details of a crime could testify in court. The twin engines of truth—cross-examination and the use of expert testimony—were sufficient to test the credibility of the testimony given by the witness. Thus, the *Harding* court took the position that whether the witness’s memory may have been impaired by hypnosis or suggestion is a matter affecting *credibility*, not *admissibility*. This judicial viewpoint, that refreshing memory with hypnosis was no different from refreshing memory by any other method, is known as the *open admissibility* rule.

Harding v. State inaugurated a decade of court decisions that followed its reasoning. From 1968 to 1978, every appellate court in the United States that addressed the issue of hypnotically refreshed recollection adopted *Harding’s* open admissibility approach. During this decade, courts did not discuss any potential dangers attendant to the use of hypnosis, and they did not discuss the relevant scientific literature on this topic.

PER SE EXCLUSION RULE

As might be expected, judicial acceptance of hypnotically refreshed recollection invited its increased use. Police officers by the thousands received training in

hypnosis. To counter these developments, defense attorneys collaborated with hypnosis experts to turn the tide of cases involving hypnotically refreshed recollection. In 1978, the Ninth Circuit Court of Appeals in *United States v. Adams* sounded a warning that the use of hypnosis to refresh memory may contain special dangers. That warning led to the development of two alternative rules, each of which restricted the admission of hypnotically refreshed testimony. The most restrictive rule is the opposite of the *Harding* open admissibility approach. Known as the *per se* exclusion rule, it prohibits hypnotically refreshed testimony in all cases. In other words, a person will not be allowed to testify about any memory that first surfaces during or after a hypnosis session.

The Minnesota Supreme Court, in *State v. Mack* (1980), was the first court to prohibit the admissibility of hypnotically refreshed recollection into evidence. It was soon followed by the California Supreme Court in its highly influential decision in *People v. Shirley* (1982).

The California Supreme Court's opinion relied heavily on Diamond's (1980) belief that hypnosis inevitably causes memory hardening ("concreting") that results in a witness having an increased confidence in the veracity of the memories that occur after a hypnotic memory refreshing technique had been used. Even worse, this increase in confidence in the accuracy of a memory applies to all memories—including false ones. As a consequence of this hardening, previously hypnotized subjects cannot be cross-examined effectively about these false memories, thereby depriving a defendant in a criminal case of the constitutional right to confront adverse witnesses.

Mack and *Shirley* also raised concerns about other possible dangers when hypnosis triggers memory retrieval. These courts concluded that hypnotically refreshed memories inherently lack reliability and that the use of hypnosis creates undue suggestibility that would lead to confabulated testimony.

Interestingly, the *Mack* and *Shirley* courts both based their reasoning on *Frye v. United States* (1923), a case that set the test for the admissibility of *expert* testimony when that testimony was based on a *new* or *novel* scientific device, instrument, or procedure. Technically, *Frye* had no application to the testimony of *witnesses*, or to the use of hypnosis, which was hardly new or novel (D. Spiegel, 1987).

In *People v. Williams* (1982), Judge Gardner strongly objected to the reasoning of *Shirley*, which he described as "really more of a polemic than an opinion":

I am troubled by the concept that the testimony of a percipient witness as to relevant facts be deemed inadmissible simply because he has undergone hypnosis.

What next? Once we begin to rule evidence inadmissible because of our dissatisfaction with the witness' credibility based on improper memory jogging, where do we stop? What about witnesses who have been brainwashed, coached, coerced, bribed or intimidated? Are we going to reject all this testimony because it is suspect? I have no doubt that a corrupt polygraph operator could convince a witness of limited intelligence that his accurate memory is actually faulty and thus persuade him to testify to an untruth. The same is true with the so-called truth serums, hallucinogenic drugs

or other exotic drugs only hinted at in C.I.A. [Central Intelligence Agency] suspense fiction. I have no doubt that through the misuse of these drugs a witness' testimony may become faulty and even suspect. Once having undergone exposure to something of this nature is the witness still going to be allowed to give his best recollection, or be precluded from testifying?

I am firmly of the belief that jurors are quite capable of seeing through flaky testimony and pseudo-scientific clap-trap. I quite agree that we should not waste our valuable court time watching witch doctors, voo-doo practitioners or brujas go through the entrails of dead chickens in a fruitless search for the truth. However this is only because the practice is too time consuming and its probative value is zilch. I like the rule established in *Frye v. United States* . . . on the basis that it is a good pragmatic tool to keep out unnecessary, time consuming and nonproductive evidence. . . . However, the idea that an eyeball witness to a transaction be denied the opportunity to tell a jury his recollections of what he saw is disturbing to me whether that recollection has been refreshed by hypnosis, truth serum, drugs, intimidation, coercion, coaching, brainwashing or impaired by the plain old passage of time. (pp. 926–928)

The per se exclusion rule prohibits anything remembered during or after a hypnosis session from being admitted into evidence. The *Shirley* rule was even more severe. In its initial opinion, the *Shirley* court held that any witness or victim who had been hypnotized for forensic purposes would not be allowed to testify about *any* of the facts of the case. Once a person was hypnotized, that person was disqualified from testifying even about matters remembered and recorded *before* hypnosis was used. Thus, this initial ruling disqualified the *witness*, not just the *posthypnotic testimony*. However, the Supreme Court of California modified its *Shirley* ruling to indicate that memories that had been recorded before the hypnosis session would be admissible. Thus, a witness or victim who has provided a recorded prehypnosis statement is permitted to testify about memories revealed in this statement. However, the per se exclusion rule bars all testimony that is recalled during and/or after the hypnosis session, and it prevents the witness or victim from testifying about a posthypnotic identification of the defendant (McConkey & Sheehan, 1995).

The initial *Shirley* opinion was also modified with regard to defendants. The court ruled that a defendant in a criminal case would be permitted to testify about matters remembered during or after hypnosis. This modification later achieved constitutional status in *Rock v. Arkansas* (1987), in which the U.S. Supreme Court held that the Sixth Amendment prohibited a state from automatically excluding, by way of a per se inadmissibility rule, a criminal defendant's testimony.

Commentators who support the per se exclusion rule cite *Shirley* as an example of why the hypnotically refreshed recollection should be prohibited. They claim that hypnosis substantially altered the victim's testimony and was used successfully to get her to tell a coherent story. For example, Karlin and Orne (1996) have stated that after Catherine, the alleged victim in the *Shirley* case, was hypnotized, she stopped offering varying accounts of how she was sexually assaulted by the defendant.

Instead, her testimony “did not waiver from her final version, which she told in court with considerable certainty” (p. 57). The implication here is that the hypnosis shaped and solidified a now consistent, but completely false, story that was told with confidence and coherence. However, although it is true that Catherine did not tell a consistent and coherent story before hypnosis and probably was hypnotized by the prosecutor to get her to do so, the opposite of what Karlin and Orne report is true (Schefflin, 1997b). According to the California Supreme Court’s recitation in *People v. Shirley* (1982) of what actually occurred at trial:

The jury believed part of Catherine’s story, as it convicted the defendant of rape; but it also apparently found that she was lying when she described in detail the alleged act of oral copulation, as it acquitted the defendant of that charge. The jury doubtless had a difficult task, since Catherine’s performance as a witness was far from exemplary: the record is replete with instances in which her testimony was vague, changeable, self-contradictory, or prone to unexplained lapses of memory. Indeed, on occasion she professed to be unable to remember assertions that she had herself made on the witness stand only the previous day. (p. 245)

Thus, the hypnosis was ineffective in fabricating a coherent, consistent, and false story. Furthermore, the jury was not awed by the use of hypnosis—the jurors reached a reasoned approach that accepted some parts of the hypnotically refreshed testimony and rejected other parts of it, as juries do with testimony that was not hypnotically refreshed.

For many reasons, *Shirley* was a poor vehicle for articulating a per se exclusion rule. First, Catherine had been drinking heavily before the crime, and she also drank immediately afterward and took a sedative, which suggest that her memory was already impaired. Hypnosis is not effective under this condition. Second, the prosecutor, who was hardly neutral, performed the hypnosis and did so on the eve of Catherine’s testimony. Third, the hypnosis was not for the purpose of *solving* a crime but rather for *enhancing* Catherine’s credibility as a witness. It can be argued that the misuse of hypnosis in *Shirley* did not require a per se exclusion rule in *all* cases, any more than the fact that the police may unfairly question a witness in one case should lead to a rule that they may never question a witness in any case. It should be noted that none of the alleged dangers of using hypnosis to refresh memory was present in *Shirley*, because Catherine, before and after the hypnosis, had a slipper memory and lacked confidence in her recollections.

By the time *Shirley* was decided, police across the country were being trained by the thousands in hypnosis. The inadequacy of much of this training began to alarm the courts and became the subject of frequent media stories. The California Supreme Court was clearly influenced by the expanding use of police hypnotists, especially considering that the country’s leading training school was located in Los Angeles (Reiser, 1980). The rise of police hypnotists, coupled with the improper use

of hypnosis in the *Shirley* case, shaped the law far more than did the actual science of hypnosis used with memory.

GUIDELINES TEST AND THE “TOTALITY OF THE CIRCUMSTANCES” TEST

In between the *Mack* and *Shirley* cases, the New Jersey Supreme Court in 1981 rejected the wide-open *Harding* ruling and the completely closed per se rule of *Mack*. Instead, the justices in *State v. Hurd* (1981) permitted hypnotically refreshed recollection to be used in court with these guidelines:

First, a psychiatrist or psychologist experienced in the use of hypnosis must conduct the session. . . .

Second, the professional conducting the hypnotic session should be independent of and not regularly employed by the prosecutor, investigator or defense. . . .

Third, any information given to the hypnotist by law enforcement personnel or the defense prior to the hypnotic session must be recorded, either in writing or another suitable form. . . .

Fourth, before inducing hypnosis the hypnotist should obtain from the subject a detailed description of the facts as the subject remembers them. . . .

Fifth, all contacts between the hypnotist and the subject must be recorded. . . .

Sixth, only the hypnotist and the subject should be present during any phase of the hypnotic session, including the pre-hypnotic testing and the post-hypnotic interview. . . . (p. 545)

This “admissibility with guidelines” test required courts to hold a pretrial hearing to ascertain whether the requirements had been met, which in turn provided a foundation for believing that the hypnotically refreshed recollection was reliable enough to be admitted into evidence. At trial, experts and cross-examination could further support or challenge the memory’s reliability.

But what if the guidelines had not been scrupulously followed? Should courts strictly adhere to the guidelines, or should they use them more flexibly? Adopting the latter approach, most courts adopted a “totality of the circumstances” test, whereby the reliability of the hypnotically refreshed memory is evaluated pretrial by examining whether, under the totality of the circumstances surrounding the hypnosis sessions, the hypnotically refreshed testimony appears sufficiently free of undue suggestion or other taint so that its reliability should be tested in court rather than excluded in its entirety (*Borawick v. Shay*, 1995; Clemens, 1991).

The main difference between the per se exclusion rule and the totality of the circumstances test is the former’s rejection of *every* case in which hypnosis has been utilized, compared to the latter’s requirement that every case have a pretrial hearing to determine if the hypnosis sessions were likely to be unduly suggestive (Schefflin, 1994a,b).

In *State v. Armstrong* (1983), the Supreme Court of Wisconsin established a slightly more comprehensive set of guidelines for the admissibility of hypnotically refreshed recollection:

1. The person administering the hypnotic session ought to be a mental health person with special training in the use of hypnosis, preferably a psychiatrist or a psychologist.
2. This specially trained person should not be informed about the case verbally. Rather, such person should receive a written memorandum outlining whatever facts are necessary to know. Care should be exercised to avoid any communication that might influence the person's opinion.
3. Said specially trained person should be an independent professional not responsible to the prosecution, investigators or the defense.
4. All contact between the specially trained person and the subject should be videotaped from beginning to end.
5. Nobody representing the police or the prosecutor or the defendant should be in the same room with the specially trained person while he is working with the subject.
6. Prior to induction a mental health professional should examine the subject to exclude the possibility that the subject is physically or mentally ill and to confirm that the subject possesses sufficient judgment, intelligence, and reason to comprehend what is happening.
7. The specially trained person should elicit a detailed description of the facts as the subject believes them to be prior to the use of hypnosis.
8. The specially trained person should strive to avoid adding any new elements to the subject's description of her/his experience, including any implicit or explicit cues during the pre-session contact, the actual hypnosis and the post-session contact.
9. Consideration should be given to any other evidence tending to corroborate or challenge the information garnered during the trance or as a result of post-hypnotic suggestion. (p. 394 fn. 23)

HYPNOSIS WITH DEFENDANTS

The three rules just discussed deal with the admissibility of the hypnotically refreshed recollection of witnesses and victims of crimes. Each state has the authority to utilize whichever rule it prefers. However, when dealing with the person accused of committing the crime, *Rock v. Arkansas* (1987) held that the Constitution mandates that a state cannot automatically exclude the defendant's hypnotically refreshed testimony.

In several cases, hypnosis has been misused by the police in an effort to obtain information from a defendant about a crime. For example, in *Leyra v. Denno* (1954), hypnosis was used in an attempt to coerce a confession from Leyra, who was accused of killing his parents. After hours of intense interrogation shortly after the murders, a doctor offered to treat Leyra for a headache. The doctor covertly hypnotized Leyra and told him that he might as well confess to the murders.

The doctor assured Leyra that he would see to it that the police would “go easy” on him. Leyra confessed to the doctor and then to the police. He was convicted and sentenced to the electric chair, but this conviction was reversed on appeal, because the confession had been coerced by the hypnosis. At the next trial Leyra was again found guilty based on a second confession, but the U.S. Supreme Court reversed this decision. The majority opinion held that the second confession should be considered part of a continuum clearly related to the first confession and thereby similarly coerced. By the time a third trial began, the remaining evidence was largely circumstantial and inadequate. Despite this fact, Leyra was again convicted, but, because of the sparse evidence, the case yet again was reversed on appeal. Leyra was eventually set free by a judge because of the unconscionable police interrogation practices, which included a physician’s coercive misuse of hypnosis to elicit his confessions. The law is clear that hypnosis may be used by the police with criminal defendants, but only with their informed consent and only if proper guidelines have been followed to avoid undue influence or impermissible suggestion.

Apart from the police use of hypnosis with suspects, defendants may volunteer to be hypnotized to assist their own defense. In *State v. Papp* (1978; Orne, 1979), a defendant who reported amnesia for certain details of the crime underwent hypnosis, and his account suggested his innocence. Expert witnesses for the prosecution testified, however, that his behavior was typical of someone simulating rather than actually experiencing hypnosis. On the strength of this testimony, the hypnosis session was interpreted as self-serving and was not introduced in court. In another case, *People v. Ritchie* (1977; Orne, 1979), a defendant undergoing hypnosis implicated his wife rather than himself, but the court eventually decided to exclude the hypnotic evidence. Mutter (1984, 1990), however, reported cases in which hypnosis of the defendant produced exonerating statements that were later independently corroborated. In *Rock v. Arkansas* (1987), the U.S. Supreme Court dealt with exactly this situation.

May a per se exclusion rule be applied to criminal defendants to prohibit them from testifying after their memories were hypnotically refreshed? In *Rock v. Arkansas* (1987), a wife and her husband began an argument that eventually led to a physical struggle. A gun somehow was produced, and the husband was shot and killed. The wife, charged with his death, reported she could not remember the actual shooting, although she did remember that they were arguing. Limited amnesia (Schacter, 1986) and situation-specific amnesia (Gudjonsson, 1992) following traumatic events has been well documented (Loftus & Burns, 1982; Schefflin, 2004). Defense counsel sent his client to a mental health professional for hypnosis to determine whether any memories could be accessed. During the hypnosis session, the wife reported that she remembered that her finger was never on the trigger. As a result of her statement under hypnosis, defense counsel had the gun tested at a laboratory. The results showed that the gun was defective and could discharge even though the trigger had not been pulled. Thus, the wife’s exculpatory testimony was

at least partially corroborated. However, Arkansas followed a *per se* exclusion rule that prevented the wife from testifying about her hypnotically refreshed recollection. After being convicted, she appealed on the grounds that the *per se* exclusion violated her constitutional right to testify. The U.S. Supreme Court held that application of a *per se* rule of inadmissibility “does not extend to *per se* exclusions that may be reliable in an individual case. Wholesale inadmissibility of a defendant’s testimony is an arbitrary restriction on the right to testify in the absence of clear evidence by the State repudiating the validity of all posthypnosis recollections” (p. 61).

The *Rock* ruling was reaffirmed in *United States v. Scheffer* (1998), which upheld a *per se* exclusion rule where the defendant in a court-martial proceeding sought admission of polygraph results to support his testimony that he had not knowingly used drugs. The Supreme Court specifically distinguished *Rock* because, in that case, the exclusion of evidence “significantly undermined” fundamental elements of the defense. The defendant was unable to testify that the killing was accidental, because this memory had been refreshed by hypnosis. Thus, the *per se* rule “deprived the jury of the testimony of the only witness who was at the scene and had firsthand knowledge of the facts. . . . Moreover, the rule infringed upon the accused’s interest in testifying in her own defense—an interest that we deemed particularly significant, as it is the defendant who is the target of any criminal prosecution” (pp. 315–316). By contrast, in *Scheffer*, the exclusion of the polygraph results “did not implicate any significant interest of the accused” (pp. 316–317), nor did it significantly impair the defense’s presentation of its case. As noted by the Court, the members of the general court-martial “heard all the relevant details of the charged offense from the perspective of the accused, and the [per se] Rule did not preclude him from introducing any factual evidence. Rather, [defendant] was barred merely from introducing expert opinion testimony to bolster his own credibility” (p. 317). Thus, a *per se* exclusion rule is unconstitutional if it undermines the ability of an accused to present a defense (*Paxton v. Ward*, 1999).

In *Newman v. Hopkins* (2001), a woman who had been sexually assaulted provided the police with a description of her assailant, whom she claimed spoke with a Hispanic accent. The defendant wanted permission to read a neutral statement aloud in court to demonstrate to the jury that he did not speak with an accent. However, under Nebraska’s evidentiary rulings, such voice exemplars were *per se* inadmissible, because they were unreliable for the reason that an accent could be easily manipulated and the circumstances under which the victim heard her attacker’s voice could not be replicated. The federal court overruled the defendant’s conviction and held that a *per se* exclusion of such evidence was unreasonable, because “the reliability determination was based not on an individualized inquiry into the facts and circumstances of [defendant’s] proposed voice exemplar but rather on characteristics common to all voice exemplars” (p. 852). The categorical ban on voice exemplar evidence “prevented [the defendant] from offering factual

evidence and significantly undermined his ability to establish the essential elements of his defense” (p. 853). According to the court:

We recognize full well that the state of Nebraska has a legitimate interest in the reliability of evidence, and if the facts and circumstances surrounding a particular voice exemplar make it so unreliable as to render it inadmissible under Nebraska’s evidentiary rules, a defendant would have no absolute right to introduce it. . . . Rather than balancing the state’s concern with the defendant’s Sixth Amendment rights, however, Nebraska’s *per se* rule bars not only unreliable evidence but also evidence that may, in individual cases, be reliable. This the state may not constitutionally do, for “[a] State’s legitimate interest in barring unreliable evidence does not extend to *per se* exclusions that may be reliable in an individual case.” *Rock*, 483 U.S. at 61, 107 S.Ct. 2704. . . . The decisions of the Supreme Court clearly establish that such a *per se* rule is unconstitutional. *Rock*, 483 U.S. at 61, 107 S.Ct. 2704. (p. 853)

Thus, even after *Rock*, trial judges may still rule that a defendant’s hypnotically refreshed recollection is inadmissible if such testimony is deemed to be unreliable (*State v. Butterworth*, 1990; *State v. L.K.*, 1990; *Tumlinson v. State*, 1988), provided there has been a pretrial hearing on this issue.

Orne (1982, cited in Perry & Laurence, 1990) has pointed out that “the risk to the legal system that a defendant’s memory be distorted by hypnosis in his favor is probably disproportionately small. Thus, judges and juries expect defendants’ statements to be self-serving and designed to present him in the best possible light” (p. 267). Consequently, he supported the *Rock* opinion and argued in favor of a “double standard” whereby defendants are permitted to have access to hypnosis for their defense but the hypnotically refreshed recollection of witnesses and victims is excluded (Orne, Dinges, & Orne, 1990). Perry and Laurence (1990) have expressed concern that permitting defendants to testify about their hypnotically refreshed memories “may ultimately reverse the trend in American courts to proscribe the admission of testimony derived from hypnosis” (p. 281). However, in the past two decades, courts have not been so inclined.

If the *per se* exclusion rule is to be altered in states that follow it, it will be for reasons other than that defendants may testify about their hypnotically refreshed recollection. For example, in *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993), the U.S. Supreme Court rejected the *Frye* rule and changed the test for the admissibility of scientific evidence in all federal courts. According to *Daubert*, the admissibility of expert testimony in all cases is determined by the trial judge, who must use a flexibly applied four-factor test to evaluate the reliability of an expert’s opinion about a scientific theory or technique:

1. Whether the theory or technique has been tested or can be tested (whether it is verifiable or falsifiable), or is otherwise derived by a scientific method.
2. Whether the theory or technique has been subjected to peer review and/or publication.

3. Whether the theory or technique has a known or potential error rate.
4. Whether the theory or technique is generally accepted within the relevant scientific community.

The *Daubert* test has now been adopted by most state courts. Because this test is built on flexibility, several courts have held that per se exclusion rules are no longer defensible. In two cases involving polygraphs, federal courts have held that evidence obtained from a lie detector cannot automatically be excluded (*United States v. Pettigrew*, 1996; *United States v. Posado*, 1995). In another polygraph case (*United States v. Cordoba*, 1996), the court held that its per se exclusion rule against the admission of polygraph evidence was “effectively overruled” by the “flexible inquiry assigned to the trial judge by *Daubert*” (p. 227). The court further noted that other per se rules were equally as vulnerable to abolition and had already been overthrown.

More directly on point, in *Rowland v. Commonwealth* (1995), a stepmother witnessed her stepson shoot her and her daughter in the back. The stepmother’s physician diagnosed her as having posttraumatic stress disorder and recommended she be treated by Dr. William Wester, a psychologist. Wester agreed with the diagnosis and decided to treat her with hypnosis. Before beginning the hypnosis treatment, Dr. Wester, a former president of the ASCH with extensive forensic hypnosis experience, took complete statements from the stepmother about the shooting incident. The first statement was audiotaped, and the second statement was videotaped. Following the videotaping, Wester used hypnosis for the first time. The stepmother’s statement while in trance was virtually identical to her recorded prehypnotic statements. After the defendant was convicted, on appeal he argued for a rule of per se inadmissibility. In a 4-to-3 decision, the Supreme Court of Kentucky held that a per se inadmissibility rule was no longer appropriate and might violate *Daubert*.

SUPPORT FOR THE PER SE EXCLUSION RULE

No commentator has defended the application of the open admissibility rule, which is still followed in three states (North Dakota, Oregon, and Wyoming). The hypnosis literature debates whether the per se exclusion rule or the totality of the circumstances rule is preferable. Faigman, Kaye, Saks, and Sanders (2002), following the lead of other commentators (e.g., Giannelli, 1995; Giannelli & Inwinkelried, 1999; Karlin & Orne, 1996; Laurence & Perry, 1988), repeated the common assertion that the “majority of courts employ a per se rule of inadmissibility for hypnotically refreshed testimony” (p. 272). However, the claim that most courts, or most jurisdictions, follow the per se exclusion rule is not accurate. As of 2012, 25 states have adopted the per se exclusion rule; 14 states and 11 federal courts of appeals have adopted the totality of the circumstances test; three states have adopted an open admissibility rule; and eight states plus the District of Columbia plus one federal court of appeals have no definitive court rulings (see the appendix to this

chapter, which lists the hypnosis admissibility rules adopted in each state and federal jurisdiction). Thus, the per se exclusion rule applies in only 25 of the 63 jurisdictions in the American legal system.

Proponents of the per se exclusion rule (Karlin & Orne, 1997; Laurence & Perry, 1988) point out that it is economical, because it saves court time and judicial resources, and they argue that hypnosis *inevitably* contaminates memory, thereby constituting, in essence, a form of tampering with evidence. According to Schefflin and Frischholz (1999, pp. 93–94), courts, based on the claims of experts who support automatic exclusion of hypnotically refreshed recollection, have identified eight dangers associated with hypnotically refreshed recollection:

A. Suggestibility

1. The subject becomes “suggestible” and may try to please the hypnotist with answers the subject thinks will be met with approval.
2. The subject is highly responsive for the creation (“implantation”) of pseudomemories.

B. Reliability

3. The subject is likely to “confabulate,” that is, to fill in details from the imagination, in order to make an answer more coherent and complete.
4. The subject experiences “memory hardening,” which gives him or her great confidence in both true and false memories, making effective cross-examination more difficult.
5. The subject has source amnesia that prevents properly identifying whether a memory occurred before or during hypnosis, or whether the memory is real or suggested.
6. The subject experiences a loss of critical judgment.

C. Believability

7. Juries will disproportionately believe testimony that is the product of hypnosis.
8. The subject can easily feign hypnosis and can be deceptive in trance.

Proponents of the per se exclusion rule offered two additional arguments against the use of hypnosis with memory. First, Perry (1995) claimed that, even when hypnosis is not being utilized directly, its contaminating effects occur with techniques that are actually “disguised” hypnosis. Disguised techniques, which are prevalent in stage hypnosis shows, have been learned and used by “recovered memory” therapists who have added their “New Age ideology, which argues that insight into the cause of symptoms leads to their alleviation” (p. 196). For Perry, “disguised” hypnosis is any use of guided imagery, relaxation, imagination, or visualization. Perry described a student who discussed with her mother whether she should participate in a hypnosis experiment. On the morning of the experiment, the mother said that maybe the daughter should not participate because she might never come out of the trance. Perry claimed this statement acted as a prehypnotic

suggestion, and the student had great difficulty coming out of the trance. By contrast, H. Spiegel (1997) classified the mother's behavior as a "noecebo," a negative message that inhibits healing. Thus, according to Perry's position concerning "disguised" hypnosis, even therapists or police who do not intentionally use "hypnosis" as a specialized technique, and who may have no training in hypnotic procedures, are nevertheless using hypnosis, and any memories that are recalled should be inadmissible.

Another objection to hypnosis is that hypnotic consequences may affect people who are not responsive to hypnotic suggestion. Orne and his colleagues (Orne, Whitehouse, Dinges, & Orne, 1996; Orne, Whitehouse, Orne, & Dinges, 1996), based on retrospective analyses of earlier research, argued that low- and medium-hypnotizable individuals are vulnerable to contamination from the inherently corrupting influence of hypnosis. Thus, even those who are not affected by hypnosis are affected by hypnosis. Brown, Schefflin, and Hammond (1998) rejected this position based on methodological flaws in the Orne research design. No court has yet dealt with these two additional objections to the admissibility of hypnotically refreshed recollection.

The Supreme Court of New Jersey, in *State v. Fertig* (1996), indicated that it might be inclined to revisit its ruling in *Hurd* in light of the number of state cases that had rejected a guidelines approach in favor of a total ban on testimony that had been hypnotically refreshed. That opportunity came in *State v. Moore* (2006), in which the court ruled:

Based on the record . . . , and the substantial body of case law that has considered the question since *Hurd* was decided, we have determined that a change in course is now warranted. We are no longer of the view that the *Hurd* guidelines can serve as an effective control for the harmful effects of hypnosis on the truth-seeking function that lies at the heart of our system of justice. Most important, we are not convinced that it is possible to know whether post-hypnotic testimony can ever be as reliable as testimony that is based on ordinary recall, even recognizing the myriad of problems associated with ordinary recall. We therefore conclude that the hypnotically refreshed testimony of a witness in a criminal trial is generally inadmissible and that *Hurd* should no longer be followed in New Jersey. (p. 1213)

Hurd was a sensible and fair rule, but it did require pretrial hearings to determine whether the guidelines had been met. The court noted that *Hurd* was more costly to administer and cited that reason in support of its reversal. The opinion may very well be influential in turning the tide against the admission of hypnotically refreshed recollection, but it can be faulted for its selective citation of the small number of scientific articles supportive of its perspective. Indeed, the two major analyses of the scientific studies representing the consensus of the hypnosis community at that time were not cited at all (Brown et al., 1998; Hammond et al., 1995). A dissenting judge found another basis to be unhappy with the court's ruling: The majority

opinion disenfranchises victims of crimes of their constitutional rights. According to the dissent:

In my view, a defendant's constitutional right to testify on his own behalf and a victim's constitutional right to testify against the one who stands accused of harming that victim cannot be of unequal constitutional dignity. If, then, the Constitution allows the barring of hypnotically refreshed testimony but requires that an exception be made for a defendant's hypnotically refreshed testimony, there is no principled basis on which to treat a victim's hypnotically refreshed testimony any differently than that of the defendant. In other words, if we abandon the principles of *Hurd* and conclude that, in general, hypnotically refreshed testimony is forbidden in New Jersey because of its demonstrated unreliability, then the same logic that compels an exception for a defendant's hypnotically refreshed testimony likewise compels an exception for a victim's hypnotically refreshed testimony. (*State v. Moore*, 2006, p. 1231)

SUPPORT FOR THE TOTALITY OF THE CIRCUMSTANCES RULE

Proponents of the totality of the circumstances test reject the view that hypnotically refreshed testimony should never be admitted in court, for four reasons: (1) objections to hypnosis are based on logical errors; (2) a per se exclusion rule is unfair in general and to real victims of crime in particular; (3) a per se exclusion rule is not practical; and (4) the relevant science fails to support the claim that hypnosis inevitably contaminates memory or that hypnotically refreshed recollection is necessarily unreliable.

Logical Errors. Proponents of the totality of the circumstances test argue that experts who support the per se exclusion rule commit four major logical errors. First, they confuse the *use* of hypnosis with the *misuse* of hypnosis. The fact that a police officer may use undue suggestion during a particular interrogation does not mean that all police interrogations are improperly suggestive and should be prohibited. The same argument applies to the use of hypnosis. When the hypnotist follows strict guidelines to prevent undue suggestion or influence, memory is not contaminated. When hypnosis is used improperly (Coons, 1988; Schefflin, 2012; *State v. Zimmerman*, 2003), the reliability of resulting memories is subject to question, just as they should be with any other memory retrieval method not properly applied. Lynn, Neuschatz, Fite, and Kirsch (2000) correctly point out that "it would be wrong to scapegoat hypnosis while ignoring or minimizing the potentially misleading and hazardous effects of a variety of non-hypnotic memory enhancement techniques (e.g., leading questions, reinforcement for recall)" (p. 120).

Second, experts supporting the per se exclusion rule mistakenly attribute to hypnosis phenomena that are really aspects of memory. Thus, confabulation, memory hardening, and postevent misinformation are all attributes of *memory*; they are not created solely by hypnosis, and they occur without the use of hypnosis. Memory

research has shown that confabulation is a natural way in which memory works rather than a by-product of hypnotic trance (Loftus, 1980). Experiments with eyewitness testimony have conclusively demonstrated confabulation in nonhypnotic settings and have also demonstrated that hypnotically refreshed recollection is not necessarily confabulated (Brown et al., 1998; Hammond et al., 1995; Loftus, 1975, 1979a, 1979b, 1979c). Scientific studies demonstrate that the inherent memory problems of confabulation and postevent misinformation effects are not enhanced by hypnosis if appropriate guidelines have been followed (Hammond et al., 1995). Furthermore, memory hardening without hypnosis may be achieved by repetition and rehearsal, as trial lawyers demonstrate on a daily basis.

Third, proponents of the *per se* exclusion rule mistakenly assume that phenomena that may be more prevalent in persons who are highly hypnotizable are equally plausible in those who are moderately to low hypnotizable (Diamond, 1980). Most of the scientific studies reporting memory distortion with hypnosis have involved highly hypnotizable subjects, the population most vulnerable to memory distortion. The problems of confabulation and an artificial sense of confidence are especially applicable to the small subgroup of the population who measure as highly hypnotizable, whether hypnosis formally has been used or not (D. Spiegel & Spiegel, 1984). An intense, structured, and leading police interrogation, or pretrial preparation for testimony by an attorney, can have a more powerful adverse effect than any formal hypnotic ceremony would ever have in producing a false confession (Connery, 1977; Gudjonsson, 1992) or a false memory (Gudjonsson, 2003). Some investigators have concluded that high hypnotizability may be a factor equal to or more important than the formal use of the hypnotic ceremony (Zelig & Beidleman, 1981). Barnier and McConkey (1992) showed 30 high- and 30 low-hypnotizable subjects slides of a purse snatching. The subjects then asked to imagine seeing the slides in hypnosis; other subjects were asked to imagine seeing the slides in a waking condition (which in fact they had done). The experimenter suggested that the offender had a mustache (true), wore a scarf (false), and picked up flowers (false). Memory was tested by the experimenter after the suggestion, by another experimenter during an inquiry session, and again by the second experimenter after he or she appeared to have ended the session. More high- than low-hypnotizable subjects reported false memories. Barnier and McConkey concluded that the trait of hypnotizability, not hypnosis itself, was associated with false memory reports.

Other studies have argued that both high hypnotizability and a formal induction of hypnosis are necessary to produce an alteration in the recall of information (Dywan & Bowers, 1983). Courts generally have overlooked the significance of individual differences in evaluating hypnotically refreshed memory. Use of the standardized hypnotizability scales (e.g., Stanford Hypnotic Susceptibility Scales, Weitzenhoffer & Hilgard, 1959; Stanford Hypnotic Clinical Scale, Hilgard & Hilgard, 1975; Harvard Group Scale, Perry, Nadon & Button, 1992; Hypnotic Induction Profile, H. Spiegel & Spiegel, 1978; Barber Creative Imagination Scale, Barber & Wilson, 1978–1979) to document the subject's degree of hypnotic responsiveness should be an essential

part of the forensic use of hypnosis. Indeed, if a subject fails to demonstrate any hypnotic responsiveness on formal testing, the person conducting the session would be well advised to forgo any further hypnotic induction ceremonies as the subject is unlikely to respond, and the problems inherent with the appearance of having induced hypnosis can be avoided.

Fourth, experts who seek to ban hypnotically refreshed recollection from courtrooms overgeneralize the dangers of hypnotically created false memories. The fact that memory *can* be distorted with hypnosis does not mean that it *will* be so distorted. The claim that hypnosis always contaminates memory is disproven by the many cases in which hypnosis has not altered the memory of a witness or victim.

FAIRNESS

Schefflin (1997b) has offered the next true case as an example of how a total exclusion of hypnotically refreshed recollection is unfair to victims:

A four-year-old girl went to her mother and said "Daddy's touching me in my private parts." The mother had a breakdown and was hospitalized. The child, now in the custody of daddy, learned not to talk about this; look what happened to mommy when she was told. Several years pass, and the molestings continued. Medical records of the child were consistent with molesting, but the child would not talk when asked. After a year of therapy, hypnosis was used, and the child talked about the molesting. New York courts would not admit her posthypnotic testimony despite the fact that there was independent medical corroborating evidence that she was molested. Without her evidence, there was no proof that daddy was the molester. Daddy retained custody. (p. 269)

It is a matter of fundamental fairness that each case be heard on its own merits, at least at a preliminary hearing where the quality of the evidence can be judicially assessed. The *per se* exclusion rule prohibits posthypnotic memories from being admitted into evidence, even when it can be shown that the hypnosis procedural guidelines were scrupulously followed, that there was no undue suggestion by the hypnotist, and that the memories can be independently corroborated as true.

Although the *per se* exclusion rule has been defended on the grounds that it saves money because it avoids court hearings, trading judicial economy for a lesser form of justice is a poor bargain. Furthermore, because most trials have preliminary hearings anyway, the cost savings, if any, would not be substantial.

Every area of the law dealing with undue influence or suggestion, including the assessment of police lineups and interrogations, uses a totality of the circumstances test—except the issue of the admissibility of hypnotically refreshed recollection. As Schefflin (1997b) has noted, use of a *per se* exclusion rule in cases involving forensic hypnosis is hard to defend in light of how courts treat other aspects of memory:

According to the . . . *per se* exclusion rule for hypnotized witnesses, a person who has been lobotomized can testify in court, a person who has received massive electroshock

treatments can testify in court, a person who has taken enormous dosages of mind-altering psychiatric drugs or psychedelics can testify in court, a person who has suffered substantial organic brain damage can testify in court; but a person who had been competently hypnotized by an experienced licensed professional who carefully followed strict guidelines to avoid undue suggestions, cannot testify in court. (p. 207)

Although a per se rule will eliminate any false memories from evidence, it will also eliminate 100% of the true memories.

PRACTICALITY

Application of the per se rule creates several complications and undesirable consequences. Schefflin (1994a, 1994b) has criticized the per se exclusion rule by raising several troublesome scenarios concerning its application.

Sexual Seduction. Under the per se rule, an unethical hypnotist who uses trance to facilitate seduction will have committed the perfect crime. The subject-victim will be unable to testify, because all his or her memories are posthypnotic. Even the supporters of the per se rule have acknowledged that an exception must be made for illegal or unethical conduct committed while the subject was in trance (Giannelli, 1995).

Time Delays. Suppose an individual is hypnotized and has no new memories. Five years later, with no intervening hypnosis, additional memories surface. Is the witness disqualified to testify in regard to them simply because of the hypnosis 5 years earlier? Is there any time limit after which the memories cannot be attributed to the hypnosis?

Self-Hypnosis. A person learns self-hypnosis and practices the technique regularly. Are all memories of this person now contaminated?

Audiotapes. The market is flooded with commercial "self-hypnosis/subliminal message" audiotapes. Does listening to such a tape disqualify a person from testifying?

Therapeutic Hypnosis. Suppose a patient arrives at therapy and the therapist decides hypnosis would be beneficial. The hypnosis is *not* conducted for the purpose of retrieving memories. During the trance, however, some memories are voluntarily revealed. The therapist immediately terminates the hypnosis session. Should the per se rule disqualify the witness from testifying about these memories and any later memories?

Nonhypnotizable Subjects. A person who is nonhypnotizable may be subjected to a hypnotic induction ceremony, but will not experience hypnosis. Opposing counsel, however, will move to block the person's testimony on the basis of the attempted hypnosis. The per se rule may thus bar testimony even though no hypnosis actually occurred (Orne, Whitehouse, Dinges, et al., 1996; Orne, Whitehouse, Orne, et al., 1996). For this reason, measurement of hypnotizability is extremely important. Because hypnotizability is a stable and measurable trait, and because people vary in their hypnotic capacity from no responsiveness to very high responsiveness, it is important to determine whether hypnosis has indeed occurred. The California Supreme Court accepted this argument in *People v. Caro* (1988), where an expert, Dr. David Spiegel, was able to show that the witness was not hypnotizable despite police efforts to induce a trance. The court ruled that the witness could testify, because the attempt to hypnotize had not succeeded. Similarly, a person who is highly hypnotizable may be susceptible to memory errors even if no hypnosis is used. For this reason, Beahrs (1988) has argued that hypnosis can never be excluded from the legal setting—even if everyone agreed that it was a good idea to do so.

Lynn et al. (2000), who concluded that “as a general rule, hypnosis should not be used to assist recall in forensic situations,” nevertheless suggested two other situations in which hypnosis is warranted: where “desperate” circumstances are involved, such as an ongoing kidnapping, or as a “last resort” when “other recall methods have tried and have failed to elicit useful material” (p. 119). A fully documented instance of the latter situation is reported by Raginsky (1969), who used hypnosis to restore a repressed memory of an airline pilot for the events surrounding a major airline crash. The pilot, in the previous 2 years, had undergone psychoanalytic interviews, directive interviews, intravenous pentothal, and psychological interviews, all of which were conducted by highly qualified professionals. In addition, the pilot had interviews with leading airline safety investigators, all of which were focused on the goal of memory recovery and all of which were unsuccessful in restoring details of his memory. A session involving the use of hypnosis for memory recall produced new information that led safety investigators to reexamine specific portions of the plane wreckage. When they did so, they found a defective part they had previously overlooked, thereby confirming the hypnotically retrieved repressed memory and relieving the guilt feelings the pilot had that he might have been the cause of the crash.

The fourth reason why hypnotically refreshed testimony should never be admitted in court is elaborated in the following discussion of studies of hypnotic memory enhancement.

STUDIES OF HYPNOTIC MEMORY ENHANCEMENT

Hypnosis has lent itself to mystification, even in forensic settings. The dramatic and compelling examples of previously amnesic material unearthed with hypnosis, especially in a traumatized witness or victim, led to hopes that hypnosis could

be used as a kind of truth serum and that the material elicited with it had some higher order of veracity than ordinary memories. Indeed, many police officers were taught that memory acts like a tape recorder and that hypnosis facilitated the replaying of the tape. It has been known for over a century, however, that memory is reconstructive (Münsterberg, 1908). Gardner (1932–1933) put it quite eloquently when he wrote:

What memory does not recall, the imagination tends to supply unconsciously as a rule, half-consciously where bias or suggestion exists, and consciously in whole-cloth perjury. As memory fades, imagination retouches the details; where this is done unconsciously, therefore honestly, we are apt to recall what we *think* should have normally occurred, or, if personally involved, what we *wish* had occurred, or what, from suggestions now half-forgotten, we *believe* occurred. . . . The merest skeleton of fact, repeatedly told, bodies forth as a complete, truthful narrative, “ere long fiction expels reality from memory and reigns in its stead alone” and “unconscious impressions” blend with “conscious realities”, playing havoc with objective truth. This “filling-in” of memory occurs so unconsciously that it does not even affect the positive belief or manner of the witness. . . . Memory is more than the re-instatement of the original perception; it involves the interpretation of details, judgment, estimates, and the correlation of related incidents. Imagination and suggestion are twin-artists ever ready to retouch the fading daguerrotype of memory. Just as “Nature abhors a vacuum”, the mind abhors an uncompleted picture, and paints in the details, careless indeed as to whether the old picture is reproduced faithfully. (pp. 400–401)

The question of what rule to adopt for the admission of hypnotically refreshed recollection is a policy issue, but the question of what impact hypnosis has on memory retrieval is a question of science. What does the science say about the accuracy of memories recollected with the assistance of hypnosis?

The experimental literature has attempted to answer some of the questions of the effectiveness of hypnosis on improving recall, but these studies are limited by the problem of ecological validity—the strained analogy between the laboratory and the forensic setting, especially when the topic to be studied is the effect of hypnosis on the memory of *traumatized* witnesses and victims. As an expert witness in *United States v. Hall* (1997) correctly noted, “People think of experiments as the be all and end all of science. . . . That is simply not the case. . . . I’m simply saying that . . . given the problems that arise in doing particularly social science research, we cannot do the things to people that we can even do to animals” (p. 1204). It is neither legal nor ethical to traumatize people for the purpose of conducting laboratory research, and it is not clear that nontraumatized subjects will behave the way traumatized subjects will behave. There is a difference between watching a carefully crafted film about a crime and being a witness to or a victim of an actual crime as it occurs. Thus, while the laboratory permits controlled studies to be conducted, whether the results from the laboratory generalize to real-life crime settings is an important question.

Even in experiments that attempt to replicate the kind of emotional arousal that may occur in a rape, an assault, a staged mock assassination (Timm, 1981), or a gory film (Putnam, 1979), such artificial settings cannot reproduce the sense of fear, pain, and helplessness that real victims and witnesses may experience during a crime. The intertwined roles of emotion and content in memory retrieval cannot be adequately replicated in a laboratory.

Another confounding effect involves the motivation involved in testimony (McConkey & Sheehan, 1995). It is far different for a college student to attempt to recall information as part of an experiment than for a witness or victim to provide information about a traumatic event that may lead to someone's incarceration. These motivational factors are crucial, especially when they affect the response criterion (i.e., the willingness of the subject to report something as a memory), and it cannot be assumed that factors that influence the response criterion in a laboratory experiment are the same as those that affect a witness's willingness to testify. Furthermore, there is always the problem of demand characteristics when evaluating laboratory experiments (Kihlstrom, 2002; Orne, 1959; Perry, 2002). Finally, none of the laboratory studies deals with amnesia for an event and the spontaneous recovery of memory by hypnosis.

However, clinical studies, though they have real-world validity, lack scientific rigor. Furthermore, as pointed out by McConkey and Sheehan (1995), in cases involving crimes, excessive motivation to remember details actually may hinder accurate recall. It is possible for a witness to come up with "remembered" material that is more responsive to internal needs (to be helpful in solving the crime) or external factors (to please the investigator) than it is to the truth. A number of reports illustrate either self-serving and feigned stories elicited under hypnosis (Orne, 1979; D. Spiegel & Spiegel, 1984) or an artificially induced experimental confabulation in a highly hypnotizable subject instructed to stick by an invented story (H. Spiegel, 1980). Thus, it is clear from the clinical literature that it is possible for hypnotized individuals to come up with compelling stories that are not necessarily true.

When faced with hypnotically refreshed recollections, courts have been concerned primarily that hypnosis inevitably produces confabulation, pseudomemories, and an enhanced confidence about the accuracy of the hypnotically retrieved memory, which makes cross-examination less effective. Results of experimental investigation fail to support this view, although the danger of confabulation, pseudomemories, and enhanced confidence is always present when any testimony is presented in court (*State v. Ture*, 1984).

In *State v. Mena* (1981), the court ruled, "It is generally agreed that hypnosis is a state of altered consciousness and heightened suggestibility in which the subject is prone to experience distortions of reality, false memories, fantasies and confabulation ('the filling in of memory gaps with false memories or inaccurate bits of information')" (p. 1276). These undesirable aspects of recall are artifacts of memory itself, not of hypnosis in particular. Hypnosis as a procedure does not contribute in a significant way to the production of pseudomemories, except

perhaps in a small proportion of persons who are highly hypnotizable responding to leading suggestions for peripheral details about a non-emotionally charged event. Barnier and McConkey (1992) found that hypnotizability, but not hypnosis, was associated with false memory reports. Scoboria, Kirsch, Mazzoni, and Milling (2002) pointed out that “per se exclusion of testimony from witnesses who have been hypnotized, but not of testimony from witnesses who have been asked leading questions, would be supported empirically only if the effects of hypnosis were more pernicious than those of misleading questions” (p. 32). Although they state that hypnosis and leading questions may have a negative effect on the accuracy of memory reports, they conclude that “the effects of prior exposure to misleading questions are more pernicious than those of prior exposure to hypnosis” and that “misleading questions produce significantly more errors than hypnosis” (p. 32). Scoboria, Mazzoni, and Kirsch (2006) revisited this study and obtained new results. Significantly, they report in the abstract to their paper:

In 2002, the first author and colleagues reported data indicating that both hypnosis and misleading questions decreased the accuracy of memory reports and decreased “don’t know” response rates, that the effects of misleading questions were significantly greater than those of hypnosis, and that the two effects were additive. . . . [T]he present study replicated the findings that misleading questions reduce accuracy and “don’t know” responding but failed to replicate the negative effect of hypnosis on memory reports. (p. 340)

These conclusions are consistent with those of Schefflin, Brown, Frischholz, and Caploe (2002), who explained that “the problem of memory distortion has much less to do with the use of the technique. . . per se and much more to do with the context, the pattern of expectancies, the quality of the interviewing, and the personality characteristics of the subject” (p. 415).

Thus, pseudomemory is largely a function of the interviewer’s questions and the characteristics of the subject, not of hypnosis (Goldstein, 2012). Leading questions may alter memory with or without hypnosis. A simple example illustrates this point. An interviewer asks a crime victim if her attacker looks like her neighbor. This question is unduly leading because it adds content by suggesting a possible answer. The question is equally as objectionable if the victim had first been hypnotized and then asked to respond. If, however, the interviewer says, “What do you remember about the person who attacked you?”, there is no improper suggestion and thus no possible contamination. Whether the subject had been hypnotized makes no difference.

CONFABULATION

Memory research has shown that confabulation is a function of many complex variables, including a personality trait of responsiveness to confabulation, the nature of the social context, and the strength of the memory. Confabulation is a natural way

in which memory works; it is not an inevitable by-product of hypnotic trance (Brown et al., 1998; Loftus, 1980). Experiments with eyewitness testimony have conclusively demonstrated confabulation in nonhypnotic settings and have also demonstrated that hypnotically refreshed recollection is not necessarily confabulated (Loftus, 1975, 1979a, 1979b, 1979c).

PSEUDOMEMORIES

The judicial belief that hypnosis itself produces pseudomemories is erroneous (Brown et al., 1998; McConkey, Barnier, & Sheehan, 1998). As noted by Lynn and Kirsch (1996):

False memories can be created with or without hypnosis, and the role of hypnosis in their creation is likely to be quite small. Similarly, the available data suggest that the trait of fantasy proneness is not likely to be of great importance. . . . Hypnosis does not reliably produce more false memories than are produced in a variety of nonhypnotic situations in which misleading information is conveyed to participants. (p. 151)

Similarly, Beahrs, Cannell, and Gutheil (1996) concluded that “false memories are more likely to arise from social influence, either inside or outside of hypnosis or psychotherapy; intrinsic suggestibility (especially interrogative) and dissociative potential; and less so, simply from being hypnotized” (p. 50). Lynn et al. (2000) accurately concluded:

The effects of leading or suggestive questions is [*sic*] probably much greater than the effects of the administration of a hypnotic induction which, in general, increases suggestibility to only a small degree. After all, nonhypnotic suggestive procedures can result in the production of very unlikely or false memories that equal or exceed those elicited by hypnosis. (p. 120)

ENHANCED CONFIDENCE (THE “CONCRETING/HARDENING” EFFECT)

Many courts have accepted the view expressed by the Minnesota Supreme Court that “effective cross-examination of a previously hypnotized witness is virtually impossible” (*State v. Ture*, 1984, p. 32). This has been labeled the “concreting” or “hardening” effect (Diamond, 1980). The only forensic studies on point, however, directly contradict the assertion that hypnotically enhanced pseudomemories are more resistant to cross-examination than are pseudomemories produced by skillful, suggestive interrogation techniques (Spanos, Gwynn, Comer, Baltruweit, & de Groh, 1989; Spanos, Quigley, Gwynn, Glatt, & Perlini, 1991). In *State v. Dreher* (1991), the court observed:

The defendant’s argument that, because of the hypnosis session, [the hypnotized subject’s] trial testimony was delivered with an aura of confidence which it would not otherwise have had is not persuasive. The memory-hardening process is an intrinsic

part of a witness's preparation for trial. While ordinarily it takes the form of numerous pre-trial interviews and interrogations by counsel, the result is the same as that which defendant claims occurred here: a witness who testifies with conviction and believability. The fact that the witness has been prepped to testify effectively does not disqualify his evidence so long as it has not been falsified. (pp. 220–221)

As noted by Lynn and Kirsch (1996), "The role of hypnosis in enhancing confidence in false memories is also exaggerated" (p. 152).

OPINIONS IN THE SCIENTIFIC COMMUNITY

The *legal* question of the admissibility of hypnotically refreshed recollections depends in part on the *scientific* question of whether hypnosis is a reliable means for facilitating accurate recall. Courts for decades have been influenced by whether a method, technique, or procedure has been generally accepted in the relevant scientific community. This was the crucial factor under the *Frye* test, and it is one of the four factors under the *Daubert* ruling. What do experts in the field of hypnosis believe about the science of hypnotically refreshed recollection and the admissibility in court of hypnotically retrieved memories?

REPORT OF THE AMERICAN MEDICAL ASSOCIATION

The Council on Scientific Affairs of the AMA convened an eight-member panel to prepare a report on the scientific status of refreshing recollection by the use of hypnosis. This report was approved by the House of Delegates (AMA, 1985). The panel concluded that there is no evidence that hypnosis enhances recall of meaningless or nonsense material (Barber & Calverley, 1966; Dhanens & Lundy, 1975; Rosenhan & London, 1963), and there is no enhancement of the *recognition* of meaningful material, such as a photo identification (Timm, 1981).

The more interesting, and more significant, area of research for the evaluation of investigative hypnosis is the study of enhancement of *recall* of *meaningful* material. Indeed, "the AMA report noted that the effectiveness of the use of hypnosis to uncover emotionally arousing memories had not been scientifically proven or disproven" (D. Spiegel, personal communication, November 16, 2004). Some early studies indicated greater recall of meaningful material under hypnosis, but the price paid is an increase in incorrect recall and an increased sense of confidence (Stebly & Bothwell, 1994) not justified by the ratio of incorrect to correct new material. However, these studies did not account for response bias. This is important, because it is clear that repeated trials, even without hypnosis, can result in an increase in the reporting of new correct and incorrect information (Erdelyi, 1970, 1996). Indeed, the proportion of correct–incorrect responses is similar in hypnosis and nonhypnosis recall conditions; there is simply more productivity in the hypnosis condition (Dywan & Bowers, 1983). However, the research strategy that has been used to control for productivity encourages subjects to guess; therefore, the significant

increase in inaccurate responses in research using that paradigm is likely to be, in part, an artifact of the research design.

The Dywan and Bowers study (1983) is particularly interesting, because it demonstrates that low-hypnotizable persons in the hypnotic condition perform no differently from high or low hypnotizables who are not hypnotized. Only the high hypnotizables in the hypnosis condition showed an increase in productivity and confidence. However, even with this group, no follow-ups were conducted to determine whether the reports were actually believed to be memories.

A study by Laurence and Perry (1983) relied on by the AMA panel stated that hypnotized individuals, told that they heard something while they were sleeping that in fact they did not hear, tended to report as real memories this hypnotically induced memory 7 days later. This study is identical to the experiment conducted by Bernheim (1891/1980) a century ago and by Orne in 1982 (Barnes, 1982). It is also reminiscent of the “honest liar” experiment reported by H. Spiegel (1980). However, even in the Laurence and Perry (1983) study, only a minority of the highly hypnotizable subjects were affected 7 days later. McCann and Sheehan (1988) criticized the study on methodological grounds, and the results have been overgeneralized in media accounts and in courtroom testimony by experts who report that hypnosis inevitably contaminates memory by causing implanted false memories. Similar studies of postevent misinformation that did not use hypnosis and did not use only highly suggestible subjects yield the same percentage of memory distortion. The number of these reports of memory distortion, however, may be the product of experimental design, because there is a substantially lower report rate when the research design is not a forced choice and when demand characteristics are controlled (Brown et al., 1998; Orne, 1962).

Frischholz (1996) presented an effective critique of the 1985 AMA report, noting that it used an outmoded definition of hypnosis, relied primarily on memory research using nonsense material in laboratory settings and on anecdotal reports in legal cases, and failed to recognize the distinction between memory and hypnotically assisted memory. Precisely the same contaminations some experts (Diamond, 1980) cautioned against—confabulation, undue self-confidence, increased responsiveness to suggestion and/or social influences, and demand characteristics—may be produced in memory without hypnosis, as memory researchers (Brown et al., 1998) have repeatedly demonstrated. Sheehan (1996) is correct in observing that “an important conclusion that has emerged from the literature is that memory contamination is a function of memory and influence and not a danger specific to the use of hypnosis” (p. 13).

In 1994, the AMA reaffirmed its 1985 report, but it did so without a single citation to the extensive post-1985 scientific literature (American Medical Association, 1995).

EXPERT OPINION

In 1995, the ASCH, after thoroughly examining the research literature and canvassing the opinions of approximately 80 experts on hypnosis, suggestion, and memory,

concluded that a per se inadmissibility rule was not scientifically warranted (Hammond et al., 1995). The Society for Clinical and Experimental Hypnosis gave this publication the Arthur Shapiro Award for the “Best Book of the Year on Hypnosis,” providing additional support that it accurately represents the scientific views of the majority of the hypnosis community. Brown et al. (1998) conducted a subsequent and more detailed examination of the scientific literature and also concluded that the per se exclusion rule was unnecessarily harsh. Their book received many awards, including the American Psychiatric Association’s prestigious Manfred S. Guttmacher Award.

Although a minority of experts continue to support the per se exclusion rule (Karlin, 1997; Karlin & Orne, 1996, 1997; Perry, Orne, London, & Orne, 1996), most forensic hypnosis authorities favor the totality of the circumstances test (Brown et al., 1998; Hammond et al., 1995; McConkey & Sheehan, 1995; Schefflin, 1996, 1997b; Schefflin & Shapiro, 1989). As Australian researcher Peter Sheehan (1996) has noted, the consensus “appears to be that the courts must decide on a case by case basis the admissibility of hypnotically recalled material” (p. 13). Sheehan and McConkey (1993) concluded, “It seems extreme to take the view that all hypnotically obtained information should be ignored” (p. 720); 2 years later they expressly rejected the per se exclusion rule in favor of the “admissibility with safeguards” approach (McConkey & Sheehan, 1995). Sadoff and Dubin (1990) reached the same conclusion: “The court must decide on a case by case basis about the admissibility of the hypnotic recall. We are opposed to the admissibility per se and the exclusion per se rules” (p. 121). Wagstaff (1996), one of Britain’s foremost forensic experts, concluded:

Instead of a blanket rejection of anything said by the witness in such cases, we must judge each case individually. Perhaps we might more usefully ask, what might be the effect of this *particular* hypnosis session, on *particular* statements, made by this *particular* witness? (p. 189)

Some supporters of a per se exclusion rule nevertheless recognize that hypnotically refreshed recollection may be accurate. According to Lynn et al. (2000), “Although we argue against the use of hypnosis to bolster recall, it does not necessarily follow that all hypnotically-elicited testimony is, by its very nature, inaccurate, and that hypnosis inevitably corrupts a person’s memory” (p. 120).

Significantly, four members (Gravitz, Mutter, D. Spiegel, and H. Spiegel) of the original eight-member committee that drafted the 1985 AMA report have concluded that their report should not be used by courts to support a per se exclusion rule (Hammond et al., 1995; D. Spiegel, personal communication, November 16, 2004).

HYPNOSIS AND THE LEGAL STANDARD OF MENTAL HEALTH CARE

Hypnosis currently is under siege in the therapeutic setting in three different ways. First, per se rules serve as a threat to therapists using hypnosis, because the therapist may be depriving the patient of the ability to testify in court (Schefflin, 1993).

Thus, the rules developed in forensic settings to govern the admissibility of testimony find application in therapy settings to govern the treatment provided to the patient. For example, suppose a patient during a hypnosis session reports memories of child abuse that occurred years earlier. The patient now wants to sue the alleged perpetrator. Even though the hypnosis was competently conducted for therapeutic purposes, and even though memory retrieval was not the purpose for the hypnosis, and even though leading and suggestive questions were not used, the patient in states following a *per se* rule might (but should not) be held disqualified from testifying about those memories. Such a patient may then decide to sue the therapist for disenfranchising him or her of the opportunity to bring a successful lawsuit against an assailant. Even if the court concludes that the patient can testify, the alleged perpetrator's attorney may use the fact that the memory surfaced after hypnosis to discredit the reliability of the patient's memory.

Schefflin and Shapiro (1989) first raised this issue and suggested the use of informed consent forms that protect the patient and the therapist (Schefflin, 1993). The guidelines for forensic and clinical hypnosis developed by the ASCH (Hammond et al., 1995) now strongly urge that therapists obtain the written informed consent of the patient and that the form specifically identify the risk if a legal action might be brought after the hypnosis is conducted. In this litigious climate, with lawsuits against therapists on the rise, failure to obtain this written and signed informed consent before using hypnosis in therapy increases the therapist's risk of malpractice liability. The difficulty of obtaining an informed consent to hypnosis in therapy was explored by Zeig (1985). The manner in which therapists utilize informed consent (the "event" model versus the "process" model) is discussed in Berg, Applebaum, Lidz, and Parker (2001).

Second, beginning in the early 1990s, hundreds of lawsuits challenged the validity of repressed memories (Taub, 1996). For more than a decade, lawyers suing therapists have questioned the legitimacy of using hypnosis to retrieve memories. In other cases, lawyers have defended alleged perpetrators by claiming that the memories of their accusers were tainted by the hypnosis used in the course of the therapy. In essence, it is argued that because hypnosis involves suggestion, hypnotic sessions must inevitably be *unduly* suggestive, resulting in memories that are not accurate but rather the product of such improper suggestion. A variant of this argument states that hypnosis is always an exercise in fantasy and imagination and therefore cannot result in historically accurate recollection.

Third, lawyers have extended the argument against hypnosis beyond the confines of its use to facilitate memory recall. In these cases, the attack is on hypnosis as a form of *treatment* (*Storm v. Legion Insurance Company*, 2003). These novel legal theories have their basis in *Daubert* (1993), where the U.S. Supreme Court instructed trial judges to utilize the previously discussed four factors in evaluating the admissibility of an expert's proposed testimony.

Although *Daubert* was intended to address only the admissibility of expert scientific testimony, lawyers have extended its four factors to challenge the use

of any treatment or therapeutic technique that fails to satisfy each factor. In most cases, proper informed consent may serve as an adequate defense (Beahrs & Gutheil, 2001). The argument is that any treatment or therapy that cannot be proven to be effective is an “experimental” or “dangerous” procedure. In most cases, even if this legal argument had merit, an appropriate informed consent form detailing the alleged nature and risk of the procedure would serve as an adequate defense to a malpractice claim.

Of course, therapists are well advised to follow current standards of care (Axelrad, Brown & Wain, 2009; Schefflin & Spiegel, 1998; Hammond et al., 1995). However, it is likely that the legal challenge to hypnosis eventually will claim that it should not be used at all, even when the patient has been fully informed. This drastic legal challenge should ultimately be resolved by the cumulation of evidence-based studies on the efficacy, or lack thereof, of hypnotic procedures. Ultimately, science should determine whether hypnosis should be used in therapy and, if so, for what conditions and under what circumstances.

HYPNOSIS AND ADVOCACY

The fifth area of intersection between law and hypnosis concerns *hypnosis and advocacy*. Should attorneys be permitted to use hypnotic techniques in the courtroom to persuade judges and juries? Little has been written on this intriguing subject (Schefflin, 1998).

The first formal relationship between hypnotists and lawyers seems to have occurred in the 1950s when Harry Arons (1967), a lay hypnotist based on the East Coast, began teaching hypnosis to physicians, psychologists, law enforcement officers, and attorneys. Meanwhile, on the West Coast, attorney Melvin Belli (1976) invited Dr. William J. Bryan Jr. to speak at the 1961 annual Belli Seminars being held that year in St. Louis. Bryan, the only American physician to limit his practice exclusively to hypnosis cases, explained how trial lawyers could benefit from hypnosis training.

Attorney F. Lee Bailey (1971) met Bryan at the Belli Seminar. Bryan hypnotized Bailey, and the next thing Bailey knew there was a hypodermic needle sticking through his hand. Bailey then enrolled in Bryan’s American Institute of Hypnosis as a student, graduating from the advanced class in 1964 (*Hypnosis Quarterly*, 1965).

Reference to the use of hypnosis or hypnotic technique makes an occasional appearance, usually unflattering, in court opinions. In *Mason v. Underwood* (1992), after the plaintiff was awarded a sum of money for her claim that the defendant had falsely imprisoned her, a concurring judge complained that the verdict was

so grossly excessive that it must have been given under some evil influence. Quite possible the plaintiff cast a spell over the jurors; quite possible her skillful counsel made some passionate, unfair, hypnotic, and prejudicial appeal to the jurors, asking them for smart money, while he posed as a disinterested minister of justice (p. 950).

In *State v. Rameau* (1996), after a jury acquitted the defendant on the charge of unlawful sexual contact, the trial judge offhandedly commented that defense counsel “had hypnotized the jury.” The possibility that a lawyer might actually possess and use hypnotic powers arose in the case of *Wilburn v. Reitman* (1939), where plaintiff’s counsel claimed that the attorney for the defense dominated the court by “will power and by the power of suggestion” (p. 34). Noting that this argument raised “a most serious question,” Chief Justice Ross responded:

It is perfectly ethical for an attorney to use logic, oratorical skill, persuasive power and magnetism to gain his point, but according to the accusations of counsel the opposing counsel employed hypnosis to win his case, that is, he put the judge in “a state resembling normal sleep, differing in being induced by the suggestions and operations of the hypnotizer, with whom the hypnotized subject remains in rapport, responsive to his suggestions.” (Webster’s New Int’l Dict.) It is true that the law is a social mechanism that grows and expands to keep pace with social progress and needs, but we feel that the practice of hypnotism, at least for the present, should not be permitted in a court of justice. (p. 35)

How realistic is it to believe that skills used by hypnotists would find no place in trials? Is it possible, or desirable, to keep the practice of hypnotism out of the courtroom? Bryan (1962) argued that hypnotic techniques can be used to influence jurors. After noting that this subject had not previously been discussed in hypnosis or legal publications because “it has been considered unthinkable that an ethical lawyer would employ hypnotic techniques in order to influence a jury” (pp. 215–216), Bryan made a distinction between “hypnotizing the jury and the use of hypnotic techniques in [an attorney’s] courtroom presentation” (p. 216). Whereas the former is clearly unthinkable, the latter is in fact inevitable. Bryan (1971) later described his techniques of jury selection, which he applied in several high-visibility trials handled by attorney F. Lee Bailey. If hypnosis is understood as an aspect of persuasive communication, techniques used by hypnotists are relevant to the courtroom tasks that lawyers perform. Thus, lawyers may learn from hypnotists how to:

1. assess the suggestibility of the audience (jurors, judges, witnesses);
2. attract and sustain attention;
3. concentrate the mind;
4. reduce peripheral awareness;
5. lower critical thinking to increase reception to suggestion;
6. use metaphors and sensory language;
7. give positive or negative suggestions;
8. build unconscious associations;
9. motivate posthypnotic action; and
10. facilitate selective amnesia.

Bryan (1962) also suggested that hypnosis might be used to relax a nervous witness before he or she takes the stand to testify:

There comes a time in the life of every trial lawyer when he has at his disposal a witness who knows the truth, understands the truth, but is simply too nervous and full of stage-fright to give an acceptable performance upon the witness stand. It is in such cases that a qualified medical hypnotist can be of great value to the attorney. (p. 193)

Although Bryan assumed the attorney would hire the services of a licensed hypnotist, he provided no reason why the lawyer with hypnosis training could not use relaxation inductions and techniques without professional assistance. However, there are problems with this use of hypnosis to relax witnesses, because judges might conclude that hypnosis used for relaxation unduly interferes with the cross-examination of the witness.

Related to the issue of lawyers utilizing hypnotic techniques in advocacy is the question of whether hypnotists should teach these skills to attorneys. Codes of ethics of hypnosis associations generally prohibit instructing or training laypersons in the use of hypnosis, with self-help techniques taught to patients constituting the major exception. Principle III-1 of the Code of Ethics of the Society for Clinical and Experimental Hypnosis (1993) states that a member "shall not give courses in hypnosis to lay people" (p. 2). Sections 4.1 and 4.2 of the International Society of Hypnosis Ethics Code (October 2002) prohibit teaching hypnotic techniques to persons not otherwise eligible for membership in its organization. There is no showing that attorneys with training in hypnosis are injuring the judicial process by misusing their skills, but the thought of lawyers manipulating juries with covert suggestions is somewhat intriguing. Lawyer ethics codes, not surprisingly, do not address this issue.

CONCLUSION

Hypnosis remains a subject of interest to the law, in part because it engenders much fascination and in part because it involves issues central to the resolution of legal disputes: memory, free will, choice, voluntariness, and responsibility. There is every reason to believe that legal cases involving various aspects of hypnosis will continue to receive judicial and legislative attention. Although the legal rules involving hypnotically refreshed recollection were largely formulated in the 1980s, most of the scientific literature concerning the use of hypnosis with memory has appeared since that time. The Minnesota Supreme Court, in *State v. Blanchard* (1982), wrote, "We are, of course, willing to consider future developments in this area" (p. 430). However, no modern court has adequately considered the recent scientific studies, a fact that may prompt lawyers to seek judicial review of evidentiary rules that were developed long before the science had spoken.

Hypnosis will continue to be used by mental health professionals, and some patients will complain that they have been harmed by it. Lawyers have been

carefully crafting new legal theories to challenge the use of hypnosis in therapy settings. These legal battles generally have pitted laboratory researchers in hypnosis acting as experts testifying against clinicians who use hypnosis in their practice. The court cases have as their backdrop the evolving standard, derived from the four-part *Daubert* factors, that only evidence-based therapies be utilized with patients (Barden, 2001; Gutheil, 2001). In particular, the legal debate involves how clearly informed consent forms should articulate the extent to which a hypnotic intervention can be empirically justified and the extent to which patients must be told that some treatments involving hypnosis may be, in the words of some lawyers and their experts, “dangerous and experimental” (Cannell, Hudson, & Pope, 2001; Davis, Loftus, & Follette, 2001; Lynn, 2001).

Questions concerning the misuse or abuse of hypnosis for antisocial purposes, which involves intentional violation of civil or criminal laws, will also receive judicial attention. The most frequent abuse of hypnosis for antisocial purposes involves seduction, but most of these cases are resolved in disciplinary proceedings rather than in courts. When such cases appear in civil trials, they are usually settled or not appealed, so it is difficult to know how often hypnosis is misused for prurient purposes. As it has in the past, forensic hypnosis will remain a subject of complication and fascination for jurists.

APPENDIX

State and Federal Hypnosis Rules

PER SE EXCLUSION RULE:

Alaska
Arizona
Arkansas
California
Connecticut
Florida
Georgia
Hawaii
Illinois
Indiana
Kansas
Maryland
Massachusetts
Michigan
Minnesota
Missouri

Nebraska
New Jersey
New York
North Carolina
Oklahoma
Pennsylvania
Utah
Virginia
Washington

OPEN ADMISSIBILITY RULE:

North Dakota
Oregon
Wyoming

GUIDELINES/TOTALITY OF THE CIRCUMSTANCES RULE:

Alabama
Colorado
Idaho
Iowa
Kentucky
Louisiana
Mississippi
New Mexico
Ohio
South Carolina
South Dakota
Tennessee
Texas
Wisconsin
First Circuit Federal Court of Appeals
Second Circuit Federal Court of Appeals
Third Circuit Federal Court of Appeals
Fourth Circuit Federal Court of Appeals
Fifth Circuit Federal Court of Appeals
Sixth Circuit Federal Court of Appeals
Seventh Circuit Federal Court of Appeals
Eighth Circuit Federal Court of Appeals
Ninth Circuit Federal Court of Appeals

10th Circuit Federal Court of Appeals
11th Circuit Federal Court of Appeals

NO DEFINITIVE COURT RULING:

Delaware
District of Columbia
Maine
Montana
Nevada
New Hampshire
Rhode Island
Vermont
West Virginia
District of Columbia Circuit Federal Court of Appeals

REFERENCES

- American Law Institute. (1985). *Model penal code and commentaries*. Philadelphia, PA: Author.
- American Medical Association. (1958). Council on Mental Health, Medical use of hypnosis. *Journal of the American Medical Association*, 168, 186–189.
- American Medical Association. (1985). Council on Scientific Affairs, Scientific status of refreshing recollection by the use of hypnosis. *Journal of the American Medical Association*, 253, 1918–1923.
- American Medical Association. (1995). Council on Scientific Affairs. Report on memories of childhood abuse. CSA Report 5-A-94. *International Journal of Clinical and Experimental Hypnosis*, 43, 114–117.
- American Psychiatric Association. (1961, February 15). Regarding hypnosis: Statement of position. APA Document Reference No. 610001. Washington, DC: Author.
- Arons, H. (1967). *Hypnosis in criminal investigation*. Springfield, IL: Charles C Thomas.
- Ault, R. L., Jr. (1979). FBI guidelines for the use of hypnosis. *International Journal of Clinical and Experimental Hypnosis*, 27, 449–451.
- Ault, R. L., Jr. (1980). Hypnosis—The FBI's team approach. *FBI Law Enforcement Bulletin*, 49(1), 5–8.
- Axelrad, A. D., Brown, D., & Wain, H. J. (2009). Hypnosis in forensic psychiatry. In B. J. Saddock, V. A. Saddock, & P. Ruiz (Eds.), *Kaplan and Saddock's comprehensive textbook of psychiatry* (9th ed., Vol. 2, pp. 2804–2831). Philadelphia, PA: Lippincott Williams & Wilkins.
- Bailey, F. L. (1971). *The defense never rests*. New York, NY: New American Library.
- Barber, T. X., & Calverley, D. S. (1966). Effects of recall of hypnotic induction, motivational suggestions, and suggested regression: A methodological and experimental analysis. *Journal of Abnormal Psychology*, 71, 169–180.
- Barber, T. X., & Wilson, S. C. (1978–1979). The Barber Suggestibility Scale and the Creative Imagination Scale: Experimental and clinical applications. *American Journal of Clinical Hypnosis*, 21, 84–96.

- Barden, R. C. (2001). Commentary: Informed consent in psychotherapy—A multidisciplinary perspective. *Journal of the American Academy of Psychiatry and the Law*, 29(2), 160–166.
- Barfield v. State*, 54 Ala. App. 15, 304 So.2d 257 (1974).
- Barnes, M. (Producer and Director). (1982). *Hypnosis on trial*. [Television documentary]. London, England: British Broadcasting Company.
- Barnier, A. J., & McConkey, K. M. (1992). Reports of real and false memories: The relevance of hypnosis, hypnotizability, and context of memory test. *Journal of Abnormal Psychology*, 101, 521–527.
- Behrs, J. O. (1988). Hypnosis cannot be fully nor reliably excluded from the courtroom. *American Journal of Clinical Hypnosis*, 31, 18–27.
- Behrs, J. O., & Gutheil, T. G. (2001). Informed consent in psychotherapy. *American Journal of Psychiatry*, 158(1), 4–10.
- Behrs, J. O., Cannell, J. J., & Gutheil, T. G. (1996). Delayed traumatic recall in adults: A synthesis with legal, clinical, and forensic recommendations. *Bulletin of the American Academy of Psychiatry and Law*, 24, 45–55.
- Belli, M. M. (1976). *My life on trial*. New York, NY: Popular Library.
- Berg, J. W., Applebaum, P. S., Lidz, C. W., & Parker, L. S. (2001). *Informed consent: Legal theory and clinical practice* (2nd ed.). Oxford, England: Oxford University Press.
- Bernheim, H. (1980). *New studies in hypnotism*. Trans. R. S. Sandor. New York, NY: International Universities Press. (Original work published in 1891).
- Block, E. B. (1976). *Hypnosis: A new tool in crime detection*. New York, NY: McKay.
- Bonnema, M. C. (1993). "Trance on trial": An exegesis of hypnotism and criminal responsibility. *Wayne Law Review*, 39, 1299–1334.
- Borawick v. Shay*, 68 F.3d 597 (2nd Cir. 1995), cert. denied, 517 U.S. 1229, 116 S. Ct. 1869, 134 L. Ed. 2d 966 (1996).
- British Medical Association. (1955, April 23). Supplementary annual report of council, 1954–5. *British Medical Journal*, (Supp., pp. 190–193).
- Brodie-Innes, J. W. (1891). Legal aspects of hypnotism. *Juridical Review*, 3, 51–61.
- Brown, D., Schefflin, A. W., & Hammond, D. C. (1998). *Memory, trauma treatment, and the law*. New York, NY: Norton.
- Bryan, W. J., Jr. (1962). *The legal aspects of hypnosis*. Springfield, IL: Charles C Thomas.
- Bryan, W. J., Jr. (1971). *The chosen ones: The psychology of jury selection*. New York, NY: Vantage Press.
- California Business & Professions Code*, Section 2908.
- Cannell, J., Hudson, J. I., & Pope, H. G., Jr. (2001). Standards for informed consent in recovered memory therapy. *Journal of the American Academy of Psychiatry and the Law*, 29(2), 138–147.
- Clemens, B. (1991). Hypnotically enhanced testimony: Has it lost its charm? *Southern Illinois University Law Journal*, 15, 289–320.
- Condon, R. (1959). *The Manchurian candidate*. New York, NY: McGraw-Hill.
- Connery, D. S. (1977). *Guilty until proven innocent*. New York, NY: Putnam.
- Coons, P. M. (1988). Misuse of forensic hypnosis: A hypnotically elicited false confession with the apparent creation of a multiple personality. *International Journal of Clinical and Experimental Hypnosis*, 36(1), 1–11.
- Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 113 S.Ct. 2786, 125 L. Ed.2d 469 (1993).

- Davis, D., Loftus, E. F., & Follette, W. C. (2001). Commentary: How, when, and whether to use informed consent for recovered memory therapy. *Journal of the American Academy of Psychiatry and the Law*, 29(2), 148–159.
- Denis v. Commonwealth*, 144 Va. 559, 131 S.E. 131 (1926).
- Deyoub, P. L. (1984). Hypnotic stimulation of antisocial behavior: A case report. *International Journal of Clinical & Experimental Hypnosis*, 33, 301–306.
- Dhanens, T. P., & Lundy, R. M. (1975). Hypnotic and waking suggestions and recall. *International Journal of Clinical and Experimental Hypnosis*, 23, 68–79.
- Diamond, B. L. (1980). Inherent problems in the use of pretrial hypnosis on a prospective witness. *California Law Review*, 68, 313–349.
- Diggett, C., & Mulligan, W. C. (1982). *Hypno-Cop*. Garden City, NY: Doubleday, Inc.
- Dressler, J. (2012). *Understanding criminal law* (6th ed.). New York, NY: LexisNexis.
- DuMaurier, G. (1894). *Trilby*. New York, NY: Harper & Row.
- Dywan, J., & Bowers, K. S. (1983). The use of hypnosis to enhance recall. *Science*, 222, 184–185.
- Echterling, L. G., & Emmerling, D. A. (1987). Impact of stage hypnosis. *American Journal of Clinical Hypnosis*, 29, 149–154.
- Eimer, B. N. (2012). Inadvertent adverse consequences of clinical and forensic hypnosis: Minimizing the risks. *American Journal of Clinical Hypnosis*, 55, 8–31.
- Ellenberger, H. F. (1970). *Discovery of the unconscious: The history and evolution of dynamic psychiatry*. New York, NY: Basic Books.
- Erdelyi, M. H. (1970). Recovery of unavailable perceptual input. *Cognitive Psychology*, 1, 99–113.
- Erdelyi, M. H. (1996). *The recovery of unconscious memories: Hypermnnesia and reminiscence*. Chicago, IL: University of Chicago Press.
- Erickson, M. H. (1962). Stage hypnosis back syndrome. *American Journal of Clinical Hypnosis*, 3, 141–142.
- Ewin, D. (2008). Hypnosis and death: Two remarkable cases. *American Journal of Clinical Hypnosis*, 51(2), 65–71.
- Faigman, D. L., Kaye, D. H., Saks, M. J., & Sanders, J. (2002). *Modern scientific evidence: The law and science of expert testimony*. St. Paul, MN: West.
- Finkelstein, S. (1989). Adverse effects after exposure to lay hypnosis in a group setting: A case report. *American Journal of Clinical Hypnosis*, 32(2), 107–109.
- Frischholz, E. J. (1996, November 11). *Latest developments in forensic hypnosis and memory*. Paper presented at the 47th Annual Workshops and Scientific Program of the Society for Clinical and Experimental Hypnosis, Tampa, FL.
- Frischholz, E. J., & Schefflin, A. W. (2009). A comment on an alleged association between hypnosis and death: Two remarkable cases. *American Journal of Clinical Hypnosis*, 52(1), 45–67.
- Frye v. United States*, 54 App. D.C. 46, 293 F. 1013 (1923).
- Gardner, D. S. (1932–1933). The perception and memory of witnesses. *Cornell Law Quarterly*, 18, 391–409.
- Giannelli, P. C. (1995). The admissibility of hypnotic evidence in U.S. courts. *International Journal of Clinical and Experimental Hypnosis*, 43, 212–233.
- Giannelli, P. C., & Inwinkelried, E. J. (1999). *Scientific evidence* (3rd ed.). Charlottesville, VA: Lexis Law.

- Goldstein, R. H. (2012). A case of false confession after an inadvertent hypnotic induction. *American Journal of Clinical Hypnosis*, 55, 56–67.
- Gravitz, M. A. (1983). An early case of hypnosis used in the investigation of a crime. *International Journal of Clinical and Experimental Hypnosis*, 31, 224–226.
- Gravitz, M. A. (1995). First admission (1846) of hypnotic testimony in court. *American Journal of Clinical Hypnosis*, 37(4), 326–330.
- Gruzelier, J. (2000). Unwanted effects of hypnosis: A review of the evidence and its implications. *Contemporary Hypnosis*, 17(4), 163–193.
- Gudjonsson, G. H. (1992). *The psychology of interrogations, confessions and testimony*. New York, NY: Wiley.
- Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions*. Hoboken, NJ: Wiley.
- Gutheil, T. G. (2001). Standards for informed consent in recovered memory therapy. *Journal of the American Academy of Psychiatry and the Law*, 29(4), 504.
- Haberman, M. A. (1987). Complications following hypnosis in a psychotic patient with sexual dysfunction treated by a lay hypnotist. *American Journal of Clinical Hypnosis*, 29(3), 166–170.
- Hammond, D.C., Garver, R.B., Mutter, C.B., Crasilneck, H.B., Frischholz, E., Gravitz, M.A., . . . Wester, W. (1995). *Clinical hypnosis and memory: Guidelines for clinicians and for forensic hypnosis*. Des Plaines, IL: American Society of Clinical Hypnosis Press.
- Harding, E. C. (1978). Complications arising from hypnosis for entertainment. In F. H. Frankel & H. S. Zamansky (Eds.), *Hypnosis at its bicentennial* (pp. 163–167). New York, NY: Plenum Press.
- Harding v. State, 5 Md. App. 230, 246 A.2d 302 (1968), *cert. denied*. Harding v. Maryland, 395 U.S. 949, 89 S. Ct. 2030, 23 L. Ed.2d 468 (1969), *overruled by State v. Collins*, 296 Md. 670, 464 A.2d 1028 (1983).
- Harris, R. (1985). Murder under hypnosis in the case of Gabrielle Bompard: Psychiatry in the courtroom in Belle Époque Paris. In W. F. Bynum, R. Porter, & M. Shepherd (Eds.), *The anatomy of madness: Vol. 2: Institutions and society* (pp. 197–241). London, England: Tavistock.
- Harris, R. (1989). *Murder and madness: Medicine, law, and society in the fin de siecle*. Oxford, England: Clarendon Press.
- Hibbard, W. S., & Worring, R. W. (1981). *Forensic hypnosis*. Springfield, IL: Charles C Thomas.
- Hibler, N. S., & Schefflin, A.W. (2012). Maximizing the usefulness of hypnosis in forensic investigative settings. *American Journal of Clinical hypnosis*, 55, 32–55.
- Hilgard, E. R. (1977). *Divided consciousness: Multiple controls in human thought and action*. New York, NY: Wiley.
- Hilgard, E. R., & Hilgard, J. R. (1975). *Hypnosis in the relief of pain*. Los Altos, CA: Kaufmann.
- Hohe v. San Diego Unified School District, 224 Cal.App.3d 1559, 274 Cal.Rptr. 647 (4th Dist. 1990).
- Hypnosis Quarterly*. [Photograph]. (1965). 10(1), 40.
- Hypnotism and the law. (1893, October 14). *Law Times*, 500.
- Hypnotism in criminal defence. (1894). *Barrister*, 249–251.
- International Society of Hypnosis (2002). *Ethics code*. Retrieved from <http://www.ish-hypnosis.org/ish-codeethics.htm>
- Johnson v. Gerrish, 518 A.2d 721 (Me. 1986).

- Karlin, R. A. (1997). Illusory safeguards: Legitimizing distortion in recall with guidelines for forensic hypnosis—Two case reports. *International Journal of Clinical & Experimental Hypnosis*, 45(1), 18–40.
- Karlin, R. A., & Orne, M. T. (1996). Commentary on *Borawick v. Shay*: Hypnosis, social influence, incestuous child abuse, and satanic ritual abuse: The iatrogenic creation of horrific memories for the remote past. *Cultic Studies Journal*, 13, 42–94.
- Karlin, R. A., & Orne, M. T. (1997). Hypnosis and the iatrogenic creation of memory: On the need for a per se exclusion of testimony based on hypnotically influenced recall. *Cultural Studies Journal*, 14, 172–206.
- Kihlstrom, J. F. (1985). Hypnosis. *Annual Review of Psychology*, 36, 385–418.
- Kihlstrom, J. F. (2002). Demand characteristics in the laboratory and the clinic: Conversations and collaborations with subjects and patients. *Prevention & Treatment*, 5, Article 36c. Retrieved from <http://socrates.berkeley.edu/~kihlstrm/PDFfiles/OrneCommentary.pdf>
- Kleinhauz, M., Dreyfuss, D. A., Beran, B., Goldberg, T., & Azikri, D. (1979). Some after-effects of stage hypnosis. *International Journal of Clinical & Experimental Hypnosis*, 27, 219–226.
- Kroger, W. (1977). *Clinical and Experimental Hypnosis* (2nd ed.). Philadelphia, PA: Lippincott.
- Kuhns, B. W. (1981). *Hypnosis and the law*. Glendale, CA: Westwood.
- Laurence, J.-R., & Perry, C. (1983). Hypnotically created memory among highly hypnotizable subjects. *Science*, 222, 523–524.
- Laurence, J.-R., & Perry, C. (1988). *Hypnosis, will, and memory: A psycho-legal history*. New York, NY: Guilford Press.
- Leyra v. Denno, 347 U.S. 556, 74 S.Ct. 716, 98 L. Ed. 948 (1954).
- Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, 7, 560–572.
- Loftus, E. F. (1979a). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.
- Loftus, E. F. (1979b). The malleability of memory. *American Scientist*, 67, 312–320.
- Loftus, E. F. (1979c). *The manipulative uses of language*. [Audiotape No. 20234]. Psychology Today Cassette. New York, NY: Ziff-Davis.
- Loftus, E. F. (1980). *Memory*. Reading, MA: Addison-Wesley.
- Loftus, E. F., & Burns, T.E. (1982). Mental shock can produce retrograde amnesia. *Memory & Cognition*, 10(4), 318–323.
- Lynn, S. J. (2001). Informed consent and uninformed clinical practice: Dissociation, hypnosis and false memories. *American Journal of Clinical Hypnosis*, 43(3–4), 311–321.
- Lynn, S. J., & Kirsch, I. (1996). Alleged alien abductions: False memories, hypnosis, and fantasy proneness. *Psychological Inquiry*, 7, 151–155.
- Lynn, S. J., Neuschatz, J., Fite, R., & Kirsch, I. (2000). Hypnosis in the forensic arena. *Journal of Forensic Psychology Practice*, 1(1), 113–122.
- Lynn, S. J., & Rhue, J. (Eds.). (1991). *Theories of hypnosis: Current models and perspectives*. New York, NY: Guilford Press.
- Marks, J. (1978). *The search for the Manchurian candidate*. New York, NY: New York Times Press.
- Mason v. Underwood, 49 N.D. 243, 191 N.W. 949 (1992).
- Masters v. State, 170 Tex. Crim. 471, 341 S.W.2d 938 (Crim. App. 1960).

- McCann, T., & Sheehan, P. W. (1988). Hypnotically induced pseudomemories—Sampling their conditions among hypnotizable subjects. *Journal of Personality & Social Psychology*, *54*, 339–346.
- McConkey, K. M., Barnier, A. J., & Sheehan, P. W. (1998). Hypnosis and pseudomemory: Understanding the findings and their implications. In S. J. Lynn & K. M. McConkey (Eds.), *Truth in memory* (pp. 227–259). New York, NY: Guilford Press.
- McConkey, K. M., & Sheehan, P. W. (1995). *Hypnosis, memory, and behavior in criminal investigation*. New York, NY: Guilford Press.
- McIlwain v. State, 402 So.2d 1194 (Fla. App. 5th Dist. 1981), *rehearing denied*, August 25, 1981, *petition for review denied* 412 So.2d 467 (1982).
- Mirowitz v. State, 449 S.W.2d 475 (Tex. Crim. App. 1969).
- Moll, A. (1958). *The study of hypnosis*. New York, NY: Julian Press. (Original work published 1889)
- Monaghan, J. (1980). *Hypnosis in criminal investigation*. Dubuque, IA: Kendall/Hunt.
- Montana Crimes Code, Section 45-2-101(3)(c).
- Münsterberg, H. (1908). *On the witness stand*. New York, NY: Doubleday.
- Mutter, C. B. (1984). The use of hypnosis with defendants. *American Journal of Clinical Hypnosis*, *27*, 42–51.
- Mutter, C. B. (1990). The use of hypnosis with defendants: Does it really work? *American Journal of Clinical Hypnosis*, *32*, 257–262.
- Newcomb, A. (2012, January 12). Hypnotist principal faces one year in jail, if guilty. Retrieved from <http://abcnews.go.com/US/florida-hypnotist-principal-charged/story?id=15346388>
- Newman v. Hopkins, 247 F.3d 848 (8th Cir. 2001).
- Orne, M. T. (1959). The nature of hypnosis: Artifact and essence. *Journal of Abnormal and Social Psychology*, *58*, 277–299.
- Orne, M. T. (1962). On the social psychology of the psychology experiment: With particular reference to demand characteristics and their manipulation. *American Psychologist*, *17*, 776–783.
- Orne, M. T. (1979). The use and misuse of hypnosis in court. *International Journal of Clinical and Experimental Hypnosis*, *27*, 311–341.
- Orne, M. T. (1982). Affidavit to *People v. Shirley*, April 28, 1982, quoted in C. Perry and J.-R. Laurence (1990). Hypnosis with a criminal defendant and a crime witness: Two recent cases. *International Journal of Clinical and Experimental Hypnosis*, *38*(4), 266–281.
- Orne, M. T., Dinges, D. F., & Orne, E. C. (1990). *Rock v. Arkansas*: Hypnosis, the defendant's privilege. *International Journal of Clinical and Experimental Hypnosis*, *38*(1), 250–265.
- Orne, M. T., Whitehouse, W. G., Dinges, D. F., & Orne, E. C. (1996). Memory liabilities associated with hypnosis: Does low hypnotizability confer immunity? *International Journal of Clinical and Experimental Hypnosis*, *44*, 354–367.
- Orne, M. T., Whitehouse, W. G., Orne, E. C., & Dinges, D. F. (1996). "Memories" of anomalous and traumatic autobiographical experiences: Validation and consolidation of fantasy through hypnosis. *Psychological Inquiry*, *7*(2), 168–172.
- Paxton v. Ward, 199 F. 3d 1197 (10th Cir. 1999).
- People v. Baldi, 80 Misc. 2d 118, 362 N.Y.S. 2d 927 (1974).
- People v. Cantor, 188 Cal. App. 2d Supp. 843, 18 Cal. Rptr. 363 (App. Dept., Super. Ct., L.A. County, 1961).

- People v. Caro, 46 Cal. 3d 1035, 251 Cal. Rptr. 757, 761 P. 2d 680 (1988).
- People v. Dunigan, 52 Ill. Dec. 247, 96 Ill. App. 3d 799, 421 N.E. 2d 1319 (1st Dist. 1981).
- People v. Ebanks, 117 Cal. 652, 49 P. 1048 (1897), *overruled on other grounds*.
- People v. Flannelly, 128 Cal. 83, 60 P. 670 (1900).
- People v. Marsh, 170 Cal. App. 2d 284, 338 P. 2d 495 (4th Dist. 1959).
- People v. Ritchie, No. C-36932. Super. Ct. Orange Co., California (April 7, 1977) (unreported).
- People v. Shirley, 31 Cal. 3d 18, 723 P.2d 1354, 181 Cal. Rptr. 243, *stay denied California v. Shirley*, 458 U.S. 1125, 103 S. Ct. 13, 73 L. Ed. 2d 1400, *cert. denied California v. Shirley*, 459 U.S. 860, 103 S. Ct. 133, 74 L. Ed. 2d 114 (1982).
- People v. Smith, 117 Misc. 2d 737, 459 N.Y.S. 2d 528 (1983).
- People v. Sorscher, 151 Mich. App. 122, 391 N.W. 2d 365 (1986).
- People v. Williams, 132 Cal. App. 3d 920, 183 Cal. Rptr. 498 (4th Dist. 1982).
- People v. Worthington, 105 Cal. 166, 38 P. 689 (1894).
- Perry, C. (1979). Hypnotic coercion and compliance to act: A review of evidence presented in a legal case. *International Journal of Clinical & Experimental Hypnosis*, 36, 187–218.
- Perry, C. (1995). The False Memory Syndrome (FMS) and “disguised” hypnosis. *Hypnos*, 22, 189–197.
- Perry, C. (2002). Hypnosis, demand characteristics, and “recovered memory” therapy. *Prevention & Treatment*, 5, Article 40. Retrieved from <http://journals.apa.org/prevention/volume5/pre0050040c.html>
- Perry, C., & Laurence, J.-R. (1990). Hypnosis with a criminal defendant and a crime witness: Two recent cases. *International Journal of Clinical and Experimental Hypnosis* 38(4), 266–281.
- Perry, C., Nadon, R., & Button, J. (1992). The measurement of hypnotic ability. In E. Fromm & M. R. Nash (Eds.), *Contemporary hypnosis research* (pp. 459–490). New York, NY: Guilford Press.
- Perry, C., Orne, M. T., London, R. W., & Orne, E. C. (1996). Rethinking per se exclusions of hypnotically elicited recall as legal testimony. *International Journal of Clinical and Experimental Hypnosis*, 44, 66–80.
- Putnam, W. H. (1979). Hypnosis and distortions in eyewitness memory. *International Journal of Clinical and Experimental Hypnosis*, 27, 437–448.
- Raginsky, B. B. (1969). Hypnotic recall of air crash cause. *International Journal of Clinical and Experimental Hypnosis*, 17, 1–19.
- Reiser, M. (1980). *Handbook of investigative hypnosis*. Los Angeles, CA: LEHI.
- Rock v. Arkansas, 483 U.S.44, 107 S. Ct. 2704, 97 L. Ed. 2d 37 (1987).
- Rogers v. State, 105 S.W.3d 630 (Crim. App. Tex. 2003).
- Rosenhan, D., & London, P. (1963). Hypnosis in the un hypnotizable: A study in rote learning. *Journal of Experimental Psychology*, 65, 30–34.
- Rowland v. Commonwealth, 901 S.W. 2d 871 (Ky. 1995).
- Sadoff, R., & Dubin, L. (1990). The use of hypnosis as a pretrial discovery tool in civil and criminal lawsuits. In C. H. Wecht (Ed.), *Legal medicine* (pp. 105–124). Dayton, OH: Butterworth-Heinemann.
- Schacter, D. L. (1986). Amnesia and crime: How much do we really know? *American Psychologist*, 41(3), 286–295.
- Scheflin, A. W. (1993). Avoiding malpractice liability. *American Society of Clinical Hypnosis Newsletter*, 34(1), 6.

- Schefflin, A. W. (1994a). Forensic hypnosis: Unanswered questions. *Australian Journal of Clinical and Experimental Hypnosis*, 22, 23–34.
- Schefflin, A. W. (1994b). Forensic hypnosis and the law: The current situation in the United States. In B. J. Evans and R. O. Stanley (Eds.), *Hypnosis and the law: Principles and practice* (pp. 25–48). Heidelberg, Victoria, Australia: Australian Society of Hypnosis.
- Schefflin, A. W. (1996). Commentary on *Borawick v. Shay*: The fate of hypnotically retrieved memories. *Cultic Studies Journal*, 13, 26–41.
- Schefflin, A. W. (1997a). Ethics and hypnosis: Unorthodox or innovative therapies and the legal standard of care. In W. Matthews & J. Edgette (Eds.), *Current thinking and research in brief therapy: Solutions, strategies, narratives* (Vol. 1, pp. 41–62). New York, NY: Brunner/Mazel.
- Schefflin, A. W. (1997b). False memory and Buridan's ass: A response to Karlin, R. A., and Orne, M. T. (1996). Hypnosis, social influence, incestuous child abuse, and satanic ritual abuse: The iatrogenic creation of horrific memories for the remote past. *Cultic Studies Journal*, 14, 207–289.
- Schefflin, A. W. (1998). Ethics and hypnosis: A preliminary inquiry into hypnotic advocacy. In W. Matthews & J. Edgette (Eds.), *Current thinking and research in brief therapy: Solutions, strategies, narratives* (Vol. 2, pp. 307–328). New York, NY: Brunner/Mazel.
- Schefflin, A. W. (2004, January 5). *Amnesia and homicide: Do killers forget their crimes?* Presented at the annual meeting of the Association of American Law Schools, Atlanta, GA.
- Schefflin, A. W. (2012). How not to conduct a forensic hypnosis interview: A case study. *American Journal of Clinical Hypnosis*, 55, 68–84.
- Schefflin, A. W., Brown, D., Frischholz, E. J., & Caploe, J. (2002). Special methodologies in memory retrieval. In R. I. Simon and D. W. Shuman (Eds.), *Retroactive assessment of mental states in litigation: Predicting the past* (pp. 369–423). Washington, DC: American Psychiatric Publishing.
- Schefflin, A. W., & Frischholz, E. J. (1999). Significant dates in the history of forensic hypnosis. *American Journal of Clinical Hypnosis*, 42(2), 84–107.
- Schefflin, A. W., & Opton, E., Jr. (1978). *The mind manipulators*. London, England: Paddington Press.
- Schefflin, A. W., & Shapiro, L. (1989). *Trance on trial*. New York, NY: Guilford Press.
- Schefflin, A. W., & Spiegel, D. (1998). From courtroom to couch: Working with repressed memory and avoiding lawsuits. *Psychiatric Clinics of North America: Diagnostic Dilemmas, Part II*, 21(4), 847–867.
- Scoboria, A., Kirsch, I., Mazzoni, G., & Milling, L. S. (2002). Immediate and persisting effects of misleading questions and hypnosis on memory reports. *Journal of Experimental Psychology, Applied*, 8(1), 26–32.
- Scoboria, A., Mazzoni, G., & Kirsch, I. (2006). Effects of misleading questions and hypnotic memory suggestion on memory reports: A signal-detection analysis. *International Journal of Clinical & Experimental Hypnosis*, 54(3), 340–59.
- Sheehan, P. W. (1996, August 17). *Contemporary trends in hypnosis research*. State-of-the-art address presented at the 26th International Congress of Psychology, Montreal, Canada.
- Sheehan, P. W., & McConkey, K.M. (1993). Forensic hypnosis: The application of ethical guidelines. In J. W. Rhue, S. J. Lynn, & I. Kirsch, *Handbook of clinical hypnosis* (pp. 719–738). Washington, DC: American Psychological Association.

- Society for Clinical and Experimental Hypnosis (1993). *Code of ethics*. <http://www.sceh.us/assets/Documents/sceh%20code%20of%20ethics.pdf>
- Spanos, N. P., Gwynn, M. I., Comer, S. L., Baltruweit, W. J., & de Groh, M. (1989). Are hypnotically-induced pseudomemories resistant to cross-examination? *Law and Human Behavior*, 13, 271–289.
- Spanos, N. P., Quigley, C. A., Gwynn, M. I., Glatt, R. L., & Perlini, A. H. (1991). Hypnotic interrogation, pretrial preparation, and witness testimony during direct and cross-examination. *Law and Human Behavior*, 15, 639–653.
- Spiegel, D. (1987). The Shirley decision: The cure is worse than the disease. In R. W. Rieber (Ed.), *Advances in forensic psychology and psychiatry* (Vol. 2, pp. 101–118). Norwood, NJ: Ablex.
- Spiegel, D., & Spiegel, H. (1984). Hypnosis in psychotherapy. In *The psychiatric therapies: Report of the American Psychiatric Association Committee on Psychiatric Therapies* (Part II, pp. 701–737), Washington, DC: American Psychiatric Association.
- Spiegel, H. (1980). Hypnosis and evidence: Help or hindrance? *Annals of New York Academy of Science*, 347, 73–85.
- Spiegel, H. (1997). Nocebo: The power of suggestibility. *Preventive Medicine*, 26(5), 616–621.
- Spiegel, H., & Spiegel, D. (1978). *Trance and treatment: Clinical uses of hypnosis*. New York, NY: Basic Books.
- State v. Armstrong, 110 Wis.2d 555, 329 N.W. 2d 386, cert. denied, Wisconsin v. Armstrong, 461 U.S. 946, 103 S. Ct. 2125, 77 L. Ed.2d 1304 (1983).
- State v. Blanchard, 315 N.W.2d 427 (Minn. 1982).
- State v. Butterworth, 246 Kan. 541, 792 P.2d 1049 (1990).
- State v. Donovan, 128 Iowa 44, 102 N.W. 791 (1905).
- State v. Dreher, 251 N.J. Super 300, 598 A.2d 216 (1991).
- State v. Fertig, 143 N.J. 115, 668 A.2d 1076 (1996).
- State v. Harris, 241 Or. 224, 405 P.2d 492 (1965).
- State v. Hurd, 86 N.J. 525, 432 A.2d 86 (1981).
- State v. L. K., 244 N.J. Super. 261, 582 A.2d 297 (1990).
- State v. Mack, 292 N.W.2d 764 (Minn. 1980).
- State v. Mena, 128 Ariz. 226, 624 P.2d 1274 (1981).
- State v. Moore, 188 N.J. 182, 902 A.2d 1212 (2006).
- State v. Papp (1978). No. 78–02–00229. C.P. Summit Co., Ohio; Lorain Co. No. 16862, March 23, 1978; unrep.; app'd U.S. Sup. Ct. No. 79–5091, cert. denied October 27, 1979.
- State v. Rameau, 685 A.2d 761 (Me. 1996).
- State v. Ture, 353 N.W. 2d 502 (Minn. 1984).
- State v. Zimmerman, 266 Wis.2d 1003, 669 N.W.2d 762 (2003).
- Stebly, N. M., & Bothwell, R.K. (1994). Evidence for hypnotically refreshed testimony: A view from the laboratory. *Human Behavior*, 18(6), 635–651.
- Storm v. Legion Insurance Company, 265 Wis. 2d 169, 665 N.W. 2d 353 (2003).
- Taub, S. (1996). The legal treatment of recovered memories of child sexual abuse. *Journal of Legal Medicine*, 17, 183–214.
- Timm, H. W. (1981). The effect of forensic hypnosis techniques on eyewitness recall and recognition. *Journal of Police Science and Administration*, 9, 188–194.
- Tumlinson v. State, 757 S.W.2d 440 (Tex. App. 1988).
- Tyrone v. State, 77 Tex. Crim. 493, 180 S.W. 125 (1915).

- United States v. Adams, 581 F.2d 193 (9th Cir. 1978), *cert. denied* 439 U.S. 1006, 99 S. Ct. 621, 58 L. Ed.2d 683 (1978).
- United States v. Cordoba, 104 F.3d 225 (9th Cir. 1996).
- United States v. Hall, 974 F. Supp. 1198 (C.D. Ill. 1997).
- United States v. McCollum, 732 F.2d 1419 (9th Cir.), *cert. denied* *McCollum v. United States*, 469 U.S. 920, 105 S.Ct. 301, 83 L. Ed.2d 236 (1984)
- United States v. Mest, 789 F.2d 1069 (4th Cir. 1986).
- United States v. Pettigrew, 77 F.3d 1500 (5th Cir. 1996).
- United States v. Phillips, 515 F. Supp. 758 (E.D. Ky. 1981).
- United States v. Posado, 57 F.3d 428 (5th Cir. 1995).
- United States v. Scheffer, 523 U.S. 303, 118 S.Ct. 1261, 140 L. Ed.2d 413 (1998).
- Venn, J. (1988). Misuse of hypnosis in sexual contexts: Two case reports. *International Journal of Clinical & Experimental Hypnosis*, 36(1), 12–18.
- Vingoe, F. J. (1997). A note on “experts” in forensic hypnosis. *Contemporary Hypnosis*, 14(1), 48–52.
- Wagstaff, G. F. (1996). Should “hypnotized” witnesses be banned from testifying in court? Hypnosis and the M50 murder case. *Contemporary Hypnosis*, 13, 186–190.
- Watkins, J. G. (1972). Antisocial behavior under hypnosis: Possible or impossible? *International Journal of Clinical & Experimental Hypnosis*, 20, 95–100.
- Weitzenhoffer, A. M., & Hilgard, E. R. (1959), *Stanford Hypnotic Susceptibility Scales: Forms A and B*. Palo Alto, CA: Consulting Psychologists Press.
- Wester, W. C. II, and Hammond, D. C. (2011). Solving crimes with hypnosis. *American Journal of Clinical Hypnosis*, 53(4), 255–269.
- Wilburn v. Reitman, 54 Ariz. 31, 91 P.2d 865 (1939).
- Zeig, J. K. (1985). Ethical issues in hypnosis: Informed consent and training standards. In J. K. Zeig (Ed.). *Ericksonian psychotherapy, Vol. 1: Structures* (pp. 459–473). New York, NY: Brunner/Mazel.
- Zelig, M., & Beidleman, W. B. (1981). The investigative use of hypnosis: A word of caution. *International Journal of Clinical and Experimental Hypnosis*, 24, 401–412.

PART FIVE

COMMUNICATING
EXPERT OPINIONS

Writing Forensic Reports

IRVING B. WEINER

EFFECTIVE consultation flows from effective communication. Applied psychologists have usually learned this lesson well from their training and professional experience. They know that their opinions and recommendations are valuable only when they are conveyed meaningfully in reports to those who have sought their services.

No matter how sharply practitioners have honed their communication skills as consultants in other contexts, they typically must learn some new ground rules when they undertake forensic consultation. As helping professionals, they have been accustomed to serving the needs of all parties involved in their cases, without knowingly or intentionally operating to anyone's disadvantage. The administration of civil and criminal justice marches to a different drummer, however, known as the *adversarial system*. As exemplified by customary courtroom proceedings, the adversarial system pits verbal combatants against each other to produce a winner and a loser. Whereas clinicians provide help in their cases, litigating attorneys win or lose when they take on a case. Litigants air opposing views before the bench, and the arguments that hold sway result in judgments that gratify some parties to a case and dismay others.

The adversarial system calls on attorneys to promote the interests of their clients while trying to prevent the clients of opposing attorneys from keeping or getting something they want to have, such as a sum of money or a conviction. The concerns of the judge rest not with who gets what or which party feels better or worse but with safeguarding due process and adherence to the rules of evidence. Impartial judicial oversight ensures a full and equal hearing before the bench—every person's day in court—whatever the outcome. The familiar statue of the blindfolded goddess of justice, allowing the scales to balance as they may, vividly portrays this feature of the judicial process (see Barrett & Morris, 1993; Bennett et al., 2006; Sales & Shuman, 2005).

Given the nature of the adversarial system, psychologists embarking on forensic work are commonly what Brodsky and Robey (1972) described many years ago as “courtroom-unfamiliar.” To provide effective consultation, newcomers to the field must become familiar with the adversarial system and comfortable with offering opinions that may contribute to severe penalties and crushing disappointments on the losing side of a case.

In addition, psychologists undertaking forensic consulting must learn to deal with impersonal clients. Forensic clients are not individuals seeking service directly on their own behalf. In some cases, they are an entity, such as a court seeking advice, a prosecutor’s office seeking a conviction, or a company seeking to defend itself against charges of negligence or malfeasance. In other cases, forensic clients are attorneys acting on behalf of a person or entity they are representing. Forensic psychologists usually have some direct contact with the plaintiff or defendant in a case on which they are consulting, especially for purposes of conducting a formal evaluation. However, this plaintiff or defendant will not ordinarily be the person with whom they make arrangements for the evaluation or discuss the nature and import of their findings. Instead, such matters are discussed with the attorney or other entity who constitutes the psychologist’s client.

The general implications of identifying impersonal clients accurately and working with them appropriately were first delineated by Monahan (1980) in a widely cited monograph titled *Who Is the Client?* Greenberg and Shuman (1997) and Heilbrun, Grisso, and Goldstein (2009), among others, subsequently have elaborated the distinctly different roles and responsibilities that characterize conducting forensic evaluations and providing psychotherapy.

This chapter addresses specific implications that working within the adversarial system and with impersonal clients have for the writing of forensic reports. Especially important in this regard are deciding whether a report should be written and, if so, how it should be focused. Along with presenting guidelines for making these decisions, the text that follows provides suggestions for writing forensic reports in a clear, relevant, informative, and defensible manner.

DECIDING WHETHER A REPORT SHOULD BE WRITTEN

Deciding whether to write a report may seem a frivolous consideration. Psychologists are accustomed to writing reports as a necessary and expected culmination of providing consultative services. In forensic work, however, the inevitability of a written report is tempered by rules concerning the nature of evidence. Expert opinions become evidence not when they are formulated in a consultant’s mind but only when they are stated orally under oath or written down, whether in formal reports that are voluntarily testified to or informal notes and electronic communications that can be subpoenaed during the discovery process. For this reason, attorneys typically advise persons involved in litigation to think whatever they wish but write down only what they are prepared to justify in testimony.

As an exception in this regard, communications to an attorney about the psychologist's arrangements for their work together may be considered part of the attorney's "work product" and therefore privileged, which means that these communications are protected against disclosure. A psychologist's findings and opinions in a case ordinarily remain privileged as well, so long as the psychologist is not identified to the court as an expert witness who might be called to give testimony. Once psychologists are included in a list of potential witnesses, however, they may be required to submit a written report for purposes of discovery, thereby allowing opposing counsel an opportunity to ferret out the likely content of their testimony.

When the option is available, deciding whether to write a report should be based on the preferences of the client and certain ethical principles and professional realities that shape the practice of law and psychology, including specific legal requirements that may obtain in a particular jurisdiction.

RESPECTING THE PREFERENCES OF THE CLIENT

Most commonly in forensic cases, the psychologist's client is the court or an attorney. Even when the fees are being paid by other agencies or private parties, judges and lawyers are the persons to whom forensic psychologists are most directly responsible and with whom they communicate most directly. These two main types of forensic clients typically differ in the kinds of information they seek. Judges request psychological evaluations to help them reach decisions, and they usually want to receive as much relevant information as possible and hear any opinions, whatever their implications, that might be useful to them. Hence, a written report typically is expected in court-ordered evaluations, and a thorough elaboration of the nature and significance of the psychologist's findings is welcome.

By contrast, attorneys trying a case are looking only for facts and opinions that will strengthen arguments on their client's behalf. If they consider the psychologist's conclusions damaging to their case, attorneys may prefer not to have a report written and may omit the psychologist from their witness list. The next three cases, each of a type common in the practice of forensic psychology, illustrate circumstances in which the attorney exercised these no-report options.

Case 1. Mr. A was a 33-year-old systems analyst who had undergone surgery in connection with an accurately diagnosed medical condition. Apparently as a consequence of some careless surgical procedures, he suffered some unanticipated postoperative complications. Although not permanently disabling, these complications had prolonged Mr. A's recovery, delayed his return to work, and required him to undergo physical rehabilitation. His attorney believed that medical malpractice could be demonstrated and that Mr. A was entitled to compensatory and punitive damages. He believed further that his client must have suffered psychological and physical distress, which would warrant a larger damage award than if his iatrogenic problems were only physical.

Results of the psychologist's evaluation suggested that Mr. A was an emotionally resourceful individual who was coping effectively with his unfortunate medical situation. Compared to most people with his illness who had required surgery, he seemed to be adjusting well psychologically. He was, in fact, the kind of patient one hopes to see on a rehabilitation service, a patient with a good prognosis for physical recovery with minimal attendant psychological distress. As a compassionate individual, the attorney was pleased to receive an oral report to this effect. As Mr. A's representative in a medical malpractice suit, however, he recognized that the psychologist's opinion, if introduced as evidence, might reduce the amount of damages that would be awarded. Hence, he did not request a written report or any further input from the psychologist, nor was he obliged by the rules of evidence to disclose to the court that this evaluation had been done.

Case 2. An attorney sought a psychologist's opinion while preparing to defend Mr. B, a 37-year-old elementary school teacher accused of sexually molesting several girls in his fifth-grade class. Mr. B had allegedly fondled these girls during class sessions by reaching into their underclothes while he was sitting at his desk and they approached him to ask a question or turn in an assignment. Aside from doubting that his client would have committed such acts so publicly, the attorney was puzzled by the manner in which the complaints had emerged. The initial allegation of molestation was made by just one girl, who spoke to her parents about it. Later, after this girl's parents had talked with the school principal and the police had been called in to investigate, several other girls in Mr. B's class told their parents that they had also been fondled by him.

Interestingly, reports from the school indicated that the girl who had complained first was socially popular and a leader among her peers. Also of note was the impression of several interviewers that none of the girls seemed particularly upset while talking about having been molested. To the contrary, they told their stories as if they were pleased and proud to relate them. As for Mr. B, he had for many years been a highly admired teacher in this school, known especially for a warm and caring attitude toward his students.

The attorney wondered whether there was any reason to think that a group of girls who had not been molested would allege as much. Some possibilities will come quickly to the minds of psychologists familiar with the romantic fantasies and peer-group interactions that commonly characterize prepubescent development in 10- and 11-year-old girls. Consider, for example, the possibility of a young girl first fantasizing about an appealing, perhaps paternal male teacher making a sexual overture toward her and then fabricating such a story as a way of feeling attractive and grown up and impressing her parents and peers in certain ways. Consider further the possibility of other girls in the class, having heard this story from a popular trendsetter, deciding to claim "He did it to me, too." This is the well-known stuff of which mass hysteria is made, as described in such classic papers as "The 'Phantom Anesthetist' of Mattoon" (Johnson, 1945) and in the recounting of the

Salem witch trials (Starkey, 1949) and their dramatization in Arthur Miller's play *The Crucible*. More recent scientific publications have described group episodes of psychological distress and neurological dysfunction apparently triggered by some event or one person's complaint (Balaratnasingam & Janca, 2006; Bartholomew & Sirois, 2000; Hatfield & Rapson, 2008; Roach & Langley, 2004).

In mentioning these possibilities to the attorney, the psychologist identified them as clinical formulations that could not be substantiated with solid empirical evidence. Having a possible explanation that child specialists would find plausible is different from having a line of defense that will stand up in court. The psychologist advised the attorney that his speculations in this case, if offered in testimony, could be made to look foolish under skillful cross-examination and might thereby detract from other aspects of the defense he was preparing. Hence, a report was not written, and the psychologist did not participate further in the case.

Case 3. An attorney representing a young man charged with a serious crime believed that his client was incompetent to stand trial. The attorney had been struck by his client's strange and disturbed behavior and expected that a psychological evaluation would provide supporting evidence for this claim. The psychologist concluded from his evaluation that the young man was indeed acting strangely but was feigning impairment. When he conveyed this opinion to the attorney in a telephone conversation, the attorney indicated that he should send a bill and would be paid for his time but that he would not be asked to testify or submit a report.

ETHICAL PRINCIPLES AND PROFESSIONAL REALITIES

For psychologists unfamiliar with forensic consulting, the preceding three illustrations of being dismissed from a case prior to preparing a report might raise some unsettling questions about proper practice. One might be especially concerned about Case 3, in which the psychologist provided an expert opinion that a man accused of committing a felony was feigning mental disturbance. How could the attorney ignore this finding and continue constructing a case for incompetency, and should the psychologist allow this to happen? The answers to these questions touch on some ethical and realistic considerations in the intersection of law and psychology.

Quality of Expert Opinions. To prepare themselves for sometimes unenthusiastic responses to their opinions, forensic consultants need to remain sufficiently humble to recognize that they may at times be in error, or at least not possessed of all the answers. Clinicians must appreciate that their skill and judgment do not transcend all of the imperfections in their assessment methods, and there may be critical facts in a case that have not come to their attention. As Shapiro (1991, chapter 4) reminds forensic consultants, expert psychological opinions are not statements of fact but

only reasonable conclusions based on careful analysis of the information they have been able to obtain.

The psychologist in Case 3 was reasonably certain that the defendant was malingering, but he would not have been prepared to testify that he was absolutely certain—nor indeed should he have been, given the difficulty of establishing malingering with absolute certainty (see Berry, Sollman, Schipper, Clark, & Shandera, 2009; Rogers, 2008). From the attorney's point of view, then, the opinion concerning malingering could be taken as a possibility, but not as the only possibility. The attorney might also have had in hand information unknown to the psychologist, perhaps even another expert opinion that in his view argued strongly against malingering.

Like the imperfections of assessment methods, the existence of sharply differing expert opinions brings a sobering measure of reality into forensic consulting. One expert's opinion is neither the only nor the last word. There are no obligations that would have prevented the attorney in Case 3 from listening to the psychologist's opinion concerning malingering and then turning to a different consultant, or perhaps a string of consultants, until he found a qualified psychologist who held the opinion that the defendant was truly incapable of understanding the proceedings against him and consulting effectively with his attorney, and therefore incompetent to stand trial according to the *Dusky* standard (see Zapf, Roesch, and Pirelli, Chapter 11 this volume; see also Grisso, 2003).

Learning of such an outcome, the first consultant could feel strongly that this last expert lacked sufficient experience or diagnostic acumen to recognize a clear case of malingering. However, the court in such a case would ordinarily accept a licensed psychologist with some experience in assessing competence in criminal defendants as qualified to offer such an opinion. To be sure, issues of competency and criminal responsibility commonly feature expert testimony from both parties, and arguments may ensue concerning which of several qualified professionals is best qualified to give reliable testimony. The point remains, however, that it is entirely appropriate and consistent with prevailing practice for attorneys to challenge or reject the opinions of a consultant they have retained and to seek other consultants whose opinions will provide better support for their case.

Considerations in Practicing Law. Instead of questioning the quality of expert opinions that fail to meet their needs, attorneys may decide on the basis of a consultant's conclusions to change their approach to a case or withdraw from it. For example, becoming convinced that a client in a criminal case has been faking emotional disturbance, is lying about guilt, or has in other ways behaved in a reprehensible manner may lead an attorney to decline to represent that person further or, if court appointed, to ask to be excused from the case.

Yet our system of criminal justice entitles everyone to a defense, no matter how reprehensible the offense, how despicable the alleged offender, or how guilty the accused appears to be. Regardless of how many attorneys choose not to represent

certain kinds of clients, every criminal defendant has the right to be represented by a member of the bar. Furthermore, the attorney eventually retained or appointed to provide this representation is ethically responsible for presenting the strongest possible case on behalf of the defendant. A weak or halfhearted defense of a defendant whom an attorney regards as guilty or as having few redeemable qualities can result in the attorney's appearing inept or unethical in the eyes of the legal community. Ineffective or unprepared trial lawyers may even risk being publicly chastened by the bench for having done a poor job on their client's behalf. Moreover, a decision reached in such an instance could well be reversed on appeal to a higher court on the basis of the defendant having had ineffective counsel.

In addition to preventing such negative consequences, a strong case presented on behalf of a client considered difficult to defend can enhance an attorney's professional reputation. Hence, trial attorneys may enjoy or even seek out opportunities to take on challenging cases and construct convincing briefs in them, especially in trials that capture media attention. Cases in point are such media events as the murder trials of O. J. Simpson, Timothy McVeigh (Oklahoma City bomber), Theodore Kaczinski (Unabomber), Lee Boyd Malvo (D.C. sniper), and, more recently, Casey Anthony (alleged daughter killer) and George Zimmerman (stand-your-ground shooter).

Forensic psychologists need to appreciate these realistic and proper motivations for attorneys to continue building a case that an expert consultant considers flawed. Moreover, there are instances in which attorneys have no choice but to continue with a case, regardless of reservations about the worthiness of the client or the weight of the evidence. For example, defense attorneys appointed by the court are rarely given an option to withdraw from the case, and they can expect the court to have little tolerance for a lackluster effort on their part. Likewise, public defenders and prosecuting attorneys may be assigned cases by the office for which they work, without being given much latitude to choose which ones they would prefer to try. These various considerations provide ample basis for conscientious and ethical attorneys to decline having their consultants furnish evidence that would damage their case.

Considerations in Practicing Psychology. Turning now to the second troublesome question raised by Case 3, how could the psychologist allow pertinent information to be suppressed? Being reasonably certain from his data that the accused was malingering psychosis, how could he sit silently while a competency hearing was taking place? To make matters worse, suppose that media accounts of the hearing were predicting that the defendant would be found incompetent to stand trial because of emotional disturbance.

Psychologists struggling with this kind of question must recognize that their dismay derives from their professional experience with the case conference model. In the case conference model, all relevant information is sought and alternative opinions considered in arriving at a diagnostic formulation and treatment plan.

This model is seldom approximated in forensic consultation, except when the client is the court. Then, as noted earlier, any testimony that helps the court reach its decisions is usually welcome.

For attorneys, however, who in conformance with the adversarial system argue just one side of a case, the only welcome testimony consists of evidence and opinions that support their arguments. If expert opinions exist that would support the other side of the case, it is up to opposing counsel to discover and produce these opinions. Confronted with such realities on cases in which they have consulted, psychologists may experience disappointment, anger, or perhaps even a sense of outrage. They may feel that situations like Case 3 call for their findings to be brought to light, to prevent a malingering criminal from escaping justice. They may even be tempted to blow the whistle by calling the prosecuting attorney or the judge and volunteering their opinion, or by informing the media that critical information concerning the case had been suppressed.

Except in extraordinary circumstances, responsible psychologists must resist any such temptations. To do otherwise would abuse the defendant's right to confidentiality and thereby violate the Ethical Principles of Psychologists and Code of Conduct adopted by the American Psychological Association ([APA], 2002, 4.01 & 4.05), which state:

Psychologists have a primary obligation and take reasonable precautions to protect confidential information obtained through or stored in any medium, recognizing that the extent and limits of confidentiality may be regulated by law or established by institutional rules or professional or scientific relationship . . . Psychologists disclose confidential information without the consent of the individual only as mandated by law, or where permitted by law for a valid purpose. (p. 1066)

Finally, with respect to deciding whether a report should be written following a forensic consultation, psychologists should heed the urging of Goldstein (2007) and Packer and Grisso (2011) to be sufficiently aware of applicable case and statutory law and the implications of their findings to advise their attorney clients concerning whether a report is likely to be helpful to them. In addition to being much appreciated, informed opinions of this kind sometimes can result in attorneys requesting a report of findings that would appear to weaken their position, as in the next case.

Case 4. An attorney representing Ms. C requested an evaluation of the extent to which she was suffering from posttraumatic stress disorder or any other psychological problems as a consequence of an automobile accident in which she was involved. On the basis of ample historical data, a detailed clinical interview, and findings from a comprehensive battery of psychological tests, the psychologist concluded that the plaintiff was not experiencing any emotional or behavioral difficulties that were attributable to the accident. Discussing his impressions in an

informal conversation with the attorney, the psychologist told her, "I don't think I can help you." Surprisingly, the attorney replied that she could put the psychologist's negative findings to good use. She explained that Ms. C had been pressing her to seek compensation for psychic damage and disability as well as for her well-documented physical injuries. An unambiguous written statement concerning the negative results of the psychological examination would help her convince her client to drop this part of her claim, she said, thereby sparing her from having to pursue a weak part of her case and allowing her to focus on the strong part. A written report was accordingly prepared.

DETERMINING THE FOCUS OF FORENSIC REPORTS

Once a decision has been made that a report will be written, the forensic psychologist must decide what to include in it. Now is the time to keep in mind that whatever is written down is discoverable and may become entered into evidence. Moreover, when psychologists are called to testify in a legal proceeding, their testimony ordinarily will be based on their written report, which means that everything in the report will be open to questioning on cross-examination. As a basic principle, then, forensic psychologists should limit their written reports to statements they will feel comfortable hearing read aloud in the courtroom and to conclusions they believe they can defend against reasonable challenge.

Beyond this preliminary consideration, the appropriate focus of forensic reports varies from one case to the next in relation to the needs of the client. As in providing other kinds of psychological services, forensic consultants should be guided by the familiar principle of giving clients what they want, within appropriate limits dictated by their professional judgment, ethical standards, and legal requirements. In this last regard, forensic examiners should be mindful of any specific court provisions in jurisdictions in which they testify. In federal court, for example, the Federal Rules of Civil Procedure (U.S. Government, 2010) provide that experts' written reports must include:

- a complete statement of all opinions the witness will express and the basis and reasons for them;
- the facts or data considered by the witness in forming these opinions;
- any exhibits that will be used to summarize or support the opinions;
- the witness's qualifications, including a list of all publications authored in the previous 10 years;
- a list of all other cases in which, during the previous 4 years, the witness testified at trial or by depositions; and
- a statement of the compensation to be paid for the study and testimony in the case.

To return to the principle of giving clients what they want, this guideline does not imply that psychologists should provide attorneys whatever opinions or

conclusions the attorneys would like to have. Meeting the client's needs refers to providing the desired services, not the desired findings. Reports should accordingly focus on matters of concern to the client without including all of the psychological observations that could be made about a person or situation being evaluated. This aspect of focusing forensic reports is embodied in the APA (2002, §4.04, p. 1066) Code of Conduct: "Psychologists include in written and oral reports and consultations only information germane to the purpose for which the communication is made." The Specialty Guidelines for Forensic Psychology state similarly that "[f]orensic examiners seek to assist the trier of fact to understand evidence or determine a fact in issue, and they provide information that is most relevant to the psycholegal issue" (APA, 2013, 10.01).

As a further note concerning propriety in this regard, it is helpful to distinguish between acts of *commission* and acts of *omission* in writing reports. Regarding the former, forensic psychologists should under no circumstances compromise their integrity by knowingly making inaccurate or misleading statements. Regarding omission, however, it is rarely warranted or necessary to answer questions that the client has not asked. Forensic reports focused within the limits described thus far will nevertheless vary in breadth as a function of the nature of the case and the line of attack or defense the client is intending to pursue.

PROVIDING NARROWLY FOCUSED CONSULTATIONS

In some forensic cases, the questions being asked by the client call for fairly limited data collection and a rather narrowly focused written report. The next two cases illustrate such circumstances.

Case 5. A young man accused of burglarizing some homes in his neighborhood had signed a confession. His attorney believed that he had been frightened into signing a confession that he was incapable of understanding. The psychologist who was asked to assess this possibility administered the Wechsler Adult Intelligence Scale–IV (WAIS-IV) and the Shipley Institute of Living Scale. The accused appeared unfamiliar with many of the vocabulary items in these two tests, including several that are listed in the Thorndike-Lorge index as being more frequently used than some of the key words in the confession he had signed. The psychologist's report stated this finding and suggested that it might have implications for considering whether the young man fully understood the text of his confession.

Case 6. An attorney preparing to plead diminished capacity in defending a man charged with attempted murder had received some divergent reports from several consultants concerning her client's mental status. As one step in trying to resolve this discrepancy, she asked a psychologist experienced in the Rorschach assessment of schizophrenia if he would review the defendant's Rorschach protocol and answer

two questions: Had the record been taken properly, and was it consistent with a diagnosis of schizophrenia as defined in the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; APA, 1994)*? In the consultant's opinion, the answers to these two questions were yes and no, respectively, and this is what was communicated to the attorney in a relatively brief and narrowly focused report.

These examples of narrowly focused forensic consultations may appear to reflect some undesirable clinical practices. First of all, the psychologist in both cases based his opinion on just one or two specialized tests, whereas psychodiagnostic assessment as commonly practiced involves a multifaceted test battery (see Groth-Marnat, 2009; Weiner & Greene, 2008). Second, in Case 6, the psychologist conducted a blind analysis of the test protocol without examining the accused or knowing anything about him except his age and sex. Most clinicians regard such blind analysis as appropriate only for didactic or research purposes. Moreover, most clinicians concur that diagnoses should be made by them, not by their tests, and only following thoughtful integration of test findings with relevant information about a person's history and circumstances (Harwood, Beutler, & Groth-Marnat, 2011; Shapiro, 1991; Weiner, 2005, 2013).

Psychologists concerned about such matters might be reluctant to provide the narrow kind of consultation requested in Cases 5 and 6. From the perspective of attorneys attempting to muster bits of evidence in support of their case, however, the request in both cases for a narrowly focused consultation was entirely appropriate. It was likewise appropriate for the psychologist to respond just to the questions being raised, provided that he felt confident of his ability to do so. Moreover, submitting a narrowly focused report did not prevent the psychologist in Case 6 from also assuming an educative function and pointing out to the attorney that an informed psychodiagnostic opinion concerning her client's mental status would have to be based on results of a test battery and contextual information, not just blind evaluation of a single test protocol.

BROADENING THE FOCUS OF FORENSIC CONSULTATION

Although instances of narrowly focused consultation are important to identify and put in perspective, they seldom occur in the practice of most forensic psychologists. Instead, consultative requests are likely to require attention to multiple sources of information in the preparation of reports. Even when psychologists are consulted primarily as experts in psychological testing, they should conduct themselves as broadly knowledgeable mental health professionals who integrate interview data, background information, and test findings into comprehensive opinions and conclusions concerning the case. Psychologists who testify on the basis of a written report that deals solely with test findings, and who must plead ignorance when asked about other elements of the case, are poorly prepared to present themselves effectively in the courtroom. They weaken their client's arguments by being exposed

as an expert who has only a superficial grasp of the case, and they demean their own professional status by failing to present themselves as anything more than a tester.

Adequate attention to the context of a forensic case does not always call for extensive data collection or record review, however. Sometimes just a few bits of background information suffice for preparing an effectively focused report, as in the next case.

Case 7. Mr. D was a 34-year-old man who had suffered a closed head injury in an accident for which there was alleged liability. He had been rehabilitated on a neurological service to the point where he was considered to have achieved his maximum recovery. His attorney wanted to establish how much permanent loss of function remained as a consequence of the accident. A Wechsler, administered as part of an extensive test battery, yielded a Full-Scale Intelligence Quotient (FSIQ) of 103. Although no pre-accident IQ score was on record in Mr. D's case to provide a baseline, his history documented that he had received a PhD in chemical engineering from a prestigious university and had enjoyed a successful career in his profession up to the time of his accident. Given the limited likelihood of such accomplishment by a person with an IQ of 103, and assuming that Mr. D was making a good effort when tested, these findings provided grounds for arguing that the accident had caused at least some loss of his mental capacity, probably too much for him to be able to resume his career. The written report accordingly focused on these particular findings.

Turning to more general guidelines in determining the scope of a report, the breadth of the psychologist's focus typically is influenced by the time frame of the inquiry, with specific respect to whether attention must be paid to present, past, or future circumstances.

Addressing Present Circumstances. When forensic psychological opinions must address primarily the present status of a plaintiff or defendant, the data that need to be collected are relatively limited, and the task of interpreting them is correspondingly uncomplicated. As previously noted, for example, questions of whether defendants are competent to stand trial concern mainly whether they are currently able to understand the proceedings and participate effectively in their defense (see Zapf, Roesch, and Pirelli, Chapter 11 this volume). Consultants may struggle with translating these legal criteria for competency into psychological terms, and they may encounter instances of marginal competence that are difficult to call one way or the other. Whatever the difficulty of the task in these evaluations, however, the critical data for determining a defendant's present functioning capacity will be available from currently obtainable interview, test, and observational data.

Assessment of personal injury also focuses mainly on current mental or emotional state and an individual's present functioning capacity. Evaluations of allegedly

reactive psychological conditions or loss of functioning capacity are a bit more complicated than determining competency, however, because current capacity must ordinarily be compared to some baseline of previous functioning, prior to an allegedly harmful incident (see Piechowski, Chapter 7 this volume; see also Greenberg, Otto, & Long, 2003; Kane & Dvoskin, 2011). In most cases, verifiable records of past events provide a baseline for such comparisons. These records may include previously obtained intelligence, neuropsychological, or personality test findings; documentation of a claimant's educational and occupational history (as in Case 7); clinicians' and hospital notes concerning prior medical problems and mental health services; and testimony and accounts from relatives and longtime acquaintances concerning the plaintiff's earlier patterns of behavior and adjustment.

Addressing Past Circumstances. Opinions that must address the individual's mental status at some prior point in time ordinarily require more extensive data collection than present status evaluations and a more broadly focused report in which the conclusions are less certain (see Weiner, 2003). The most commonly encountered cases of this kind involve questions of criminal responsibility. Criminal responsibility is determined by the nature of a defendant's mental state at the time of an offense and whether this mental state contributed to the commission of the offense (see Zapf, Roesch, Golding, and Pirelli, Chapter 12 this volume). Efforts to establish a prior mental state and its likely consequences require forensic psychologists to seek out information and confront uncertainties that stretch their capacities well beyond the relatively modest demands of conducting a present status evaluation.

For example, suppose an adult male defendant whose attorney is pleading him not guilty by reason of insanity to a felonious assault committed three months earlier shows historical and psychological test evidence of an apparently chronic and long-standing schizophrenic disorder. This finding may constitute good reason to believe that the accused was psychologically disturbed three months ago and probably well before that. Even if uncontested, however, this conclusion would not necessarily demonstrate that the defendant's disorder was active at the time of the offense and responsible for his having committed it.

Case 8. Just prior to the 1984 Summer Olympics in Los Angeles, a man named Daniel Lee Young drove his car recklessly onto a crowded sidewalk, killing one pedestrian and injuring 54 others. In his subsequent trial he was identified as having a chronic paranoid schizophrenic disorder. It was nevertheless found that his schizophrenia was not a contributing factor in his assaultive and homicidal behavior. He was considered to have been legally sane at the time of the crime and was sentenced to a prison term of 106 years and 4 months to life on one count of first-degree murder and 48 counts of attempted murder.

Suppose, by contrast, that an offender pleading temporary insanity or diminished capacity appears at present to be psychologically capable and well functioning,

thus making it doubtful that he or she was seriously disturbed just a few months earlier, when the crime was committed. Could it nevertheless be argued that any person, no matter how well functioning at the moment, could fall prey to an acute psychotic or dissociative episode during a time of duress? Or could it be concluded from currently obtained interview and test data that a defendant is not the kind of person who is likely to show psychotic or dissociative reactions to stress? Neither conclusion, whatever its psychological justification, would carry much evidentiary weight, unless it could be convincingly amplified with respect to (1) the nature and amount of stress the defendant was actually likely to have been experiencing shortly before or at the time of the criminal act and (2) how he or she was actually behaving prior to and while committing the criminal act. With this in mind, forensic psychologists addressing questions of criminal responsibility need to investigate carefully and report clearly the events leading up to and occurring during the commission of a crime. The defendant's own recollections, the police arrest report, and statements given by eyewitnesses and other informants should be integrated with current personality evaluations to yield informed opinions concerning whether the stresses in a defendant's life and his or her behavior while committing a crime seem consistent with applicable criteria for reduced criminal responsibility.

Addressing Future Circumstances. In other types of forensic cases, the questions being asked challenge psychologists not to reconstruct the past but to predict the future. Two of these types of cases involve requests for aid in sentencing lawbreakers, sometimes with respect to whether a newly convicted offender should be sent to prison or put on probation, and at other times concerning whether an already incarcerated offender should be paroled. Offering reasonable opinions about probation and parole requires being able to estimate (1) how likely the offender is to commit future crimes, especially violent ones; (2) how responsive the person will be to counseling, psychotherapy, job training, or other rehabilitative efforts outside of a prison setting; and (3) how adequate the available services are for providing the kinds of community interventions that offer promise of a successful outcome.

Estimates of these kinds are usually difficult to make on the basis of currently available information, and forensic reports addressed to future possibilities typically must be more detailed and tentative than reports addressing present functioning. To extrapolate accurately from current assessment data to future expectations bearing on the advisability of probation or parole, forensic consultants need to collect whatever clinical and empirical knowledge they can obtain concerning recidivism, violence risk, treatment response, available resources, and the offender's anticipated environmental circumstances and relate this information to the case at hand.

The other type of forensic case that addresses future circumstances involves contested child custody and visitation rights between parents who are separated or divorced. What arrangements will be in the best interests of the children? Which parent is likely to provide better child care and supervision? When and under what circumstances should the noncustodial parent have access to his or her children?

Like estimating the advisability of probation or parole, these child custody questions are difficult to answer with certainty from currently available data. Among other demands faced by forensic psychologists in family law cases, their examinations and reports typically must embrace all members of the family who will be affected by a custody decision, including both parents, all dependent children, and other significant figures in the home or in a child's life. Should a remarried parent be seeking to gain primary residential custody, the future stepparent may also need to be evaluated. In addition, to lend some reasonable certainty to their efforts to predict the future, psychologists offering opinions in custody cases need to draw on relevant research findings and clinical wisdom concerning developmental aspects of child–parent relationships and the impact of divorce on children and their parents (see Stahl, Chapter 6 this volume; see also APA, 2010; Condie & Condie, 2007).

ON BEING CLEAR, RELEVANT, INFORMATIVE, AND DEFENSIBLE

So far this chapter has indicated that writing useful and effective forensic reports requires psychologists to have a good grasp of the legal and behavioral issues surrounding a case, to determine what kinds of information will best help resolve these issues, and then to gather and evaluate this information. Once these tasks are accomplished, what remains is for consultants to present their impressions and conclusions in a clear, relevant, informative, and defensible manner.

BEING CLEAR

Forensic consultants ordinarily should begin their reports by indicating when, by whom, and for what purposes they were appointed or retained. They should then identify the sources of information they have utilized. When, where, for what reason, and in what fashion were parties to the case directly evaluated? What records were examined, such as depositions, police reports, medical records, and school or military files? What collateral persons were interviewed? To what extent were other discussions, reviews of psychological literature, or examinations of case law undertaken to further the consultant's knowledge and understanding of the case? Explicit answers to these questions in the introduction to a forensic report promote clarity by minimizing uncertainty concerning the basis on which consultants have formed the opinions they are about to state in their report.

In stating their findings and conclusions, forensic psychologists should write in plain English and limit their use of technical jargon. As previously mentioned, a written report may be gone over in painstaking detail by opposing counsel during a deposition or read aloud in its entirety in the courtroom. Some attorneys may even prefer on direct examination to have a consultant's report entered verbatim, to avoid having imprecise or poorly worded statements slip into the consultant's extemporaneous presentation. Hence, consultants should not plan on writing a formal, somewhat technical report for the record and then giving

their courtroom testimony in an informal, conversational manner that is easy to follow and understand. Instead, the written report itself should be as clear and conversational as the psychologist can make it. This means using unstilted and uncomplicated language that will be comfortable for consultants to repeat on the witness stand, that will be comprehensible to judge and jury, and that will limit a cross-examining attorney's opportunities to badger them with questions about what their statements mean.

Along with using ordinary language in reports, except where technical terms may be required (e.g., a formal diagnosis), psychologists should concentrate on writing about the person they have evaluated rather than about psychological processes. A statement like "Coping capacities are good" does not communicate as clearly as "Ms. E has good capacities for coping with stressful experiences without becoming unduly upset by them." When psychologists fail to guard adequately against being murky, impersonal descriptions of psychological processes often go hand in hand with jargon. Compare, for example, "Homophobia is pronounced" with "This man tends to avoid people because he is unusually fearful of being harmed or taken advantage of by others." Sometimes consultants may not realize that certain expressions commonly used by professionals are not generally understood by the public. For example, "Reality testing is poor" reads better as "Mr. F's reality testing is poor" but even better as "Mr. F often forms inaccurate or mistaken impressions of people and events, and he is consequently more likely than most people to show poor judgment and fail to anticipate the consequences of his actions." Similar recommendations for making psychological reports readable and useful are elaborated by Ackerman (2006), Blais and Smith (2008), Conroy (2006), Goldfinger and Pomerantz (2010), Greenfield and Gottschalk (2009), Harvey (2006), and Karson and Nadkarni (2013).

BEING RELEVANT

As in responding to consultation requests in other areas of practice, psychologists achieve relevance in forensic reports by addressing and attempting to answer the referral question. Being relevant means omitting much of what could be said about an individual's psychological characteristics and probable ways of responding in various circumstances and instead providing a distillate of those features of the individual that bear directly on the issues in the case and the client's questions about these issues.

But what are the client's questions? To some extent, relevance is achieved by adhering to the previously noted ethical guidelines concerning appropriately focused forensic reports. However, to translate this concept into practice—and thereby conduct an adequate evaluation and write a relevant report—forensic psychologists need to pursue some specifically stated question, such as whether an accused is competent to stand trial or an allegedly brain-injured person has suffered demonstrable loss of intellectual or cognitive function. If no such question has been

framed, one must be elicited from the client by asking “Why do you want to have this person evaluated?” or “What is it that you would like to learn from me?”

In addition to identifying what information to obtain and how best to organize and report it, specific referral questions also help psychologists anticipate the potential certainty and utility of the psychological data they will obtain. These expectations often can be shared with clients to good effect prior to conducting an evaluation. For example, alerting an attorney that evaluations of possible future behavior generate less certain results than evaluations of present status can enhance the effectiveness of an eventual report by minimizing any unwarranted expectations on the attorney’s part. Cases 2 and 4 presented earlier also illustrate how working with a clear referral question can facilitate providing a relevant consultation.

Forensic psychologists can increase the relevance of their consultations further by drawing on familiarity with applicable statutory and case law in a particular case. Awareness of applicable legal standards is an ethical responsibility in forensic practice (APA, 2002, 2.01f; APA, 2013, 2.04 [The Specialty Guidelines are reprinted as the appendix to this volume with permission of the APA]), and judicious integration of such knowledge into a report, especially with use of appropriate terminology, usually will enhance its relevance.

For example, psychologists preparing reports in personal injury cases should address in specific terms the issue of “proximate cause,” as spelled out in tort law (see Piechowski, Chapter 7 this volume), and those preparing reports in custody cases should pay particular attention to the “best interests of the child” doctrine, which is a prominent theme in family law (see Stahl, Chapter 6 this volume). The differing criteria employed in various state and federal jurisdictions for what constitute mitigating mental circumstances in criminal behavior also illustrate the necessity of adequate legal knowledge (see Clark, Chapter 13 this volume; Zapf, Roesch, Golding, and Pirelli, Chapter 12 this volume). To express a relevant opinion concerning criminal responsibility, forensic psychologists must appreciate how their findings fit with applicable ways of defining it and express themselves accordingly. In a jurisdiction in which the M’Naghten rule applies, for example, the utility of the consultant’s report is enhanced by the next kind of statement:

Mr. G frequently has difficulty perceiving events in his life realistically, and as a result he often misjudges how his behavior affects other people. The severity of this problem and strong indications that he has had it for a long time make it reasonable to think that he was not fully capable of appreciating the wrongfulness of his actions at the time of the crime. Significant in this regard, when asked directly if he thought he was doing anything wrong, he said, “I had every right to do it.”

BEING INFORMATIVE

Like clinical reports, forensic reports should be written in an informative manner that educates the nonpsychologist reader. This informational objective usually can be achieved by relating psychological data and impressions to benchmarks that

readers will recognize. For example, saying that a respondent has obtained “a WAIS FSIQ of 100” communicates adequately to other psychologists about the person’s overall IQ level, but this statement may be unintelligible to most laypeople unless it is amplified with some additional information: for example, that the Wechsler scale is the most widely used measure of adult intelligence; that it comprises several subtests sampling different kinds of abilities; that, although there is some measurement error associated with a Wechsler IQ, an obtained score of 100 gives a 95% probability that the respondent’s true IQ is between 95 and 105; and that about half of U.S. adults receive an IQ score above 100 on this test and about half a score below 100.

Similarly with respect to impressions of psychological disorder, consultants should indicate how the findings compare with the diagnostic criteria of widely used nomenclatures, as was done in Case 6. For example, the issues in a case may call for this kind of statement: “The way this person is thinking and feeling, as reflected in the interview and psychological test findings, is consistent with a *DSM-IV-TR* diagnosis of Major Depressive Disorder.” There may also be instances in which consultants find it useful to draw on textbook information to summarize a set of circumstances that point to a particular conclusion. For example, a psychologist consulting on a criminal case in which the court was considering a suspended sentence wrote this informative opinion:

I am concerned about having this man return without supervision to his previous place of residence. Being a white male in his late 50s, who would be living alone in a run-down section of town, and who has previously attempted to take his own life, he would be in a very high risk group for suicidal behavior.

An informative educational approach of this kind, in a report that is easy to understand and speaks explicitly to the issues at hand, promotes effective communication. Combined with good judgment concerning when reports should be written and how narrowly or broadly they should be focused, skills in being clear, relevant, and informative in writing reports contribute substantially to providing effective forensic consultation. What remains to be considered are some ways of writing forensic reports that can enhance their defensibility in the face of challenge.

BEING DEFENSIBLE

Unlike clinical reports, which are commonly either praised or ignored but rarely demeaned, at least not publicly, forensic reports are fair game to opposing attorneys. Their duty to their client calls for them to challenge the accuracy and import of what consulting psychologists on the other side have written while making opposing experts as uneasy, uncertain, and unbelievable as possible on the witness stand. Forensic psychologists can spare themselves such grief by reporting what they have to say in the most defensible manner possible. Weiner (2009) has elaborated

various means of minimizing ethical and legal jeopardy in conducting personality assessments, and these include four considerations in report writing that can help forensic consultants avoid potential pitfalls in giving expert witness testimony.

First, forensic psychologists are well advised to favor description over categorization in their conclusions about people they have evaluated. The ground underfoot is safer when describing how a respondent resembles certain types of people who have had certain kinds of experiences than it is when categorizing a respondent as being a specific type of person who has had specific experiences. The following pairs of statements illustrate the difference between describing and categorizing people:

“This woman shows many features in common with people who have developed a stress disorder subsequent to a traumatic experience” versus “This woman has a Posttraumatic Stress Disorder.”

“Children with the kinds of personality characteristics I found in Suzie often have not had the benefit of receiving much nurturance from their parents” versus “Suzie has not been adequately nurtured by her parents.”

“Mr. H’s attitudes and dispositions closely resemble those often seen in persons who act violently toward others” versus “Mr. H is likely to act violently toward others.”

The second part of each of the preceding pairs seldom causes problems when it appears in clinical reports of what an examining psychologist believes. In forensic reports, however, such categorization exposes consultants to thorny questions concerning how they can be sure that people have specific conditions or action tendencies being ascribed to them and whether the consultant actually was present to observe the experiences and events alleged to have occurred. By describing rather than categorizing people, psychologists can blunt the thrust of such challenges. In the example of Suzie, the consultant who describes her does not say that she was inadequately nurtured by her parents, only that she shares certain characteristics with children who have not been adequately nurtured.

Second, relative statements about people usually create fewer difficulties for forensic consultants than absolute statements. Couching statements about persons examined in forensic cases in terms of conditions they are more or less likely to have, rather than “definitely have,” and in terms of behavior they probably have shown in the past or may be inclined to show in the future, as opposed to “have no doubt shown” or “will surely show,” usually stand the consultant in good stead. Likewise, statements that paint people in relative and conjectural terms as being more or less likely than other people to show certain characteristics invite fewer challenges than pictures painted only in black-and-white certainty. For example, it is much easier to justify having written “Ms. J appears to be more self-centered than most people” than “Ms. J is a very self-centered person.”

Third, consistent with what psychologists know about the imperfect nature of their assessment tools, forensic consultants should avoid writing statements that

rule out conditions or events. The fact that certain conditions (e.g., some disorder) or events (e.g., having been sexually abused or abusive) are not suggested by a psychologist's data does not eliminate the possibility of their existing. To minimize their exposure to being challenged and possibly embarrassed as a consequence of having overstated their findings, forensic psychologists should focus their reports on what the data indicate is probably present (e.g., "The available data provide substantial indications that Mr. J is a psychologically stable, well-organized, and well-adjusted person"; "The evidence at hand suggests that Ms. K is quite depressed and possibly suicidal at the present time"). On those occasions when reporting of negative findings seems called for or is mandated, consultants will do well to exercise caution in drawing conclusions from these findings. Consider, for example, the next statement: "Although it is not possible on the basis of the test findings to rule out closed head injury, it can be said that the data obtained in the examination do not contain any evidence of neuropsychological impairment."

Fourth, forensic psychologists should avoid including illustrative test responses in their reports. Knowledgeable examiners are often tempted to illustrate their points with critical items from self-report inventories and rich content themes from performance-based tests (e.g., the Rorschach and Thematic Apperception Test). Even when certain repetitive self-reports and pervasive contents have compelling face validity, however, their presentation can cause examiners grief by opening the door to questions about what individual test responses mean. With a foot in this door, skilled cross-examining attorneys can make most psychologists and their tests look foolish by questioning the interpretive significance of subtle test items or responses. A helpful reply to courtroom questions about the meaning of individual items and responses is that they have little significance by themselves and acquire interpretive value only when combined into various multiple-item scales and configurations. Psychologists who have already identified individual responses in their written reports and given interpretive weight to them are poorly positioned to give this answer in the courtroom.

In summary, writing effective forensic reports begins with deciding whether a report should be written, which is a decision that should be based jointly on the preferences of the client and the consultant's commitment to ethical principles and professional realities that govern the practice of forensic psychology. When a written report is expected or required, as is usually the case in forensic work, consulting psychologists need to determine whether their report should be focused in a relatively narrow or relatively broad manner. This determination will be guided by the kinds of questions they are being asked in a particular case, which will have implications for how much information they collect, from what sources they collect this information, and how much of what kinds of information they include in their report. Good judgment concerning when reports should be written and how narrowly or broadly they should be focused, when combined with skill in writing clear, relevant, and informative reports that can be defended on the witness stand, are building blocks of providing effective forensic consultation.

REFERENCES

- Ackerman, M. J. (2006). Forensic report writing. *Journal of Clinical Psychology, 62*, 59–72.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.) Washington, DC: Author.
- American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57*, 1060–1073.
- American Psychological Association. (2010). Guidelines for child custody evaluations in family law proceedings. *American Psychologist, 65*, 863–867.
- American Psychological Association. (2013). Specialty guidelines for forensic psychology. *American Psychologist, 68*(1), 7–19. doi:10.1037/a0029889
- Balaratnasingam, S., & Janca, A. (2006). Mass hysteria revisited. *Current Opinion in Psychiatry, 19*, 171–174.
- Bartholomew, R. E., & Sirois, F. (2000). Occupational mass psychogenic illness: A transcultural perspective. *Transcultural Psychiatry, 37*, 495–524.
- Bennett, B. B., Bricklin, P., Harris, E., Knapp, S., VandeCreek, J., & Younggren, J. (2006). *Assessing and managing risk in psychological practice*. Rockville, MD: American Psychological Insurance Trust.
- Barrett, G. V., & Morris, S. B. (1993). The American Psychological Association's amicus curiae brief in *Price Waterhouse v. Hopkins*: The values of science versus the values of the law. *Law and Human Behavior, 17*, 201–216.
- Berry, D. T. R., Sollman, M. J., Schipper, L. J., Clark, J. A., & Shandera, A. L. (2009). Assessment of feigned psychological symptoms. In J. N. Butcher (Ed.), *Oxford handbook of personality assessment* (2nd ed., pp. 613–642). New York, NY: Guilford Press.
- Blais, M. A., & Smith, S. R. (2008). Improving the integrative process in psychological assessment: Data organization and report writing. In R. P. Archer & S. R. Smith (Eds.), *Personality assessment* (pp. 405–439). New York, NY: Routledge.
- Brodsky, S. L., & Robey, A. (1972). On becoming an expert witness: Issues of orientation and effectiveness. *Professional Psychology, 3*, 173–176.
- Condie, L. C., & Condie, D. (2007). Termination of parental rights. In A. M. Goldstein (Ed.), *Forensic psychology: Emerging topics and expanding roles* (pp. 294–332). Hoboken, NJ: Wiley.
- Conroy, M. A. (2006). Report writing and testimony. *Applied Psychology in Criminal Justice, 2*, 237–260.
- Goldfinger, K., & Pomerantz, A. M. (2010). *Psychological assessment and report writing*. Thousand Oaks, CA: Sage.
- Goldstein, A. M. (2007). Forensic psychology: Toward a standard of care. In A. M. Goldstein (Ed.), *Forensic psychology: Emerging topics and expanding roles* (pp. 3–44). Hoboken, NJ: Wiley.
- Greenberg, S. A., Otto, R. K., & Long, A. C. (2003). The utility of psychological testing in assessing emotional damages in personal injury litigation. *Assessment, 10*, 411–419.
- Greenberg, S. A., & Shuman, D. W. (1997). Irreconcilable conflict between therapeutic and forensic roles. *Professional Psychology, 28*, 50–57.
- Greenfield, D. P., & Gottschalk, J. A. (2009). *Writing forensic reports: A guide for mental health professionals*. New York, NY: Springer.
- Grisso, T. (2003). *Evaluating competencies: Forensic assessment and instruments* (2nd ed.). New York, NY: Kluwer Academic/Plenum Press.
- Groth-Marnat, G. (2009). *Handbook of psychological assessment* (5th ed.). Hoboken, NJ: Wiley.

- Hatfield, E., & Rapson, R. L. (2008). Emotional contagion. In M. Greco & P. Stenner (Eds.), *Emotions: A social science reader* (pp. 443–448). New York, NY: Routledge.
- Harvey, V. S. (2006). Variables affecting the clarity of psychological reports. *Journal of Clinical Psychology, 62*, 5–18.
- Harwood, T., Beutler, L. E., & Groth-Marnat, G. (2011). *Integrative assessment of adult personality* (3rd ed.). New York, NY: Guilford Press.
- Heilbrun, K., Grisso, T., & Goldstein, A. M. (2009). *Foundations of forensic mental health assessments*. New York, NY: Oxford University Press.
- Johnson, D. M. (1945). The “phantom anesthetist” of Mattoon: A field study of mass hysteria. *Journal of Abnormal and Social Psychology, 40*, 175–186.
- Kane, A. W., & Dvoskin, J. A. (2011). *Evaluation for personal injury claims*. New York, NY: Oxford University Press.
- Karson, M., & Nadkarni, L. (2013). *Principles of forensic report writing*. Washington, DC: American Psychological Association.
- Monahan, J. (Ed.). (1980). *Who is the client?* Washington, DC: American Psychological Association.
- Packer, I. K., & Grisso, T. (2011). *Specialty competencies in forensic psychology*. New York, NY: Oxford University Press.
- Roach, E. S., & Langley, R. L. (2004). Episodic neurological dysfunction due to mass hysteria. *Archives of Neurology, 61*, 1269–1272.
- Rogers, R. (Ed.). (2008). *Clinical assessment of malingering and deception* (3rd ed.). New York, NY: Guilford Press.
- Sales, B. D., & Shuman, D. W. (2005). *Experts in court: Reconciling law, science, and professional knowledge*. Washington, DC: American Psychological Association.
- Shapiro, D. L. (1991). *Forensic psychological assessment*. Boston, MA: Allyn & Bacon.
- Starkey, M. L. (1949). *The devil in Massachusetts*. New York, NY: Knopf.
- U.S. Government. (2010). *Federal rules of civil procedure*. Washington, DC: U.S. Government Printing Office.
- Weiner, I. B. (2003). Prediction and postdiction in clinical decision making. *Clinical Psychology: Science and Practice, 10*, 335–338.
- Weiner, I. B. (2005). Integrative personality assessment with self-report and performance-based measures. In S. Strack (Ed.), *Personality and psychopathology* (pp. 317–331). Hoboken, NJ: Wiley.
- Weiner, I. B. (2009). Anticipating ethical and legal challenges in personality assessments. In J. N. Butcher (Ed.), *Oxford handbook of personality assessment* (pp. 599–612). New York, NY: Oxford University Press.
- Weiner, I. B. (2013). The assessment process. In J. R. Graham & J. A. Naglieri (Eds.), *Assessment psychology* (pp. 3–25). Vol. 10 in I. B. Weiner (Ed.-in-Chief), *Handbook of psychology* (2nd ed.). Hoboken, NJ: Wiley.
- Weiner, I. B., & Greene, R. L. (2008). *Handbook of personality assessment*. Hoboken, NJ: Wiley.

CHAPTER 22

Testifying in Court

RANDY K. OTTO, SUSAN L. KAY, AND ALLEN K. HESS

SINCE their entry into juvenile courts at the beginning of the 20th century, psychologists have become increasingly involved in legal proceedings—particularly during the past 50 years (Otto & Heilbrun, 2002). In this chapter, we chronicle the rationale for and underpinnings of expert testimony, review key legal decisions that shape the presentation of expert testimony, and discuss strategies psychologists can employ to present their work and opinions to the court most effectively through sworn testimony.

FIRST EXPERT WITNESSES

The involvement in legal proceedings of witnesses who were called not because of what they knew about the case at hand but because of special knowledge they had about complicated issues related to the legal dispute occurred as early as the 13th century, when judges first called physicians and surgeons as consultants (Blau, 1998). Two centuries later, the work of expert witnesses took a more ominous turn when ecclesiastical courts relied on the work of Dominican friars Johann Sprenger and Heinrich Kraemer, authors of *The Malleus Malificarum* (*The Hammer of Witches*)—a treatise that instructed judges how to identify, interrogate, and convict witches (Millon, 1969).

Historically, juries were composed of people who had firsthand knowledge of the events at issue in the case. By the mid-1850s, there was a shift toward our current system of lay jurors. This shift, accompanied by the flood of technical knowledge derived from the Industrial Revolution, posed increasingly complex questions to the courts. When the professions were limited to clergy, military, lawyers, educators, and physicians, courts had little need for witnesses with special expertise. However, with the explosive development of knowledge and attendant professions, the judiciary came to conclude that, in some cases, it could benefit from the observations and opinions of persons with special knowledge or expertise about the matters at issue.

Psychologists can trace their entry into the courtroom to the beginning of the 20th century, when they began to testify about adolescents in proceedings occurring in juvenile courts, the first of which was established in Chicago in 1899 (Blau, 1998; Otto & Heilbrun, 2002). Although the myriad and varied contributions that psychology could make to the legal system were heralded by Hugo Münsterberg in *On the Witness Stand* (1908), his call for attorneys and judges to embrace researchers and practitioners of this new science went largely unheeded (Otto, 2012). After this false start, it was not until the latter half of the 20th century, roughly coinciding with the development of clinical psychology as an independently practiced profession, that psychologists regularly began to testify in legal proceedings.

Jenkins v. United States (1962) is a watershed case that many consider to be responsible for psychologists' wholesale entry into the courts. At the time of this decision, physicians typically provided expert testimony regarding the mental state of criminal defendants, as such was considered a medical matter. Nonetheless, the defense attorney called three psychologists who testified about Jenkins's mental state at the time of the offense. The trial court instructed the jury to disregard the psychologists' testimony on the grounds that these witnesses were unqualified to testify about mental disorder. Jenkins was convicted. On appeal, the United States Court of Appeals for the District of Columbia held that the trial court erred when it instructed the jury to disregard the psychologists' testimony on the grounds that some psychologists do have the requisite knowledge about mental disorders to be competent to provide expert opinions about defendants' mental states. The appeals court directed the lower court to make a case-by-case determination of the admissibility of the testimony of psychologists based on the substantive knowledge of the proffered expert.

LEGAL REFINEMENT OF THE USE OF EXPERTS

As is the case with all witnesses, the testimony of expert witnesses is shaped by various legal requirements.

GENERAL CONTEXT

Although evidence law varies across jurisdictions, many states have adopted rules of evidence that parallel, in full or in part, the Federal Rules of Evidence (FRE). The FRE most relevant to the control and presentation of expert testimony are discussed next.

With rare exceptions, fact or lay witnesses are limited to testifying about what they sensed—that is, what they saw, heard, felt, smelled, or tasted (FRE 701). In contrast, in addition to testifying about what they sensed, witnesses qualified as experts can offer their opinions. FRE 702 provides that:

A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if (a) the expert's scientific,

technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue; (b) the testimony is based on sufficient facts or data; (c) the testimony is the product of reliable principles and methods; and (d) the expert has reliably applied the principles and methods to the facts of the case.

Thus, witnesses proffered as experts can offer their opinions about complicated matters before the court provided that they (a) demonstrate their expertise, (b) have an adequate factual basis for their opinions, (c) employ valid techniques, and (d) validly apply their techniques to the facts in the case at hand. The party calling the expert witness bears the burden of proving that these requirements have been met.

FRE 703 makes clear the type and kind of information on which an expert may base his or her opinions:

An expert may base an opinion on facts or data in the case that the expert has been made aware of or personally observed. If experts in the particular field would reasonably rely on those kinds of facts or data in forming an opinion on the subject, they need not be admissible for the opinion to be admitted. But if the facts or data would otherwise be inadmissible, the proponent of the opinion may disclose them to the jury only if their probative value in helping the jury evaluate the opinion substantially outweighs their prejudicial effect.

Thus, experts are not limited to basing their opinions solely on their own observations. They may also rely on other information, including evidence that might not otherwise be admissible in court (e.g., hearsay), providing that it is the kind of information on which those in their profession typically rely.

FRE 704 provides that, with the exception of opinions about criminal responsibility, expert witnesses can offer opinions about the ultimate legal issues before the court. Thus, although some commentators argue that psychologists should not offer opinions on legal issues because these matters are ultimately legal/moral ones (see, e.g., Melton et al., 2007; Slobogin, 1989), the FRE (and many state counterparts) do not preclude psychologists from opining about the ultimate legal issue before the court (e.g., if the defendant is competent to stand trial, if the respondent meets criteria for civil commitment, if placement with mother is in the child's best interests).

FRE 705 provides direction as to the required factual basis for experts' opinions:

Unless the court orders otherwise, an expert may state an opinion—and give the reasons for it—without first testifying to the underlying facts or data. But the expert may be required to disclose those facts or data on cross-examination.

Thus, although the expert during direct examination can simply testify about the opinions he or she formed, the court or cross-examining attorney can ultimately force disclosure of the facts and information underlying the opinion. We recommend that the expert consult with retaining counsel to determine the best way to present his or

her testimony. Presentation of the facts and data underlying the opinion generally will bolster the opinion and thus, logically, should be provided. If the facts and data are composed of otherwise admissible evidence, we would recommend that the expert present those facts and data during direct testimony. If, however, the facts and data are not otherwise admissible, the attorney will need to decide whether to seek their admission under FRE 704.

BARRING THE TESTIMONY OF PROFFERED EXPERTS

FRE 104 directs the judge to make the preliminary inquiry into whether a “witness is qualified.” Based on the information provided by the proffered expert witness on the stand, the judge determines whether that witness can be qualified as an expert. In some cases, an opposing attorney may attempt to prevent a proffered expert from testifying on the grounds that the expert’s technique or method is invalid. When faced with such challenges, the majority of courts historically employed the *Frye* test, which directed the judge to admit only testimony that was based on techniques or methods “sufficiently established to have gained general acceptance in the particular field in which it belongs” (*Frye v. United States*, 1923). The *Frye* test was criticized on several grounds, including (a) there was no relationship between the FRE and the *Frye* test, (b) there was a lack of clarity as to what constituted “general acceptance,” and (c) the judge essentially abdicated the evidentiary gatekeeping role to the proffered expert’s field (insofar as the judge’s decision whether to allow the proffered testimony was based solely on whether that “field” generally relied on the techniques employed by the proffered expert).

Seventy years after the *Frye* decision, in *Daubert v. Merrell Dow Pharmaceuticals* (1993), the United States Supreme Court held that the *Frye* test was inconsistent with the FRE. In *Daubert*, the Supreme Court made clear that judges were not to abdicate their evidentiary gatekeeping role and should consider a number of factors when deciding on the admissibility of challenged expert testimony including, but not limited to, (a) whether the theory or technique upon which the expert’s opinion is based can be (and has been) tested, (b) whether the theory or technique has been subjected to peer review and publication, (c) the known or potential error rate of the technique or theory, and (d) whether the theory or technique has been accepted by the relevant scientific community. Six years later, in *Kumho Tire Co. v. Carmichael* (1999), the Supreme Court held that its ruling in *Daubert* applied not only to scientifically based expert testimony but to technically based testimony as well.

Although there has been a flood of legal scholarship speculating about the impact of *Daubert* and its progeny, there is also an emerging empirical literature that considers the effects of these cases on the presentation of expert testimony. Experts should be knowledgeable about the factors judges consider when ruling on the admissibility of challenged testimony and take them into account as they consider their methodologies and techniques.

Dixon and Gill (2002) assessed the impact of the Supreme Court's ruling in *Daubert* on the admission of expert testimony in civil cases filed in the federal courts in the decade following the decision. Challenges to proffered expert testimony dropped in the years preceding the decision, then rose in the three years after, only to decrease after 1997. Dixon and Gill suggested that, in the wake of *Daubert*, attorneys initially were more careful in presenting experts but then adapted to the courts' interpretations and returned to using experts as they had prior to *Daubert*. They also concluded that judges, while initially evaluating challenged expert testimony using the factors enunciated by the Court in *Daubert*, began using an increasing number and diversity of factors. Although one can never know whether any of the challenges and resulting decisions resulted in more just outcomes for the litigants, there may be value simply in raising courts' awareness about how to assess the value of experts and the soundness of the science on which they base their opinions. Dixon and Gill identified 17 factors utilized by these federal courts when considering the admissibility of challenged expert testimony (see Table 22.1).

Groscup, Penrod, Studebaker, and Huss (2002) analyzed 693 appellate cases in which proffered expert testimony was challenged by one party. Half of these cases were decided in the 5 years preceding the Supreme Court's decision in

Table 22.1
Factors Judges Consider When Ruling on the Admissibility of Proffered
Expert Testimony

-
1. Whether the method or technique on which the expert is relying has been subjected to testing
 2. Whether the method or technique on which the expert is relying has been subject to peer review and acceptance
 3. The error rates, if knowable and known, of the method or technique on which the expert is relying
 4. Whether the method or technique on which the expert is relying has received general acceptance within the expert's field of expertise
 5. The standards or controls in use in any measurements utilized by the expert
 6. The clarity and coherence of the expert's presentation
 7. Whether the expert used proper extrapolation (generalization)
 8. The breadth of the supporting data relied on by the expert
 9. Whether the data on which the expert has relied have been verified or can be verified
 10. The control or consideration of confounding variables in the expert's analysis
 11. Whether the expert relied on data or facts that are standard for the discipline
 12. The consistency of the expert's findings with those of other studies and other experts
 13. Whether the expert's findings have statistical significance
 14. Whether real-world data are consistent with the findings of the expert
 15. Whether the expert's opinion is in agreement with the opinion of any court-appointed expert
 16. The purpose for which research cited by the expert was conducted
 17. The reputation of the expert
-

Daubert and half were decided in the 5 years subsequent to the 1993 decision. Although the investigators found no differences in the relative frequency with which challenged expert testimony was admitted, they did find that courts were more likely to admit testimony from technical experts than from scientific experts, and they were also more likely to admit the testimony of prosecution experts than of defense experts. Groscup et al. reported that, in cases decided after 1993, the appellate courts did reference and apparently rely on the factors identified in the *Daubert* decision. In descending order, the courts focused most heavily on whether the proffered testimony would assist the trier of fact, followed by the expert's qualifications, the relevance of the testimony, and the degree to which the testimony could be construed as prejudicial. With respect to the expert's qualifications, experience and education were the two attributes most heavily weighted by the courts. These differences were even more evident in cases calling for technical rather than other types (scientific, medical/mental health, and business) of expert testimony. The investigators interpreted these findings as indicating that the courts mistakenly turned these inquiries into investigations of the proffered experts' qualifications rather than analyses of the soundness of the methods they employed. This finding may be a function of judges' limited training in and understanding of the scientific method.

Consistent with these findings are those reported by Kovera and her colleagues. In one study (Kovera & McAuliff, 2000), 144 Florida circuit court judges reviewed a hypothetical civil case in which psychological expert testimony was proffered and challenged. Despite the fact that "very few judges mentioned the internal validity threats present in the scientific research they evaluated," judges perceived themselves and attorneys as more capable of considering the soundness of the research than jurors (Kovera & McCauliff, 2000, p. 581). In another analog study employing a similar methodology, Kovera et al. found that attorneys were no better than their judicial counterparts at evaluating the quality of the research presented. This finding raises questions about attorneys' abilities as consumers of expert testimony and their ability to identify and challenge proffered expert testimony based on shaky methods or techniques (Kovera, Russano, & McAuliff, 2002). In order to provide some guidance to judges and attorneys when considering these matters, the Federal Judicial Center publishes the *Reference Manual on Scientific Evidence* (Committee on the Development of the Third Edition of the *Reference Manual on Scientific Evidence*, 2011).

Finally, Krafka, Dunn, Johnson, Cecil, and Miletich (2002) surveyed judges in 1991 and again in 1998, and later surveyed attorneys in 1999, all on issues surrounding admissibility of expert testimony. In the 1991 survey, judges reported rejecting expert testimony most often on the grounds that it did not assist the trier of fact (40%). The second most frequently cited rationale was concerns about the experts' qualifications (36%). Similarly, judges who responded to the 1998 survey identified reasons they limited or excluded expert testimony as including that (a) the evidence

was not relevant (47%), (b) the witness was not qualified (42%), and (c) the proffered testimony would not assist the trier of fact (36%). Krafka et al. reported that both judges and attorneys identified two problems in need of attention: experts becoming advocates for the parties who hired them and the expense associated with retaining experts.

PRESENTING ONE'S WORK AND OPINIONS THROUGH SWORN TESTIMONY

Regardless of the venue in which he or she offers expert testimony, the witness will not be persuasive unless the witness impresses the audience as credible and truly expert, and the witness is able to hold the audience's attention.

GENERAL CONSIDERATIONS

There are a number of rules that an expert should follow when offering sworn testimony, at either a deposition or a trial. Perhaps most important are that the expert pay careful attention to the question being asked, answer only the question asked, and, before answering, provide the nonquestioning attorney with a brief opportunity to object. The expert should seek clarification whenever he or she does not understand the question and indicate when he or she needs to explain an answer.

Foremost, the expert should understand his or her role as an educator, helping the legal decision maker understand a complicated matter so it can make a more informed, and presumably better, decision. Thus, the expert witness's goal should be to communicate what he or she did, learned, and concluded—all using language and concepts that the decision maker can understand. Depending on the case, the decision maker may be a judge or a jury. The expert witness should tailor his or her testimony to account for the difference in experience and expertise between judges, who regularly hear cases, and jurors, for whom this might be a first-time experience.

COMPONENTS OF EFFECTIVE COMMUNICATION

Put most simply, the experts' goals when testifying are to effectively communicate who they are and why the decision maker should listen to them, what they did, what they uncovered, and the opinions they formed. We address each of these issues in order.

For an expert witness to be persuasive, the legal decision maker must perceive the witness as expert, credible, trustworthy, and dynamic. Each of these characteristics is necessary but not alone sufficient (Bush, Connell, & Denney, 2006; Heilbrun, 1991). The expert's testimony will have no impact if the legal decision maker believes that the expert does not have information, insights, understanding, and

opinions about the matters in dispute that the decision maker lacks. In addition, the expert can be persuasive only if the legal decision maker perceives the expert as credible and trustworthy (Babitsky & Mangraviti, 2003). Finally, even if the witness is perceived as knowledgeable and credible, the legal decision will not be persuaded if the expert cannot gain and sustain the decision maker's attention.

COMMUNICATING EXPERTISE

Although it should be demonstrated throughout the expert's time in front of the legal decision maker, the witness's expertise is first communicated during qualification. It is during this examination, conducted by the attorney who has called the witness to testify, that there is a specific focus on the expert's education, experience, knowledge, skill, training, and techniques or methods. It also is during this time that the judge determines whether the witness satisfies the criteria posited by the FRE for testifying as an expert.

The expert should lay out all relevant educational and professional credentials to the decision maker, being neither too brief nor too detailed (Brodsky, 2012). These credentials include formal training (undergraduate and graduate degrees, postdoctoral training), specific forensic training, relevant continuing education, licensure, and board certifications. Also to be reviewed is professional experience involving matters relevant to the issue at hand (see *Testifying in Court*, further on, for more detailed discussion of these matters). Although considered broadly, professional experience laid out for the decision maker should be that most relevant to what is before the court. Thus, experience conducting evaluations or teaching about matters similar to that about which the expert is testifying is certainly important, whereas experience in clinical or other activities far removed from the topic in the case at hand can be ignored or minimized. Also important, however, is more general relevant experience. Thus, in a case involving assessment of the testamentary capacity of an aging will writer, the psychologist should emphasize her work with elderly populations more generally and spend less time discussing her experience counseling families in distress. However, if testifying in a custody proceeding, the psychologist should highlight experience counseling families in distress and training other professionals how to do this type of work.

Although it is often neglected during qualification, a discussion of the expert's techniques or methods is integral to communicating that the information gathered and opinions offered by the witness are truly expert. Toward this end, the witness, with the assistance of counsel, should clearly explain the techniques employed and how they reflect a sound approach to the task at hand. The expert can communicate the soundness of his or her techniques in a number of ways, including by describing their underlying rationale and science (in a way that is understandable to the decision maker) or explaining how the technique or method employed represents professional consensus.

COMMUNICATING CREDIBILITY AND TRUSTWORTHINESS

Credibility is key to persuasive communication. Analog research by Boccacinni and Brodsky (2002) indicates that jurors make judgments about the credibility of expert witnesses based on commonsense factors, such as how much they are being paid, how they spend their professional time, and by whom they have been retained. Jurors question the trustworthiness of expert witnesses who are extremely highly paid, who spend most of their time in the courtroom (as opposed to engagement in other pursuits, such as teaching or providing treatment), and who have a history of being retained only by one type of litigant (i.e., defense, plaintiff, or prosecution). These perceptions should be considered during testimony, with questions and answers tailored accordingly.

Ideally, experts will communicate their trustworthiness and credibility as they discuss their work more generally. Referencing one's obligation to be objective and impartial—and, more important, documenting what one did that reflects such objectivity and impartiality—is critical. Two methods of demonstrating one's objectivity or impartiality are (a) explaining how one sought and considered data from *all* potentially relevant sources, not just those that would be expected to provide information supportive of the retaining attorney's argument, and (b) searching for and referencing (in one's report and/or testimony) data that disconfirm and are inconsistent with one's opinions or that are not helpful to the retaining attorney's argument.

COMMUNICATING EFFECTIVELY—DYNAMISM

Dynamism refers to the expert's ability to gain and keep the audience's attention. As noted, no matter how expert and credible an expert witness is, the decision maker will not be persuaded if he or she is not listening.

An expert can maximize the audience's interest in and attention to his or her testimony by taking a number of basic steps, including (a) making eye contact, (b) modulating or varying the tone of voice, (c) employing understandable terms and concepts, and (d) using gestures appropriately and sparingly to reinforce important points. Particularly important is striking the right balance between being too detailed and not detailed enough. The expert must appreciate that, because he or she is only one of many witnesses who will testify, the legal decision maker's attention span will be understandably time limited. Although the witness should take as much time as necessary to explain the most important aspects of what he or she did, learned, and concluded, the expert should remember that everyone, including judges and jurors, will begin to lose interest at some point. Providing just enough detail, but not providing too much, is the challenge.

Integral to considering just how specific one must be when testifying is knowing whether the decision maker has been provided with and read the expert's report (if one has been prepared and provided). Practices vary across jurisdictions and as a function of the type of proceeding. Judges generally will have been provided

with a copy of the expert's report prior to hearing the expert's testimony (e.g., in competency, custody, and guardianship proceedings). Juries, on those rare occasions when they are provided with the report, will receive it after the expert testifies. Certainly, in cases in which the judge is the decision maker and the expert is confident that he or she read the report prior to the proceeding, the expert can provide fewer details and focus testimony on the most important data and opinions. In those cases in which the decision maker has not reviewed a report, the expert should consider providing a greater level of detail when describing what he or she did, learned, and concluded, all with an eye toward relevance and persuasiveness.

SETTINGS IN WHICH TESTIMONY IS OFFERED AND ASSOCIATED PROCESS AND CONSIDERATIONS

Sworn testimony is offered in two primary venues: in depositions and in the courtroom. Because these venues vary in important ways with respect to their purposes and processes, they are discussed separately.

Depositions. Depositions are a form of pretrial discovery that allow an attorney to gather sworn testimony from a witness in anticipation of and prior to trial. Rules of procedure vary dramatically regarding the types of cases in which depositions are allowed. For example, the vast majority of states prohibit depositions of witnesses who are to testify in criminal proceedings, while a small number of jurisdictions allow them.

Although depositions sometimes are taken so that the witness's testimony can be offered in lieu of an appearance at trial (e.g., if the witness will be unavailable for trial), in most cases depositions are conducted with the expectation that the witness will also testify at trial. Expert witnesses appearing at depositions typically are deposed by the attorney who has not retained them. However, in cases in which the expert has been retained by the court, the expert is most likely to be deposed by the attorney whose legal argument has not been supported in the report authored by the expert. Although the retaining attorney almost always will be present at his or her expert's deposition to raise objections and ask clarifying questions, the attorney who retained the expert, if he or she asks any questions at all, will likely limit them to those that clarify issues raised by the other attorney (absent the issue of unavailability raised earlier), since the expert and retaining attorney are free to meet and confer at any time.

Depositions of experts generally have three functions: information gathering, opinion commitment, and assessment of intangible factors. Depositions are most valuable insofar as they allow the attorney to learn what the expert did, what facts or data the expert was provided, what facts or data the expert gathered or generated, and what opinions the expert formed. The deposition provides a mechanism by which the attorney can gauge the nature and impact of the expert's testimony.

Moreover, if the expert offers testimony at trial that differs from that which was offered at the deposition, the opposing attorney can impeach the witness with the deposition transcript. Finally, insofar as the deposition process approximates what occurs in court, the attorney can draw some conclusions about the expert's effectiveness as a trial witness.

Aspects of the deposition are often more informal than court proceedings (e.g., the location of the deposition, the dress and style of the attorneys), and thus the expert may be tempted to adopt a more casual approach when being deposed. However, the expert should conduct him- or herself as if appearing in a courtroom, since the deposition is part of the legal process and written, audio, or video records of the deposition can be referenced during the trial. Also, the expert conveys confidence and expertise through his or her demeanor and deportment and does not want to appear to be anything less than fully professional.

As indicated, a deposition typically is conducted at the request of the nonretaining attorney and provides an opportunity for that attorney to learn what the expert was provided, did, and concluded. Because most of the questions are presented by the nonretaining attorney, a deposition can be conceptualized as a cross-examination that is not preceded by a direct examination. The testifying expert should rest assured, however, that the retaining attorney (or the other attorney in cases in which one has been court retained) has the opportunity to ask questions so as to clarify matters or issues of concern that are identified during the substance of the deposition. The retaining attorney may or may not question the expert witness, based on a variety of strategic factors, but likely will let the expert know of his or her strategy prior to the deposition.

The general rule for responding to questions at a deposition is quite simple. After providing any other attorneys present an opportunity to object to the question, the expert should simply answer the question that is asked and expand on that answer only if he or she fears that failure to do so would be misleading in some significant way. However, there may be some cases in which the retaining attorney may wish the expert to go further in his or her responses. For example, in some circumstances, the attorney may believe there is an advantage in presenting a powerful case in the deposition and may want the witness to be quite detailed in responding to questions. Thus, while the witness should produce all legitimately requested evidence and honestly answer all questions, the degree of detail or expansiveness offered can vary as a function of the retaining attorney's strategy. It is thus important that the expert consult with the retaining attorney before the deposition so the attorney can provide any important strategic or background information.

At the end of the deposition, the deposing attorney or the court reporter will generally ask the witness, "Would you like to read or waive (signature)?" This question allows the deposed witness to request a transcript of the proceeding, check it for accuracy, make any corrections, and sign and return it. Corrections are typically of three types: (a) errors made by the court reporter (e.g., spelling Wechsler Wexler, referring to the Hair Psychopathology Checklist-Revised); (b) small errors

that the expert might have made (e.g., misreporting the examinee's date of birth, incorrectly reporting that one is no longer licensed in a particular state); and (c) more substantive errors (e.g., misinterpreting a test profile, misreporting an examinee's IQ score). Whereas all jurisdictions allow deponents to correct errors made by the court reporter, whether the deponent can correct his or her own errors is a matter of procedure within the jurisdiction. If the rules of the jurisdiction do not allow the expert to correct substantive errors in this way, the expert can always provide a letter or memorandum to all counsel indicating the misstatements in the transcript. In any case, it is generally agreed that the expert should request and review the deposition transcript to prevent the uncorrected transcript from being used for impeachment at trial, ensure that the record in the case is accurate, and generate a copy of the deposition that the expert can review prior to trial.

TRIALS

Sworn testimony is offered at both trials and depositions, but the process and formats are quite different. Although there may be some variation depending on the jurisdiction and type of proceeding, the order of presentation at trial is provided in Table 22.2.

PRIOR TO TESTIFYING IN COURT

In those cases in which the expert was retained by an attorney (as distinguished from those cases in which the expert is retained by the court), the expert and the attorney should meet for a pretrial conference during which the attorney can (a) review any pretrial rulings that may affect the expert's testimony, (b) educate the expert about the styles of the judge and other attorneys, (c) reiterate case theories and share perceptions of the trial process as it has unfolded, and (d) review the nature and structure of the expert's anticipated testimony (Babitsky & Mangraviti, 2003; Brodsky, 1991; Mart, 2006). The pretrial meeting should include a review and discussion of the strengths and weaknesses of the expert's opinions, direct examination questions (crafted by the attorney, perhaps with the expert's assistance), and likely answers (provided by the expert) as well as discussion of questions that may be asked by opposing counsel (suggested by the attorney) and effective responses (crafted by the expert).

Before trial, the expert witness should review all primary materials about which he or she reasonably anticipates being questioned. The more material the psychologist can commit to memory, the more impressive he or she will appear. Preparedness, competence, and authority can be communicated to attorneys, judge, and jury by a good command of relevant case facts. But even with a strong understanding of the case matters, experts should bring to the witness stand all case materials they generated or reviewed (i.e., notes, test protocols and profiles, reports, memoranda, pleadings, depositions, records) with the understanding that they (and the

Table 22.2
Trial Process

Opening Statements

- Opening statement by the prosecutor (criminal case) or plaintiff's attorney (civil case)
- Opening statement by defense attorney

Prosecutor's or Plaintiff's Attorney's Case in Chief

- Direct testimony of each witness called by the prosecutor or plaintiff's attorney (including expert witnesses)
- Immediately after direct examination of a witness, cross-examination of that witness by the defense attorney (including expert witnesses)
- Immediately after cross-examination of a witness by the defense attorney, redirect examination of that witness by the prosecutor or plaintiff's attorney, if deemed necessary by the attorney (including expert witnesses)
- If allowed by the court and immediately after redirect examination, re-cross-examination of a witness by the prosecutor or plaintiff's attorney (including expert witnesses)

Defense Attorney's Case in Chief

- Direct testimony of each witness called by the defense attorney (including expert witnesses)
- Immediately after direct examination of a witness, cross-examination of that witness by the prosecutor or plaintiff's attorney (including expert witnesses)
- Immediately after cross-examination of a witness, redirect examination by the defense attorney, if deemed necessary by the attorney (including expert witnesses)
- If allowed by the court and immediately after redirect examination, re-cross-examination of that witness by the prosecutor or plaintiff's attorney (including expert witnesses)

Rebuttal Evidence by the Plaintiff's Attorney or Prosecutor

- Presentation of rebuttal evidence by the prosecutor or plaintiff's attorney (if deemed necessary by the attorney) (Rebuttal evidence is presented in the same manner as the prosecutor's or plaintiff's case in chief.)

Closing Arguments

- Closing argument by the prosecutor or plaintiff's attorney
- Closing argument by the defense attorney
- Rebuttal argument by the prosecutor or plaintiff's attorney

Jury Instructions (if a jury trial)

examining attorneys) can access them during testimony. These materials should be indexed in a way that allows easy access to critical data. The greater the volume of case materials, the greater the need to index. Seconds seem like minutes and minutes seem like hours to the expert who is desperately rummaging through a case file to locate and review interview notes or a test protocol that is the focus of questioning.

Some anxiety may be allayed, and the witness may be more comfortable, if, sometime before the trial, he or she visits the courthouse and courtroom—particularly

when testifying in a new or unknown venue (Blau, 1998; Brodsky, 1999; Mart, 2006). There can be value in struggling with mundane matters, ranging from the availability of courthouse parking to negotiating the security gauntlet, days before testimony rather than on the day of the testimony. In addition, the expert can discover important information about the logistics of the courtroom, such as lighting, temperature and ventilation, and the locations of the various legal actors (i.e., judge, jury, counsel, witness box). If the expert goes to the courtroom when other proceedings are taking place, he or she may also discern the flow of the process and feel more comfortable knowing how documents are routinely admitted and where the lawyers stand when asking questions. He or she may also get a feel for the demeanor and temperament—and any peculiar preferences—of the judge.

TESTIFYING IN COURT

The purpose of trial testimony is much more limited than that of deposition testimony. At trial, the expert simply needs to communicate to the legal decision maker (i.e., judge or jury) what he or she did, what he or she learned, and what opinions he or she formed that are relevant to the legal issues in dispute. In depositions, attorneys typically ask questions of witnesses to learn what they will say. At trial, attorneys believe they know what the witness will say, having either conferred with the witness or questioned the witness during a deposition. Trials are more formal than depositions with respect to process, procedure, and dress. Too much has been written about how experts should dress for court, and that which has been written often reflects nothing more the personal preferences and style of the writer. The true rule is actually quite simple. Although Brodsky (1991) suggested that experts dress in a way that makes them comfortable and cited his own practice of appearing in court without a tie, we recommend that witnesses be sensitive to local customs and convention and dress as attorneys do.

We recommend that, to the extent possible, witnesses bring to court and the witness stand any and all materials they considered during their involvement in the litigation (e.g., attorney correspondence, billing records, interview notes, test data, affidavits, depositions, collateral documents). If there are any questions about what the expert did or what the expert relied on, the data are available for reference and review. That being said, witnesses who can testify without referring to their materials, or who have organized their materials in a way that facilitates a quick fact check, are generally more persuasive. Those witnesses appear more credible, more prepared, and more knowledgeable, and they are less likely to lose the decision maker's attention while shuffling through their records.

Before trial, the expert should ascertain whether he or she is expected to be present for the testimony of others or is barred from such presence. FRE 615 provides for exclusion of witnesses upon request of either party:

At a party's request, the court must order witnesses excluded so that they cannot hear other witnesses' testimony. Or the court may do so on its own. But this rule does not

authorize excluding: (a) a party who is a natural person; (b) an officer or employee of a party which is not a natural person after being designated as the party's representative by its attorney; (c) a person whose presence a party shows to be essential to presenting the party's claim or defense; or (d) a person authorized by statute to be present.

In many, if not most, cases, the retaining attorney will request that the expert be allowed to remain in the courtroom during the trial on the grounds that the expert's presence is essential to the presentation of the party's case. This is because the attorney who retained the expert often will confer with him or her during the trial, especially around matters involving the testimony of the other party's expert(s). Although courts generally allow the expert to remain in the courtroom, the decision is made on a case-by-case basis and is dependent on the nature of the matter and the importance of expert testimony to the determination of the issues before the court. If the expert is not allowed in the courtroom, the retaining attorney is prohibited from discussing any testimony in the case with the expert—from the beginning of the trial until after the expert has testified.

The expert will be called to testify by the retaining attorney or the attorney who executed a subpoena mandating the appearance. The judge decides whether a witness is "qualified" under the rules of evidence and bases his or her decision on the testimony presented by the proffered expert during the qualification hearing. During this process, the attorney, through question and answer with the witness, educates the judge about the proffered witness's background and qualifications. Because a witness cannot testify as an expert unless and until the judge has ruled that he or she meets such requirements, the attorney must make clear to the judge through carefully selected questions the proffered expert's relevant experience, training, education, skill, and knowledge. This is no time for false modesty. During this process the attorney typically submits a copy of the witness's *curriculum vitae*, so it is important that it is up to date.

As noted, educational credentials to be referenced during the qualification process include formal training (undergraduate and graduate degrees, postdoctoral training), relevant continuing education, licensure, and legitimate board certifications (see Bartos, 2012; Dattilio, 2002; Dattilio & Sadoff, 2002; Golding, 1991; and Packard & Reyes, 2003, for problems associated with illegitimate credentials). The witness should be sure to emphasize any focused forensic training, particularly forensic training relevant to the matters about which he or she is testifying. Finally, it is important to discuss relevant professional experience, citing such things as the number of similar cases in which one has been involved or clinical experience dealing with matters that are at issue in the case.

Although considered broadly, experience laid out for the decision maker should be that most relevant to the matters before the court. While the expert should mention general relevant experience, he or she should focus on that which is most relevant to the proceeding (i.e., experience conducting evaluations similar to that about which the expert is testifying).

The expert probably knows best what aspects of his or her education, training, and experience are most relevant to the matters at issue. Although many attorneys, upon reviewing the expert's *curriculum vitae*, can competently assemble a list of questions that best emphasize such matters, this is not always the case. Indeed, the experience of one of our colleagues (a highly accomplished forensic psychologist who was asked a series of questions about his attendance at medical school, psychiatry residency, and licensure as a physician by a naive and misguided attorney who had retained him) reveals that not all attorneys are up to this task.

Even when the attorney is exceedingly competent, it is important for the witness to remember that he or she is the expert in the field in which he or she is testifying, and the attorney has retained the expert in part to assist in understanding the witness's area of expertise. One of the ways the witness can assist the attorney is to discuss the relative importance of various aspects of the expert's education, training, and experience and highlight those that are most relevant to the matters in dispute. In fact, the expert may be in the best position to identify the questions that focus the court on the important and relevant aspects of his or her qualifications. These questions, of course, will vary as a function of the issues in dispute in the case. Table 22.3 is a list of questions that the first author of this chapter provided an attorney who sought to qualify him as an expert witness in a guardianship proceeding.

It would be folly for an expert witness to fail to discuss his or her techniques or methods during the qualification process. Failure to establish the techniques or methods employed can result in the court's refusal to allow the expert to testify. The judge needs to be apprised of the techniques and methods so that he or she can determine that they are "generally accepted" (in those jurisdictions that use some version of the *Frye* rule) or are reliable and valid (in those jurisdictions that use some version of the *Daubert* rule). Discussing the (appropriate) techniques employed can also increase the expert's persuasiveness with the legal decision maker (judge or jury). Such a discussion communicates that the expert's opinions are based on a sound methodology. Thus, the expert, with the assistance of counsel, should explain the methods employed and how they reflect a sound approach to the task the expert was called to perform. This can be communicated in a number of ways, including a description of the rationale and science of the method/technique or a description of how the technique/method employed represents professional consensus about doing what was done.

Opposing counsel, of course, has the opportunity to query the proffered witness during the qualification process. These cross-examination questions are likely to focus on what the witness is *not* and what the witness did *not* do. Indeed, a tone adopted by some cross-examining attorneys is that *whoever* the expert is and *whatever* the expert did is irrelevant, while whatever the expert is *not* and whatever the expert did *not* do is critical. If such an approach is anticipated by the expert and retaining attorney, they should consider addressing on direct examination what the expert did not do and who the expert is not. As part of this discussion, the attorney

Table 22.3
Sample Qualification Colloquy—Guardianship Proceeding

-
- What is your name and business address?
 - What is your occupation and profession?
 - Where and in what capacities are you currently employed?
 - How long have you practiced as a psychologist?
 - What are your primary responsibilities in your current position(s)?
 - Do you hold any licenses or advanced certifications?
 - Could you tell us a little about your formal training, including internship and fellowship training?
 - Are you a member of, or active in, any professional organizations?
 - Have you published any books or articles? In what areas?
 - Do you serve on the editorial boards of any journals?
 - Have you received any awards or honors for your career achievements?
 - What kinds of evaluations do you conduct for the courts?
 - What percentage of your time is devoted to evaluating persons in legal proceedings?
 - What percentage of your time is devoted to treatment activity?
 - What percentage of your time is devoted to research and teaching activities?
 - In what states and jurisdictions have you been qualified as an expert witness?
 - When offered to the court as an expert witness by a party in a legal proceeding, have you ever been denied?
 - Have you evaluated persons subject to guardianship proceedings? How many and under what circumstances?
 - Are there other, similar types of evaluations that you have completed? How many and in what circumstances?
 - What kind of training have you completed with respect to evaluating persons subject to guardianship proceedings?
 - In this context, approximately how many persons have you evaluated and for whom?
 - Have you conducted training on guardianship evaluations? In what circumstances?
-

and expert can focus on why these omissions are irrelevant to the issues in the case. Such a tactic can take the sting out of cross-examination and can place the seeming omissions into context.

Opposing counsel may sometimes stipulate to the witness's expertise and indicate to the court that a detailed review of the expert's education, training, and experience is unnecessary. In those cases, there need not be a qualifying process in front of the judge; instead, the expert witness is called to the stand and begins his or her direct testimony. Much of the time, such a course of action by opposing counsel is valid on its face; that is, the attorney recognizes the proffered witness's expertise and considers challenging such to be a waste of time and of little value. In some cases, however, opposing counsel's acknowledgment of the expert's qualifications may be strategic rather than a matter of professional courtesy. Thus, opposing counsel may seek to avoid this component of qualification because he or she fears that the proffered witness's credentials eclipse those of his or her own expert. In all cases in which the other side stipulates to the expertise of the witness, the attorney who

is proffering the expert can recognize opposing counsel's acknowledgment of the witness's expertise but should still take the expert through the process in order to communicate the high level of the witness's expertise. The witness should provide as much information during this process as he or she would during the qualification process. Even if the witness is permitted to testify as an expert and provide an opinion, the decision maker still must decide how much weight to give to the opinion. The witness's credentials are critical to the weight the decision maker gives the witness's testimony.

DIRECT EXAMINATION STRATEGIES AND APPROACHES

The retaining attorney (or the attorney who issued a subpoena to appear) will conduct direct examination. During direct examination, all questions must be nonleading (i.e., questions that do not suggest an answer). The expert should be competent, articulate, and knowledgeable but never condescending, paternalistic, or supercilious. The expert's role is to explain complex issues and concepts in ways that are easily understood by people who are not familiar with the expert's area of specialization. In order to do this, the expert needs, above all, to be credible.

As previously noted, the expert and retaining attorney should have had a pretrial conference during which they develop a strategy and outline for direct testimony. Indeed, the expert should know ahead of time what questions the attorney will ask so that he or she can prepare appropriate responses. There is no excuse for being ill-prepared for the trial. It is crucial, however, that the expert's answers not be (or appear to be) scripted or rehearsed. The expert should listen carefully to the questions and answer clearly. He or she should make regular eye contact with the jury or, in the case of a bench trial, the judge, if possible. The most effective witnesses are those who tell a story in a way that is coherent, easily followed, and expressed in lay terms. McElhane (1997) advised avoiding words like *elucidate*, *illuminate*, *discern*, *explicate*, and *expound*. Instead, the expert should use words like *teach*, *tell*, *explain*, *help us understand*, *show us*, and *untangle*. Demonstrative words, such as *show*, *see*, *watch*, *picture*, *view*, and *look at* will draw the listener to the expert's story. At the same time, although the expert should speak plainly, he or she should avoid being too colloquial. Some use of professional jargon, with appropriate explanations, may be appropriate. After all, the expert is in the courtroom because he or she has specialized knowledge to share with the decision makers. However, the witness should avoid too much jargon and should especially avoid using jargon that is not self-explanatory or that the expert fails to explain.

Often it is helpful to use analogies as a way of communicating complicated findings, using a context with which the decision maker is familiar (Brodsky, 1991). For example, one expert was testifying about a jail suicide. The victim was an intoxicated 25-year-old male with marital troubles and a few prior arrests. The expert was confronted with the fact that most jail suicides fit the victim's pattern. The expert referred to a public service advertisement popular at that time that

stated that the vast majority of traffic accidents occur within 25 miles of one's home. However, what the ad did not say was that most driving occurs within 25 miles of one's home. The same faulty reasoning was being applied to the jail inmate; that is, the majority of people in jail fit the decedant's profile. Thus, there was no profile that the jailers could use to isolate this inmate as being particularly prone to suicide. As the ad was mentioned, the jurors' heads bobbed in recognition, they carefully listened and understood the underlying concepts of base rates, and they rejected the plaintiff's claims. Although a good analogy is often helpful, an expert should avoid relying on an analogy without first consulting with the retaining attorney. The expert needs to make sure that the analogy could not be used to undercut the attorney's theory of the case or the legal arguments being made.

It is important for the expert to understand the legal theories underlying the case being presented by the retaining attorney. Doing so allows the expert to gear his or her testimony to the issues in the case and the elements that need to be proven. It is not necessary, however, for the witness to be an expert on the law. As a friend of the second author has stated, "The witness should relax into the lawyer." In other words, the expert should rely on the lawyer to object when necessary, to respond to objections, and to make the legal arguments. Most important, when a witness is testifying and an objection is made, the witness should stop talking immediately and wait for the judge to make a ruling. Only after the lawyers have finished arguing, and the judge has decided whether to admit the evidence, should the witness speak. If unsure of the ruling and its effect on the question being asked, the witness should ask for the question to be restated.

Trials always involve serious matters. The litigants are highly invested in the issues at hand, as are others, including the attorneys, the litigants' family members, and the jurors whose routines have been disrupted. Few people are as funny as they think they are. Although a successful attempt at humor can reduce tension and might make the witness more likable, an unsuccessful attempt can cause the decision maker to perceive the witness as irreverent, disrespectful, or lacking the necessary seriousness for the task at hand (Brodsky, 2012). Most important, any use of humor—whether successful or unsuccessful—may cause the decision maker to question the expert's credibility and trustworthiness. Consequently, witnesses generally should avoid any attempts at humor (Blau, 1998).

CROSS-EXAMINATION STRATEGIES AND APPROACHES

Once the retaining attorney has completed direct examination, any opposing attorneys can conduct their cross-examination. During cross-examination, the attorney typically employs leading questions (i.e., questions that suggest the answer). The stereotypical leading question begins "Isn't it true that . . . ?" Indeed, a good cross-examiner asks *only* leading questions. Although the format of direct examination allows the witness to speak at length about all he or she did, learned, and concluded, cross-examination provides the witness with much less latitude to respond and is

designed to identify what the witness failed to do and the resulting limitations of what was learned or concluded.

The retaining attorney will likely have prepared the expert for direct testimony and should have identified some cross-examination questions that likely would be asked. Nonetheless, the expert cannot know with certainty what questions will be asked during cross-examination. The expert, like any witness, should listen carefully to the question, provide retaining counsel an opportunity to object, and then answer only the question that is being asked. If a query is not clear, the witness should request that the questioning attorney rephrase the question. The expert should not do the lawyer's work—in other words, it is the lawyer's job to ask the questions and the witness's job to answer them. The witness should not elaborate and provide additional information, unless the witness feels that the additional information is essential to making his or her answer correct.

There is an odd kind of balance in cross-examination. If a lawyer asks a question that can be answered yes or no, the witness must and should answer either yes or no. At the same time, the witness has the right to explain his or her answer if the explanation is necessary to make the answer complete and accurate. Strategically, witnesses should try to answer yes/no questions in that way whenever possible and save resistance to such responses to the most important matters. One approach for responding to a question when a yes or no answer is problematic was offered by Brodsky (2012), who recommended that the witness begin the response with a dependent clause (e.g., "Although it is true that . . ."). A dependent clause suggests that a yes or no answer will ultimately follow, but it pulls for the expert to be allowed to preface the response with an explanation. Another approach is for the expert to say that a simple yes or no is likely to be misleading, and let the cross-examining attorney respond accordingly.

In addition to identifying limitations in the direct examination testimony, cross-examining lawyers often try to impeach the expert witness. Impeachment refers to the process of asking questions that attack the witness's credibility. Many types of impeachment are permitted, but a few are particularly relevant. A witness may be impeached with a prior inconsistent statement. Any statement, whether sworn or not, whether in writing or oral, may be used for impeachment. FRE 613a provides the procedure that must be followed when impeaching a witness with a prior inconsistent statement:

When examining a witness about the witness's prior statement, a party need not show it or disclose its contents to the witness. But the party must, on request, show it or disclose its contents to an adverse party's attorney.

Most typically, the prior statements used for impeachment will be sworn testimony offered in the deposition in the case at issue or in other cases. However, a witness may also be impeached through professional writings and professional or personal statements. Thus, when preparing to testify at trial, an expert should, at a

minimum, review the transcript of his or her deposition in the case. In addition, the cautious expert will also review transcripts of testimony offered in similar matters as well and any of his or her professional publications and statements that are relevant to matters in dispute. When a witness is asked a question about an answer provided in a deposition, the witness is entitled to examine the deposition before answering the question. Because opposing attorneys may attempt to mislead the witness with statements taken out of context or with partial statements, the witness should always make sure that the full statement and context are made clear to the decision makers.

FRE 609 allows impeachment of any witness based on the witness's conviction of a crime that (a) involved proof or admission of dishonesty or a false statement or (b) is punishable by more than one year of imprisonment (a felony). Accordingly, an expert who has been convicted of any offense should share that information with the attorney before being retained to allow the attorney to make an informed hiring decision.

A witness also can be impeached through bias or any personal characteristic that causes the witness to prefer one litigant over the other. If an expert routinely testifies for one side in cases (e.g., the witness always testifies for the defendant physician in malpractice cases), the witness can be questioned about this aspect of his or her professional life, as it may show a bias toward defendant physicians. Likewise, an expert can be asked about his or her fees. Such inquiries may not occur when both parties have retained experts who are charging similar fees. However, attorneys may make a point of fees and use them to suggest bias, especially when the expert retained by the other side is highly paid (or more highly paid than his or her own expert), and research indicates that juries do resonate to such issues (Boccacinni & Brodsky, 2002). As with all matters, the expert witness should not argue with the lawyer who is trying to establish bias in this manner. The expert should acknowledge payment for time expended on the case, answer related questions truthfully, and allow the retaining attorney to make the necessary legal arguments.

Expert witnesses can be impeached with "learned treatises" or authoritative texts in the witness's area of expertise. FRE 803(18) allows a statement contained in the treatise, periodical, or pamphlet to be introduced into evidence if

- (a) the statement is called to the attention of an expert witness on cross examination or relied on by the expert on direct examination; and (b) the publication is established as a reliable authority by the expert's admission or testimony, by another expert's testimony, or by judicial notice. If admitted, the statement may be read into evidence but not received as an exhibit.

The expert should be fully familiar with the most widely cited and authoritative publications in his or her specialty. If the expert takes a position at odds with one of these sources, the witness should be prepared to explain why he or she has rejected the theory or conclusions contained in the treatise or publication. The best-prepared

expert will also have cites to other authorities whose opinions agree with his or hers. A “learned treatise” can be used for impeachment regardless of whether it meets the criteria of FRE 803(18). If the statement in the publication is used only to impeach the witness, the statement is not introduced into evidence but is simply used to show that the witness’s statements and opinions lack credibility. However, the statements in the publication can be introduced into evidence under FRE 803(18) if the expert or any other expert in the proceedings (including, of course, the opposing party’s expert) testifies that the publication is a reliable authority. This difference between using the document solely for impeachment and introducing it as evidence is, again, one of the issues that should be left to the attorneys. The expert’s role is to know the relevant authoritative sources and be ready to testify about them.

During cross-examination, the attorney might misrepresent case facts or what the expert did or previously testified to. When confident that a substantive issue has been misrepresented in this way, the witness should point this out to counsel, provide an accurate presentation of the matter, and indicate a willingness or desire to refer to whatever record may be available that support his or her rendition (e.g., the report, the deposition or trial transcript, a collateral record that was reviewed). Such a course of action ensures that the decision maker is not misinformed and communicates to the attorney that the expert has a command of the case facts.

Overall, the best approach to cross-examination is fourfold. The expert should:

1. Be truthful.
2. Answer only the questions asked.
3. Explain only if it is essential to providing a correct answer.
4. Never get frustrated or angry.

An expert can do more harm to a case by losing his or her composure than almost any testimony can overcome.

REDIRECT AND RE-CROSS EXAMINATION

After cross-examination, the retaining attorney may question the witness in a process called redirect examination. Generally, the questions on redirect examination are limited to the issues raised during cross-examination. As is the case with direct examination, questions on redirect examination must be nonleading and generally focus on clarifying any issues raised during cross-examination. The attorney determines whether to conduct redirect examination based on an assessment of how effective the cross-examination was and whether there are matters that need to be clarified with follow-up questions.

If the retaining attorney conducts redirect examination, the opposing attorney may then ask questions on recross examination. Questions on recross examination are generally limited to issues raised during redirect examination. Of course, attorneys engaging in recross examination are permitted to use leading questions.

When the expert is finished testifying, he or she should calmly leave the witness stand. Even if the witness was not permitted to sit in the courtroom before testifying, he or she may be permitted to remain in the courtroom after testifying. Whether to remain or leave the courtroom, however, should be discussed with the retaining attorney. There are, of course, other gambits that cross-examining attorneys may employ. Because a detailed discussion of these additional tactics is beyond the scope of this chapter, the interested reader is directed to the work of others for further consideration (e.g., Babitsky & Mangraviti, 2003; Brodsky (1991, 1999, 2012; Matson, 1994; Merenbach & Stephen, 1993).

CONCLUSIONS

The expert's goal when testifying is to educate the decision maker about complicated matters and persuade the decision maker that the opinions offered have a solid foundation and are valid as a result. However, the decision maker will not be persuaded unless the expert is perceived as knowledgeable, credible/trustworthy, and dynamic. It is the witness who is knowledgeable about and/or skilled in his or her area of expertise, in relevant rules of legal procedure, in direct and cross-examination strategies, and in effective ways of communicating who will be persuasive and helpful to the legal decision maker.

REFERENCES

- Babitsky, S., & Mangraviti, J. J. (2003). *Cross-examination: The comprehensive guide for expert witnesses*. Falmouth, MA: SEAK.
- Bartos, L. (2012, April 17). No forensic background? No problem. *Propublica*. Retrieved from <http://www.propublica.org/article/no-forensic-background-no-problem>
- Blau, T. H. (1998). *The psychologist as expert witness* (2nd ed.). New York, NY: Wiley.
- Boccaccini, M. T., & Brodsky, S. L. (2002). Believability of expert and lay witnesses: Implications for trial consultation. *Professional Psychology: Research and Practice*, 33, 384–388.
- Brodsky, S. L. (1991). *Testifying in court: Guidelines and maxims for the expert*. Washington, DC: American Psychological Association.
- Brodsky, S. L. (1999). *The expert expert witness: More maxims and guidelines for testifying in court*. Washington, DC: American Psychological Association.
- Brodsky, S. L. (2012). *Testifying in court: Guidelines and maxims for the expert* (2nd ed.). Washington, DC: American Psychological Association.
- Bush, S. S., Connell, M. A., & Denney, R. L. (2006). *Ethical practice in forensic psychology: A systematic model for decision making*. Washington, DC: American Psychological Association.
- Committee on the Development of the Third Edition of the *Reference Manual on Scientific Evidence*. (2011) *Reference manual on scientific evidence* (3rd ed.). Washington, DC: Federal Judicial Center.
- Dattilio, F. M. (2002). Board certification: Is it really necessary? *Professional Psychology: Research and Practice*, 33, 54–57.
- Dattilio, F. M., & Sadoff, R. L. (2002). *Mental health experts: Roles and qualifications for court*. Mechanicsburg, PA: Pennsylvania Bar Institute Press.

- Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993).
- Dixon, L., & Gill, B. (2002). Changes in the standards for admitting expert evidence in federal civil cases since the Daubert decision. *Psychology, Public Policy, and Law*, 8, 251–308.
- Federal Rules of Evidence (2011). Washington, D.C.: Lexis-Nexis Group.
- Frye v. United States, 293 F.1013 (D.C. Cir. 1923).
- Golding, S. L. (1991, August). *The voir dire of forensic experts: Issues of qualification and training (Subtitle: Sheepskins for sale: Shortcut to slaughter?)*. Paper presented at the annual meeting of the American Psychological Association, Boston, MA.
- Groscup, J. L., Penrod, S. D., Studebaker, C. A., & Huss, M. T. (2002). The effect of Daubert on the admissibility of expert testimony in state and federal criminal cases. *Psychology, Public Policy, and Law*, 8, 339–372.
- Heilbrun, K. (1991). *Principles of forensic mental health assessment*. New York, NY: Kluwer/Plenum Press.
- Jenkins v. United States, 307 F.2d 637 (D.C. App. 1962).
- Kovera, M. B., & McAuliff, B. (2000). The effects of peer review and evidence quality on judge evaluations of psychological science: Are judges effective gatekeepers? *Journal of Applied Psychology*, 85, 574–586.
- Kovera, M. B., Russano, M. B., & McAuliff, B. D. (2002). Assessment of the commonsense psychology underlying Daubert: Legal decision-makers' ability to evaluate expert evidence in hostile work environment cases. *Psychology, Public Policy, and Law*, 8, 180–200.
- Krafka, C., Dunn, M. A., Johnson, M. T., Cecil, J., & Miletich, D. (2002). Judge and attorney experiences, practices, and concerns regarding expert testimony in federal civil trials. *Psychology, Public Policy, and Law*, 8, 309–332.
- Kumho Tire Co. v. Carmichael, 526 U.S., 119 S. Ct. 1167 (1999).
- Mart, E. G. (2006). *Getting started in forensic psychology practice: How to create a forensic specialty in your mental health practice*. Hoboken, NJ: Wiley.
- Matson, J. V. (1994). *Effective expert witnessing* (2nd ed.). Boca Raton, FL: Lewis.
- McElhaney, J. W. (1997, May). Terms of enlightenment: Articulate expert witnesses help jurors visualize facts. *American Bar Association Journal*, 82–83.
- Melton, G. B., Petrila, J., Poythress, N., Slobogin, C., Lyons, P., & Otto, R. K. (2007). *Psychological evaluations for the courts: A handbook for attorneys and mental health professionals* (3rd ed.). New York, NY: Guilford Press.
- Merenbach, D. G., & Stephen, A. (1993). *How to be an expert witness: Credibility in oral testimony*. Santa Barbara, CA: Fithian Press.
- Millon, T. (1969). *Modern psychopathology*. Philadelphia, PA: Saunders.
- Münsterberg, H. (1908). *On the witness stand*. New York, NY: Doubleday.
- Otto, R. K. (2012). Introduction. In R. K. Otto (Ed.), *Forensic psychology* Vol. 11 in I. B. Weiner (Ed.-in-Chief), *Handbook of psychology* (2nd ed.). Hoboken, NJ: Wiley.
- Otto, R. K., & Heilbrun, K. (2002). The practice of forensic psychology: A look toward the future in light of the past. *American Psychologist*, 57(1), 5–18.
- Packard, T., & Reyes, C. (2003). Specialty certification in professional psychology. In M. J. Prinstein & J. J. Patterson (Eds.), *The portable mentor: Expert guide to a successful career in psychology* (pp. 191–212). New York, NY: Kluwer Academic/Plenum Press.
- Slobogin, C. (1989). The "ultimate issue" issue. *Behavioral Sciences and the Law*, 7, 259–266.

PART SIX

INTERVENING
WITH OFFENDERS

Practicing Psychology in Correctional Settings

PAUL GENDREAU AND CLAIRE GOGGIN

IN our contribution to the third edition of *The Handbook of Forensic Psychology* (Gendreau, Goggin, French, & Smith, 2006), we took the liberty of expanding our review beyond the parameters of what is typically understood by the phrase *correctional setting* (i.e., prison) in order to provide readers with the big picture on offender rehabilitation, a topic that had not been covered in previous editions.

We began by charting the arduous journey of the rehabilitative ideal in North American corrections from its inception in 1879 and followed with a summary of the general offender prediction and treatment literatures, primarily from the initial meta-analyses published in the mid-1980s to the *Handbook's* third edition in 2006 (e.g., Andrews & Bonta, 2003; Andrews et al., 1990; Davidson, Gottschalk, Gensheimer, & Mayer, 1984; Garrett, 1985; Gendreau et al., 2006; Gendreau, Little, & Goggin, 1996).

Next, the utility of Andrews and Bonta's (2003) risk-need-responsivity (RNR) model for offender treatment was described in detail. For readers new to this edition, RNR has several principles, of which three are critical. They are that the criminogenic needs (e.g., antisocial attitudes) of offenders at higher risk to reoffend should be targeted for treatment. Effective treatments (i.e., the general responsivity factor) fall within the radical behavioral, social learning, and cognitive-behavior therapy domains. These three principles are supported by the results of meta-analyses and assessments of offender treatment programs using the Correctional Program Assessment Inventory–2010[©] (CPAI-2010[©]; Gendreau, Andrews, & Thériault, 2010). We concluded our earlier chapter with some observations on the rise of correctional treatment quackery and the misuse of significance testing, both of which stand as barriers to progress in our field (Gendreau, Smith, & Thériault, 2009; Schmidt, 1996).

In the current chapter, we say little about the latter two subjects or the prediction and treatment literatures that were covered previously. Readers who are interested in a more comprehensive summary of these issues may consult the recent literature in this regard (e.g., Andrews & Bonta, 2010a, 2010b; Campbell, French, & Gendreau, 2009; Cumming, 2011; Dvoskin, Skeem, Novaco, & Douglas, 2012; Gendreau & Smith, 2007; Gendreau et al., 2009; Goggin & Gendreau, 2006; Hanson, Bourgon, Helmus, & Hodgson, 2009; Hanson & Morton-Bourgon, 2009; Smith, Gendreau, & Swartz, 2009; Yang, Wong, & Coid, 2010; see also Morgan, Kroner, Mills, and Batastini, Chapter 24 this volume).

Our focus here is on prison settings, where most psychologists in the correctional arena are employed. We begin by providing a brief history of prisons followed by a summary of the three major theories of the effects of imprisonment on offender behavior (i.e., misconduct, recidivism). Then we examine these theories in the light of what are considered to be two of the most pressing demands in penology: management of prisons in a safe and humane fashion and the development of prison treatment programs in order to reduce recidivism and thereby enhance public protection (Clements et al., 2007; Gendreau & Keyes, 2001). The degree to which these objectives are being met is addressed through an examination of the results of a number of key meta-analyses and primary studies, most of which have been conducted in the last several years. Finally, we address the controversy surrounding an emerging hot-button topic in penology: the effects of the most severe forms of incarceration (i.e., administrative segregation) on the well-being of inmates.

A BRIEF HISTORY OF IMPRISONMENT

The inclusion of this section requires some explanation. The appointment to faculty of a distinguished historian of crime reminded the first author that what has happened in the past serves as an important lesson that history must be respected, no less so in corrections than elsewhere. Discussion of topical correctional issues (e.g., the types and frequencies of punishment that offenders should experience, the severity of confinement, the role of treatment) might be approached in a more lucid and dispassionate manner if we realized that contemporary perspectives are not especially original. Consider some of the themes that have arisen in the past.

Although it is commonly regarded as a relatively modern invention (cf. Clemmer, 1950; Ives, 1914/1970; Sellin, 1967), the use of imprisonment has been documented for millennia. As Taylor (1899/2003) noted, "In the early cuneiform writing the symbol for prison is a combination of the symbol for 'house' and 'darkness'" (p. 21). The *Shu-king*, a Chinese text dating from the reign of Emperor Yao (2337–2258 BC), documents his punishment of four miscreants, three by banishment and one by imprisonment (Wines, 1880/1968). Although criminal punishment changed little between the classical era and the Middle Ages (i.e., generally blood sanctions,

banishment, detention), its underlying philosophy did evolve to include concepts of specific and general deterrence.

By the early 17th century, English clergy (Wines, 1880/1968) began to decry the conditions common to Britain's gaols and bridewells (Phillipson, 1923/1970), the forerunners of the modern prison (Rotman, 1990). For example, Mynshull's 1618 *Essays and Character of a Prison and Prisoners* described the English gaols as a debauched and mephetic environment in which prisoners of all ages, genders, and offense types commingled (Phillipson, 1923/1970; Wines, 1880/1968). Gambling was rampant, keepers were notoriously corrupt, and inmates were required to pay for their own food and supplies (Ives, 1914/1970; Wines, 1880/1968).

The 18th century marked a period of intellectual reform during which philosophers and scientists championed the merits of rationalism and individualism over the existing system of corrupt social and political values (Phillipson, 1923/1970). In so doing, they fueled discussion of what were, at the time, provocative questions about the role of the criminal justice system: What is the purpose of punishment? Under what circumstances is the state justified in its administration? What is the optimal offense:punishment ratio?

With the gradual replacement of capital punishment by imprisonment, demand surged for more suitable types of confinement. Beccaria's (1764/1986) treatise provided the context within which the conclusions from John Howard's (1991) review of prisons would resonate and, eventually, prompt governments, including those in North America, to remedy conditions within their gaols (Ives, 1914/1970; Meskell, 1999). Regrettably, the response fell short of Howard's recommendations (Erdahl, 2001), as his intent was to reform criminals by housing them in sanitary facilities where they would be separated by age and gender, work in communal silence by day, and be housed in solitary at night. Instead, the conditions within institutions in the early 1800s remained as harsh as those of the gaols they had replaced (Ives, 1914/1970).

DEVELOPMENTS IN NORTH AMERICA

At the conclusion of America's war with Britain in the 18th century, there were great expectations (Rothman, 1998a) that the new country, "rich in both its soil and its industry" (de Beaumont & de Tocqueville, 1833/1970, p. 69), would find answers to social problems that the Old World (Cellard, 2000), hidebound by entrenched political and social values based on the divine right of kings and class structure, had historically been unable to resolve (de Beaumont & de Tocqueville, 1833/1970; Rothman, 1998a): What is the most useful means of punishing criminals? What is the most effective model of confinement vis-à-vis behavioral reform?

The prison was no longer to be regarded as simply a place of detention but as a house of reformation (i.e., penitentiary). The relative merits of two models—the "congregate" system (i.e., Auburn, New York) and the "separate" regime (i.e., Eastern State Penitentiary in Pennsylvania)—dominated debate during most of

the 19th century. Both emphasized strict regimens of solitary confinement, silence, labor, and discipline. Their singular difference lay in the circumstances under which inmates worked, either alone in their cells (separate) or in groups (congregate). Although the two systems shared a common credo—segregation and redemption through the dual lessons of contemplation and hard work—the use of corporal punishment was considerably more frequent under the congregate system, given its greater degree of inmate contact and, consequently, higher rates of disciplinary infractions.

Notable contemporaries did not endorse either prison model. Dickens (1842) characterized the separate system as stultifying “torture of the brain” (p. 124), because of its emphasis on almost complete social isolation. Others disparaged the congregate model for its substandard dietary and sanitary conditions and its excessive use of corporal punishment (de Beaumont & de Tocqueville, 1833/1970; Dix, 1845/1967). Regardless, the congregate model was adopted as the standard throughout North America, not least because it offered penurious governments a win-win scenario by providing them with much-needed houses of correction as well as a steady source of contract labor. The operational demands of the developing prison system, often reliant on inmate labor for its very construction, left it vulnerable to excessive use of corporal punishment, a control mechanism that continued to dominate prison administration policy (Rotman, 1990), and bedevil prison reformers, throughout the 19th century and into the 20th.

RISE OF THE REHABILITATIVE IDEAL

At the same time, the late 19th century was a “golden age” for scholars interested in maximizing prison’s potential to reduce criminal behavior (Ives, 1914/1970; Johnston, 2000; Scalia, 1871). In 1879, the first of five U.S. conferences sponsored by the National Prison Association was convened in Cincinnati. Its foremost accomplishment was the development of 37 “epoch-making” (Glueck & Glueck, 1939/1965, p. 22) principles of prison management. These principles delineated standards for the treatment of inmates as well as the education and conduct of prison guards. The authors of the Cincinnati principles were insightful in their understanding of behavioral change mechanisms, given their emphasis on the negative impacts of corporal punishment and the importance of securing inmate participation and cooperation (Guilford, 1946; Wines, 1880/1968) in effecting behavior change, concepts that would resonate in 20th-century learning theory (Mowrer, 1960; Spiegler & Guevremont, 2010).

Regrettably, this determined commitment to prison reform did not immediately result in more humane conditions. Conditions at the end of the 19th century differed little from those at its beginning (American Correctional Association, 1972; Oliver, 1998; Rotman, 1998; Scalia, 1871), mostly because of the inherent punitiveness of both prison models (Carrigan, 1991) and, by midcentury, serious levels of institutional

overcrowding (Abbott, 1927; Rothman, 1998a). Moreover, despite broad circulation of enlightened prison management concepts, most governments lacked the political will to allocate scarce resources to improve standards of care to levels beyond those available to the general population (i.e., principle of less eligibility; Carrigan, 1991; Tomlinson, 1978).

As a result, any improvements in the conditions of inmate confinement and treatment in the 19th century, such as the innovative incentive-based token economy like prison models pioneered by Maconachie at Norfolk Island in the South Pacific, Sir Walter Crofton in Ireland, or Brockway at Elmira Reformatory in New York (Wines, 1880/1968) had little or no generalized impact on contemporary correctional policy (Putney & Putney, 1962; Rotman, 1990; Zubrycki, 1980), although they undoubtedly laid the groundwork for the eventual development of a prison reform agenda (Allen, 1981).

Early efforts to establish an optimal prison model that could reliably affect positive change in inmate behavior did not immediately have the system-wide impact advocated by Wines and Dwight (1867/1973) and Brockway (1871). They were, however, pivotal in helping to establish the foundation for the growth of the rehabilitative principles upon which contemporary correctional professional practice rests: (a) criminality can be accounted for by identifiable psychological and social factors, and (b) suitable treatment regimens can be designed to address deficits in these two areas. In the absence of such principles, we are left with only the historical precedents as a response to criminal behavior. As Brockway (1871) noted:

If punishment, suffering, degradation are deemed deterrent, if they are the best means to reform the criminal and prevent crime, then let prison reform go backward to the pillory, the whipping-post, the gallows, the stake; to corporal violence and extermination! But if the dawn of Christianity has reached us, if we have learned the lesson that *evil is to be overcome with good*, then let prisons and prison systems be lighted by this law of love. Let us leave, for the present, the thought of inflicting punishment upon prisoners to satisfy the so-called justice, and turn toward the two grand divisions of our subject, the real objects of the system, viz: *the protection of society by the prevention of crime and reformation of criminals*. (p. 42)

We now turn to a discussion of the major theories of the effects of imprisonment on inmate behavior that have evolved over time, with particular attention to their ability to provide insights into best practices regarding the safe and humane management of prisons and the successful reintegration of inmates into the community.

THEORIES OF THE EFFECTS OF IMPRISONMENT

We begin this section by discussing the three most prominent theories of the effects of imprisonment in terms of their chronological development.

DETERRENCE

Deterrence theory is among the oldest theories of criminal behavior and is predicated on the notion that prosocial behaviors can be elicited by exposure to selected punishers (i.e., prison time, corporal punishment, etc.). In the modern era, proponents of deterrence most commonly hail from the disciplines of criminology and economics. By contrast, psychology's understanding of the utility of punishment as a behavior change mechanism has been derived from an extensive and complex body of findings from rigorous experimental studies within the experimental learning and behavior modification domains. For example, the necessary "punishers" (e.g., physically aversive and/or painful stimuli, response cost) and conditional factors (e.g., punishment is administered immediately, at maximum intensity, with no opportunity of escape) that reliably suppress behavior have been extensively documented (see Masters, Burish, Hollon, & Rimm, 1987; Matson & Dilorrenzo, 1984).

By comparison, deterrence theorists' notions of the effectiveness of punishment, as well as those of the public (i.e., DeJong, 1997; Roberts, Crutcher, & Verbrugge, 2007), we would suggest, are founded on commonsense ideas and moral considerations (Gendreau, Goggin, & Cullen, 1999; Gendreau & Ross, 1981). Loss of income, stigmatization, and dehumanizing prison-based psychological events are tendered as putative punishers (Gendreau, Goggin, & Cullen, 1999; Nagin, 1998). In contrast to behavioral psychologists, deterrence advocates have never operationally defined or replicated the conditions under which punishment suppresses behavior. Essentially, the deterrence model has reduced behavior to a simple economic cost-benefit equation (see Listwan, Sullivan, Agnew, Cullen, & Colvin, 2013) whereby inmates are rationally capable of quantifying the exact dosage of pain that prison life has imposed on them and are then able to predict with absolute certainty whether they will desist from crime upon release. Much as psychopharmacologists do in their experiments, some economists have speculated that the deterrence puzzle will be solved when a specific "dosage" of pain (i.e., defined by a specific length of incarceration) is found that reduces criminal behavior but then only for inmates at lower risk of reoffending (Orsagh & Chen, 1988).

If the previous paragraph does not sound an alarm bell regarding the allure of the deterrence model, consider the following. Deterrence supporters have decried that modern-day U.S. prisons are not punishing enough, given the loss of their stigmatizing qualities through high rates of incarceration (Nagin, 1998). As a case in point, experiencing prison is almost a rite of passage for many African American males as 1 in 3 are estimated to be incarcerated at some point in their lifetime (Bonczar, 2003). Deterrence proponents advocate that prison conditions be made tougher by reinstating corporal punishment and reducing prison amenities such as education, health care, recreation, and visits (Finn, 1996; see Listwan et al., 2013; Nossiter, 1994). In effect, the promotion of such policies represents a return to the congregate model of the 1800s where, as noted, punishment was an integral component of inmate management.

Second, there is an extensive social psychology literature that has been cavalierly ignored. Whereas the rational choice model has grossly oversimplified matters by implicitly assuming that attitudes are perfectly correlated with behavior, the social psychology literature highlights the reality that myriad intervening variables demonstrate how the relationship can break down. In testament to this fact, it has been well documented that the correlation between holding a particular attitude and demonstrating a corresponding behavior can range from very weak ($r = .10$) to moderate ($r = .60$), depending on a variety of conditions (Eagley & Chaiken, 1993; Fishbein, 1995; McGuire, 1995). Finally, the rational choice position does not acknowledge that some commonplace offender characteristics (e.g., concrete thinking, egocentricity, impulsivity, psychopathy, being under the influence of substances) are antagonistic to sound decision making and, in all probability, decrease the likelihood of offenders being responsive to punishment in the first place (Gendreau & Suboski, 1971; Gendreau, Goggin, & Cullen, 1999; Hare, 1996).

What, then, does the empirical record say about how well deterrence speaks to the principle aims of penology? The goal of protecting the public through incarceration-related reductions in recidivism has been an unmitigated failure. There have been several meta-analyses of the effects of prison on subsequent reoffending (Gendreau, Goggin, & Cullen, 1999; Jonson, 2010, 2013; Nagin, Cullen, & Jonson, 2009; Smith, Goggin, & Gendreau, 2002; Villettaz, Killias, & Zoder, 2006). These studies involved considerable sample sizes (e.g., $n = 107, 165$, Smith et al., 2002) and different meta-analytic strategies, and the researchers involved had varying perspectives on the matter, as they represented different disciplines (e.g., criminology, economics, psychology). Nevertheless, the conclusions are remarkably similar. Imprisonment generally results in increases in recidivism in the range of several percentage points. Furthermore, when moderators such as differences in sentence length, higher versus lower prison security level, inmate age and risk level, and the quality of research design have been examined, the results of the meta-analyses do not support deterrence theory (e.g., Smith et al., 2002).

As to the second goal of making prisons safer and more humane environments, no-frills proponents surmise that making prisons less hospitable will somehow result in decreased levels of troublesome behaviors (e.g., misconduct) and provide the added bonus of reducing recidivism. The available data in this area are very limited, as prison files typically lack information that meets the standards required by researchers and published studies on prisons contain virtually no information for those wishing to code moderators in meta-analyses (Gendreau & Smith, 2012; S. Listwan, personal communication, September 22, 2012). Researchers are often left with only security level comparisons that may, in fact, be a distant proxy for "harsh conditions," a catchall label for any one of reduced privileges, corporal punishment, denigration of inmates by staff, or high-tech electronic monitoring within prisons (see Gendreau, Goggin, & Cullen, 1999; Piehl & Useem, 2011; Wortley, 2002). Among the handful of studies that bear on the topic, results have so far been generally disappointing for no-frills supporters (Briggs, Sundt, & Castellano, 2003;

see the review by French & Gendreau, 2006; Sundt, Castellano, & Briggs, 2008). As for harsher prison conditions (i.e., effects of higher security levels on recidivism), the results point to increases in recidivism of up to 14% (Gaes & Camp, 2009; Jonson, 2010).

In conclusion, the meta-analyses are conclusive: Prisons are not a deterrent to future criminal behavior. The theory also has little to offer in regard to the safe and humane management of prisons.

SCHOOLS OF CRIME

The “schools of crime” theory, likely most familiar to the media and general public (Cullen, Fisher, & Applegate, 2000; Mason, 1998), predicts an effect of prison directly opposite to that of deterrence. It is the perception that prison environments are “graduate schools” of crime, conferring the ultimate degree in learned criminal attitudes and behaviors. Moreover, the longer the period of imprisonment, the greater the extent of criminal skill acquired (Jaman, Dickover, & Bennett, 1972). The common term used to describe this process is *prisonization* (Clemmer, 1940; see also Sykes, 1958).

Various theories of prisonization have been postulated. They range from differential association, general strain, labeling, and self-control theories in criminology (Agnew, 2006; Akers, 1977; Colvin, 2000; Hirschi, 1969; Lemert, 1951) to social learning in psychology (Buehler, Patterson, & Furniss, 1966; Bukstel & Kilmann, 1980). Bukstel and Kilmann (1980) were concise in their conclusion that the results of the studies they reviewed convinced them that there was evidence of “overwhelming positive reinforcement” (p. 472) from peers and staff to promote antisocial attitudes and behaviors.

From this analysis, one would assume that all inmates, as a matter of course, would cultivate more criminal values, but the research record indicates a more complex picture. Initial studies in this area reported few consistent prisonization effects, so much so that many criminologists abandoned the topic due to a lack of support for their theory (Bonta & Gendreau, 1992; DeLisi & Walters, 2011). Unlike criminologists, who historically have been highly skeptical of topics such as individual differences and offender treatment (Andrews & Wormith, 1989; Cullen & Gendreau, 2001), psychologists and criminologists sympathetic to the rehabilitative agenda have taken the lead in continuing to search for additional factors that could moderate or be affected by the prison experience (e.g., offender risk level; stage of an inmate’s sentence; prison management style, availability, and program quality; Bonta & Gendreau, 1992; DeLisi & Walters, 2011; Lowenkamp, Latessa, & Smith, 2006).

Among the aforementioned variables, risk level has offered the most evidence in support of the prisonization hypothesis. An emerging research database suggests that the socialization experiences in prison and the exposure of low-risk inmates to higher-risk peers in prison treatment programs lead to increases in infraction rates

in prison and postrelease recidivism (Gendreau & Smith, 2012; Latessa, Lovins, & Smith, 2010; Latessa, Lovins, Smith, & Makarios, 2010; Smith & Gendreau, 2012; Tanaischuk, Wormith, & Guzzo, 2009; Wooldredge, 1998). Should this pattern be replicated in future research, the potential for serious harm to inmates is enormous, given the present high levels of incarceration and the fact that low-risk offenders are overrepresented in many prison populations (see Bonta & Motiuk, 1992; Goggin, 2008).

In summary, the schools of crime perspective has led to an important finding: Public protection is not furthered by incarcerating low-risk offenders, especially in prisons where they may be negatively influenced by their higher-risk peers. The theory has been silent, however, on precisely how to deal with this problem through classification and correctional treatment.

BEHAVIORAL DEEP FREEZE

The importation or “behavioral deep freeze” model has its roots in the work of criminologist Charles Thomas (Thomas & Foster, 1973). Thomas (1977) challenged the hegemony of the schools of crime theory by putting forth evidence in support of the role of imported inmate experiences, both pre- and postprison, in explaining inmates’ degree of adjustment to the conditions of prison life. Subsequently, Zamble and Porporino (1988, 1990) followed up on Thomas’s ideas while also drawing on the general psychological coping literature (see Lazarus & Folkman, 1984) to advance their notion that prisons were primarily behavioral deep-freeze environments.

An impressive body of research supports the work of Porporino, Thomas, and Zamble (see reviews by Bonta & Gendreau, 1990; Smith & Gendreau, 2012). Studies in this area have employed both cross-sectional (with inmates who had served varying lengths of sentence up to 14 years) and longitudinal (follow-ups of up to 7 years) designs and report negligible effect sizes between prison experience and inmates’ cognitive functioning, personality, mood (e.g., anxiety, depression), and psychophysiological functioning (Banister, Smith, Heskin, & Bolton, 1973; Bolton, Smith, Heskin, & Banister, 1976; Flanagan, 1980; Gendreau, Gibson, Surridge, & Hug, 1973; Gendreau, Grant, & Leipziger, 1979; Heskin, Bolton, Smith, & Banister, 1974; Heskin, Smith, Banister, & Bolton, 1973; MacKenzie & Goodstein, 1985; Rasch, 1981; Richards, 1978; Sapsford, 1978; Wormith, 1984, 1986; Zamble, 1992; Zamble & Porporino, 1990).

A subset of prison life research also lends credence to the deep-freeze theory. This is the issue of prison overcrowding, a condition that has the potential to produce negative psychophysiological side effects during incarceration. Bonta and Gendreau (1990) summarized this literature with a meta-analysis and reported that overcrowding correlated modestly ($r \approx .20$) with physiological indices (e.g., elevated heart rate and blood pressure) and self-report measures of stress (e.g., feelings of discomfort, of being crowded), but it was only weakly ($r < .10$) correlated with acting-out behaviors. Since then, the major thrust of their conclusions has

been sustained (e.g., Camp, Gaes, Langan, & Saylor, 2003; Gendreau, Goggin, & Law, 1997). Unless crowding is a chronic problem, it is likely that other factors, such as inmate perceptions of control, prison management style, staff supervisory practices, sudden changes in the prison population demographic (e.g., influx of younger inmates), and design capacity, are more important predictors of inmates' well-being (Bonta & Gendreau, 1990; Steiner & Wooldredge, 2008, 2009).

In summary, Zamble's (1992) conclusion regarding the effects of incarceration, that "the most striking result was in the total absence of any evidence for general or *widespread* deteriorative effects" (p. 420, emphasis added) serves as a succinct summary of the entire field to date. Zamble did leave room for the possibility that some inmates may cope poorly with confinement, an experience that could, in turn, affect postrelease recidivism. In his work with Porporino (Zamble & Porporino, 1990), it was found that inmates who recidivated were those who had coped most poorly in prison. Recidivists were also those who had the most extensive criminal histories, which implies that members of this group were higher risk.

Another research area also offers some insights into the deep-freeze construct and its potential effects on inmate behavior. Here we refer to the notion of correctional climate (CC) or prison "personality." Once fashionable (i.e., Moos, 1968), it has recently been revived by Clements et al. (2007) and Goggin (2008), due in part to concerns about the effects of more severe prison conditions (i.e., supermax prisons) on inmate adjustment (see Gibbons & Katzenbach, 2006; Haney, 1997; Smith & Gendreau, 2012; Useem & Kimball, 1989).

Goggin (2008) used ratings of perceived CC from two comprehensive surveys completed in 1995–1996 by inmates ($n = 4,283$) and staff ($n = 2,717$) in 43 Canadian federal prisons to evaluate its effects on inmate misconduct and recidivism. The content areas of the two surveys differed somewhat per sample, but each covered all of the domains relevant to an assessment of CC (e.g., living and working conditions; security, health, and safety; inmate programming; etc.). There was a considerable degree of congruent validity in the ratings of the two groups, with mean scores of 28.8% for inmates and 39.0% for staff, where higher scores indicate more negative CC. In the absence of published norms for CC, the question remains as to whether these generally positive ratings of prisons in the Canadian correctional system, until recently regarded as one of the most progressive among Western nations (Simpson, 2012), are generalizable to other federal or state equivalents.

Goggin (2008) reported that higher CC ratings were correlated with poorer prison adjustment (i.e., nonserious and serious misconducts) and postrelease recidivism (i.e., new charge), with r s ranging from .06 to .16. Results indicated that the poorest outcomes were found for moderate and low-risk inmates, with high-risk inmates seemingly little affected by prison climate. An examination of the results by security level revealed that outcomes were much worse among inmates in maximum-security settings (i.e., correlations between CC and serious misconducts and recidivism were .34 and .14, respectively). We draw attention to

the fact that these effects of CC on recidivism were similar to Jonson's (2010) meta-analytic findings for "harsher" prisons, which, in the latter case, were defined by security level.

Another source of evidence that tentatively challenges the notion that the importation model may not generalize to all inmates comes from the emerging literature on prison victimization (Listwan, Colvin, Hanley, & Flannery, 2010). In a manner similar to that of Goggin (2008), Listwan et al. (2010) evaluated the effect of prison climate, as measured by retrospective inmate ratings of negative inmate relations (i.e., inmate victimization by inmates), negative relations with correctional officers, and perceptions of negative prison environment, on indices of psychological trauma and reported effect sizes (r) in the range of .10. Granted that measures of prison climate were not collected until after release, it is quite possible that the negative feelings resulting from victimization also affected inmates' behavior during their incarceration. Although the study's authors were unable to access risk level data (S. Listwan, personal communication, October 22, 2012), we surmise that the study's low-risk inmates were likely among the most negatively affected.

Given the role of environment in the delivery of effective prison-based correctional programs (i.e., Andrews & Bonta, 2010a; Goggin & Gendreau, 2006; Palmer, 1994) and, ultimately, the safe and humane operation of prisons (i.e., Armstrong & MacKenzie, 2003; Gendreau & Keyes, 2001), prison climate studies should be pursued so researchers in the area can better understand how CC differentially affects inmate attitudes and behaviors, especially in maximum-security settings.

Thus, the deep-freeze perspective provides a cautionary message to the more dramatic claims from the deterrence and schools of crime theorists. In the strictest sense, the theory appears to offer little guidance regarding policy, but it should be recognized that deep-freeze advocates have been strong proponents of employing treatment programs to assist inmates' prosocial adjustment to prison and improve their reintegration potential (Zamble & Porporino, 1990).

In conclusion, knowledge of these three theories is an important first step for forensic psychologists in gaining an understanding of the battle lines in the debate surrounding the effects of imprisonment. The deterrence position, in our view, has no justification empirically, and the get-tough side of this perspective (i.e., mean-spirited policies toward offenders) is antithetical to psychologists' professional code of ethics. In contrast, the schools of crime and deep-freeze theories have a reasonable degree of empirical credibility and can provide a useful framework to guide forensic psychologists in their research and clinical activities within prison settings (e.g., handling of low-risk inmates, being acutely aware of the values that offenders bring to prison, need for treatment).

In order to satisfactorily address the two questions we posed at the outset, however, we must consult another source of evidence, otherwise known as what works, in the prediction and treatment of offender behaviors within the prison context.

PREDICTION AND TREATMENT

In comparison to the foregoing, the emerging database on the prediction and treatment of criminal behavior offers specific guidelines regarding the management of prisons and the reduction of recidivism.

MANAGING PRISONS SAFELY AND HUMANELY

Prisons that are managed satisfactorily typically have low rates of inmate misconducts. These misconducts may include anything from minor infractions, such as disobeying orders, to assaults that jeopardize the safety of other inmates and staff. Surprisingly, there has been relatively little research on this subject despite the fact that over 60 years ago, Alfred Schnur (1949) made some key discoveries about the importance of attending to institutional misconducts, which he found to be a proxy for criminal behavior. In his study, offenders with more extensive criminal histories had more misconducts ($n = 1,762$; $r_{range} = .17-.50$), misconducts predicted postrelease recidivism, and inmates whose misconducts increased over time in prison also had higher recidivism rates.

Why are these results so important for prison authorities? If the results can be replicated, inmates who pose the most problems in terms of prison management can be identified and closely monitored by custodial staff, be transferred to more secure units, or be placed in treatment programs that may help them to better adjust to prison life. Another motivation for reducing misconduct levels is the sizable costs that can accrue from assaults and damage to prison property (Lovell & Jemelka, 1996). Since misconducts represent criminal behavior, one can assume that, if treatment programs are effective in reducing infractions, this could translate into postrelease benefits as well (i.e., lower recidivism). Further, the above findings have profound implications for parole and probation reentry policies and practices (Glaser & Stratton, 1972).

Fortuitously, in recent years, there has been substantive confirmation of Schnur's (1949) findings (see the review by Cochran, Mears, Bales, & Stewart, 2012; Gendreau et al., 1997; Motiuk, 1991). Gendreau et al. (1997) generated a meta-analysis which reported that both nonviolent and violent misconducts could be predicted equally well. Of the individual predictors in their data set, previous misconducts produced the highest validities ($r = .21$; weighted effect size = .32). The other useful predictors of misconducts (e.g., antisocial attitudes, criminal history, social achievement) were of similar magnitude to those reported in an earlier meta-analysis of predictors of recidivism (Gendreau et al., 1996), adding further weight to the proxy argument.

Among risk measures in common usage at the time of the Gendreau et al. (1996) study, the Level of Service Inventory—Revised (LSI-R; Andrews & Bonta, 1995) produced the highest predictive validities with outcome ($r = .22$, $n = 2,252$, $k = 10$). Subsequently, Campbell et al. (2009) examined the ability of risk measures to predict violent misconducts and found that the Historical-Clinical-Risk Management-20

scale (HCR-20; Webster, Douglas, Eaves, & Hart, 1997) ($n = 758$, $k = 11$) produced the highest predictive validities ($r = .31$; confidence interval about r (CI_r) = $.21-.40$), although three other risk measures—the LSI-R (Andrews & Bonta, 1995), the Psychopathy Checklist–Revised™ (PCL-R; Hare, 2003), and the Psychopathy Checklist: Screening Version™ (PCL-SV; Hart, Cox, & Hare, 1995)—generated CIs that overlapped with those of the HCR-20 (Webster et al., 1997). Until more research is conducted, Campbell et al.’s (2009) results should be regarded as tentative. Sample sizes were quite small, and the width of the CIs indicated a lack of precision in the effect size estimates (for CIs whose width is $> .10$, see Gendreau & Smith, 2007), many studies provided very little information about their samples (i.e., violence history, nature of the violent index offense), and most samples were assessed as low risk to reoffend.

Smith and Gendreau (2012) provided further confirmation that institutional misconducts are a useful proxy for criminal behavior. They found that infractions that resulted in incident reports (e.g., refusing to work) or segregation (e.g., assault) predicted parole revocation and reincarceration among a large sample ($n = 5,038$) of Canadian penitentiary inmates ($r_{range} = .16$ to $.22$). The investigators also noted that a seven-item measure of prison adjustment (e.g., misconducts, escapes, criminal history) generated by the Canadian authorities predicted misconducts and recidivism on the order of $r_{range} = .28$ to $.31$. Finally, Cochran et al. (2012) examined the issue in an interesting way by comparing the recidivism rates of a large sample of inmates, matched on several risk factors, who had or had not committed misconducts and reported that the violent recidivism rate was 12% higher among the former.

As to the matter of the ability of misconducts to predict recidivism over time, Smith and Gendreau (2012) reported a similar finding to that of Schnur’s (1949), with one important qualification. Their study included two groups whose misconducts were charted over time: One whose sentence length was no more than 2 years, the other whose sentence length encompassed a 2- to 20-year period. Misconducts were operationalized in terms of severity, where more serious misconducts resulted in a period of segregation and less serious ones resulted in an incident report. Inmates with higher rates of both types of infractions showed higher rates of recidivism ($r_{range} = .07-.23$). Sample sizes were small ($n = 109-611$), however; thus the CIs were wider than desirable. Of note, this same pattern of results was not found for higher-risk inmates. Replication of this study is essential.

The obvious inference from the foregoing is that misconducts are a useful surrogate for criminal behavior. The next question is whether there are interventions that can deal effectively with misconducts. Antidotes to institutional infractions have ranged from situational crime-control procedures, climate control, better prison design, get-tough policies, reductions in inmate turnover, diet, and treatment programs (for a review, see French & Gendreau, 2006, pp. 187–188). Each of these proposals may have merit, but, with the exception of treatment programs, none has received more than weak empirical support to date.

The utility of treatment programs was first summarized in two meta-analyses by Keyes (1996) and Morgan and Flora (2002). The database was later greatly expanded by French and Gendreau (2006), who reported that the most effective programs were congruent with the RNR treatment model. Interventions that were behavioral in nature, targeted more versus fewer criminogenic needs (e.g., antisocial attitudes and values), and scored higher on therapeutic integrity as assessed by the CPAI-2010[©] (Gendreau et al., 2010) produced results at least twice as large as those of comparison categories. For example, effect sizes for behavioral versus nonbehavioral programs were $r = .26$ and $r = .10$, respectively. Another important result, albeit a tentative one given the small number of studies and the overlap in CIs about the results, was that programs that produced the best results in reducing misconducts also had the best recidivism outcomes ($r = .13$, CI = .04, .29, $k = 12$). In contrast, prison-based programs that were relatively ineffective in reducing misconducts were associated with slight increases in recidivism ($r = -.05$, CI = $-.16$, .07, $k = 11$).

Unfortunately, the French and Gendreau (2006) database did not provide information on offender risk level, but the Smith and Gendreau (2012) primary study was able to do so. In their survey of the effects of treatment programs within all federal prisons in Canada, they documented results that are alarming. Focusing on the most serious forms of misconduct, those that typically result in segregation (e.g., assaults), treatment programs that were consistent with the RNR's need principle decreased segregation rates among moderate- and high-risk inmates by 1% and 8%, respectively, but increased segregation rates among low-risk inmates by 8%. Programs that did not appear to adhere to the need principle increased segregation rates by 16% to 20% across all inmates regardless of risk level.

The final piece of evidence that supports the use of treatment programs to manage prison adjustment comes from a long-forgotten treatment modality that was supplanted by the cognitive-behavioral revolution. We refer to contingency management programs (e.g., token economy) that were frequently used in prisons 30 to 40 years ago (see Milan, 1987). Reading this literature, one is impressed with the magnitude of the positive effects reported in single studies, but until recently there had been no summary of their overall effectiveness.

In 2012, Gendreau et al. (2013) conducted the first meta-analysis in this area. They reviewed 29 studies, over 90% of which were token economy programs, that generated 64 effect sizes ($n = 1,033$) with various measures of institutional adjustment. Readers familiar with radical behavioral treatments of this sort will appreciate that token economy programs are notable for the replication of their findings and their high dosage levels. This group of studies had, on average, three replications and a treatment dosage of 123 days. The average reduction in target behaviors, about half of them directly affecting antisocial behaviors, was 62%. With that in mind, Gendreau, Listwan, and Kuhns (2013) outlined the pitfalls (e.g., inflation, deliberate sabotage by hostile inmates and staff) that are routinely encountered when running

such programs unless psychologists, who are usually responsible for their design and operation, pay exceptionally close attention to treatment fidelity.

MANAGING PRISONS TO PROTECT THE PUBLIC

Having established that misconducts are a reasonable facsimile of criminal behavior and that interventions employing some of the principles of RNR can reduce misconducts, one might expect that prison-based treatment programs reduce postrelease recidivism. Admittedly, we already know that prison treatment programs “work” based on the classic Andrews et al. (1990) meta-analysis. The studies from 20 to 40 years ago, however, were often short on clinical lore or details of the black-box elements that likely contributed to their success. This should not be construed as a criticism, because our knowledge base at that time was in its infancy (Gendreau, 1996). The next four studies have made significant strides in this regard.

We begin with a description of the Rideau Correctional Centre program (Bourgon & Armstrong, 2005). It was first off the mark, as the seeds of its development lie in the early 1970s. Among the noteworthy features of the program was its faithful adherence to the RNR model. Assessment of criminogenic needs was extensive in that the program utilized the LSI-R (Andrews & Bonta, 1995) and other well-validated psychological measures that examine specific responsivity issues. Treatments were cognitive-behavioral (CBT) in nature, in the domains of criminal thinking, anger management, substance abuse, and relapse prevention. Role-play, modeling, and behavioral rehearsal formed part of the essential core of CBT. Group leaders were chosen on the basis of personal qualities recommended in the general therapeutic literature. Sample sizes for the treatment group, including nonprogram completers, and the risk-controlled comparison group were 482 and 138, respectively.

Postprogram recidivism was reduced by 13% over a 2-year follow-up period. Since dosage level has historically been an issue in determining the effects of offender treatment (see Lipsey, 1995), Bourgon and Armstrong (2005) also calculated reductions in recidivism for various comparisons of treatment dosage by risk level for completers and no-treatment controls. This matching of risk and dosage resulted in a further reduction in recidivism of 1% to 7%, depending on the comparison. Finally, the authors estimated that, for each week of treatment, one could expect a 1% to 2% reduction in recidivism. They concluded that a minimum of 300 hours of treatment exposure may be optimal.

Next, we visit the Vermont program developed by Jack Bush in the late 1980s. A philosopher by training, Bush was influenced by three paradigms in offender treatment and one on relapse prevention (Andrews, 1980; Bandura, 1973; Goldstein, 1999; Marlatt & Gordon, 1985; Yochelson & Samenow, 1976). His outlook on offenders was that violence is not an isolated behavior and that offenders are not specialists in certain types of crime. Further, they are not mentally disordered

nor should they be considered patients. His treatment process emphasized the use of thinking reports to teach inmates to identify their thinking patterns and cognitive distortions and then learn specific skills to control their impulses and practice relapse prevention (Bush, 2005). As with the Rideau program, role play and behavioral rehearsal were key components. There was also a direct conduit to community-based continuation of the intervention, which likely was a vital component of its success.

Henning and Frueh (1996) provided a matched control group evaluation of the Vermont program and reported a 21% reduction in recidivism for the treatment group ($n = 28$) as compared with the nontreatment group ($n = 141$) during a 2-year follow-up. Dropouts from treatment were also included in the analysis. It is difficult to calculate the precise treatment dosage level used in this study. Sadler and Powell (2008) stated that at least 150 hours was the minimum dosage for the program; from Bush's (2005) description, time in treatment was quite a bit longer.

The third study in this series comes from a program evaluation enterprise that was impressive in its scope. Ohio officials evaluated the performance of 80 state-run halfway houses and community-based correctional facilities (i.e., equivalent to minimum-/medium-security prisons). There were two sets of evaluations, one in 2002 and the other in 2006, which were conducted by the University of Cincinnati research team (Latessa, Lovins, & Smith, 2010; Latessa, Lovins, Smith, et al., 2010; Lowenkamp et al., 2006). The programs varied tremendously in their diversity, and many of them had very limited recognition or understanding of the aforementioned research literature.

The Latessa, Lovins, and Smith (2010) evaluation reported outcomes for all program participants and successful completers of the correctional residential treatment programs on various measures of recidivism. Overall, the results provided minimal support for treatment, but, when treatment effects were evaluated on the basis of the risk principle, an interesting picture emerged. As seen in Figures 23.1 and 23.2, on average, high-risk offenders benefited more from treatment while low-risk offenders became worse. In addition, the Cincinnati group looked into the black box of what the Ohio programs purported to be doing. Following Lowenkamp et al. (2006), who found that a number of indices of therapeutic integrity as measured by the CPAI-2010[©] (Gendreau et al., 2010) produced strong correlations with recidivism ($r_{range} = .25$ to $.54$), Latessa, Lovins, Smith, et al. (2010) reported correlations in the same general range for program leadership, staff characteristics, assessment protocols, targeting of criminogenic needs, and the appropriate use of punishment with recidivism (e.g., $\approx 10\%$).

The risk principle was once more confirmed by Smith and Gendreau (2012) in their examination of the effects of prison programs on recidivism. Programs that conformed to the need principle increased recidivism by 4% for low-risk inmates but reduced it for moderate- and high-risk inmates by 7% and 11%, respectively. Programs that did not appear to be following this principle, however,

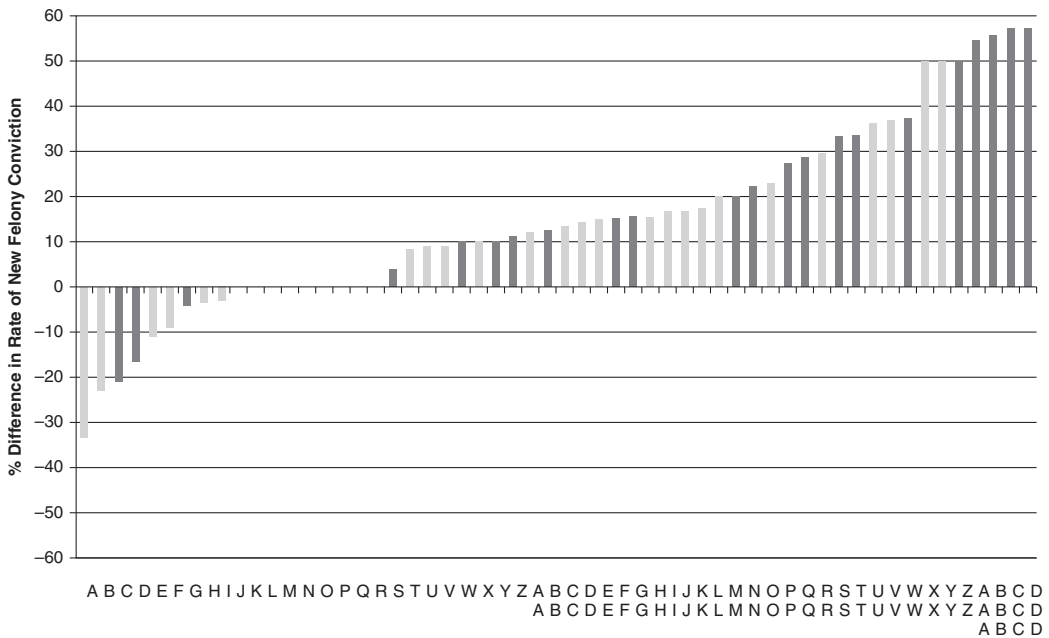


Figure 23.1 Treatment Effects Measured by New Felony Conviction for CBCF/ISP and HWH Samples—High Risk.
Note. CBCF = community-based correctional facility; ISP = intensive supervision probation; HWH = halfway house. Positive treatment effects are represented by the bars above the midline; negative treatment effects are represented by the bars below the midline. Programs with a successful completion rate above 65% are represented by black bars (whether positive or negative) while those with completion rates at 65% or below are represented by gray bars.

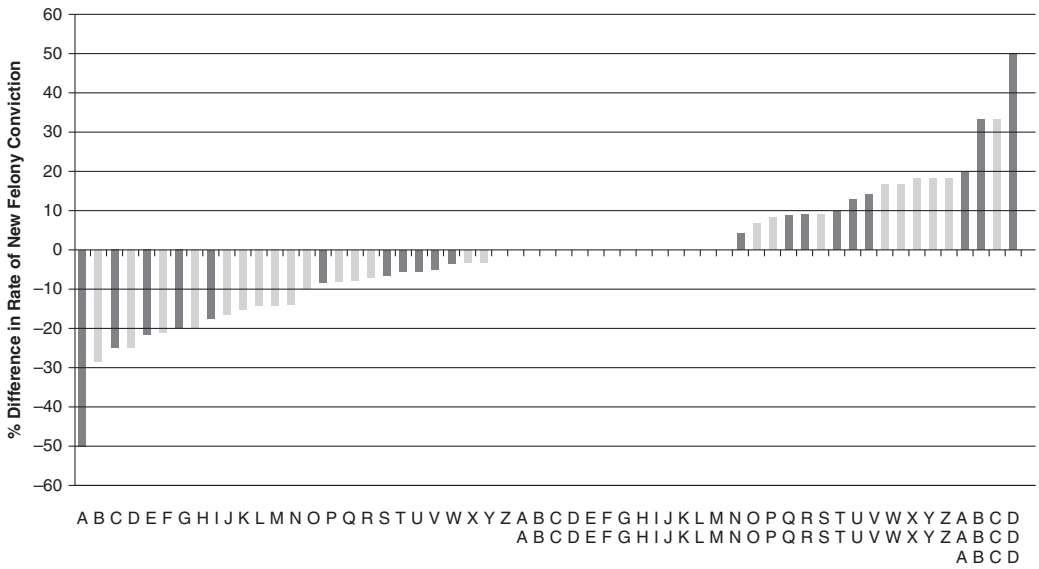


Figure 23.2 Treatment Effects Measured by New Felony Conviction for CBCF/ISP and HWH Samples—Low Risk.
 Note. CBCF = community-based correctional facility; ISP = intensive supervision probation; HWH = halfway house. Positive treatment effects are represented by the bars above the midline; negative treatment effects are represented by the bars below the midline. Programs with a successful completion rate above 65% are represented by black bars (whether positive or negative) while those with completion rates at 65% or below are represented by gray bars.

were associated with higher rates of recidivism for low-, moderate-, and high-risk inmates (14% versus 7% versus 6%, respectively).

We conclude this section with a short discussion of three issues that are frequently raised about the aforementioned results. They concern the magnitude of reductions in recidivism, cost benefits, and why low-risk inmates are adversely affected, even when programs follow some RNR principles.

The early narrative reviews and meta-analyses on correctional treatment painted a rather rosy picture of the magnitude of the effects in this matter (e.g., up to 30% reductions in recidivism; see Andrews et al., 1990; Gendreau & Ross, 1987). Critics of the rehabilitative agenda (e.g., Lab & Whitehead, 1990; Whitehead & Lab, 1989) dismissed these findings as being overly optimistic because they were based on utopian demonstration projects (e.g., programs that were generously funded and carefully implemented) and conducted by skilled clinicians, often university based, who were experienced in delivering effective offender treatments. A classic example of their point can be found in the well-known interventions by William Davidson's research group (Davidson, Jefferson, Legaspi, Lujan, & Wolf, 2001; Davidson, Redner, Blakely, Mitchell, & Emshoff, 1987). It would be reasonable to assume that programs like Davidson's would rarely be found in many real-world settings run by government or nonprofit agencies, where the outcomes, especially in prisons, could be expected to be much less robust (Gendreau, Goggin, & Smith, 1999).

Lipsey (1999) was the first to confirm the critics' prediction based on the results of what he called "routine practice" programs. He examined 196 such programs for juvenile offenders and reported a 3% reduction in recidivism, which he regarded as a beneficial result. These ballpark figures are similar to what has been reported for several adult prison treatment samples (Gendreau & Smith, 2012).

Can such "modest" reductions be fiscally meaningful? There is now a burgeoning literature in the affirmative (e.g., Drake, Aos, & Miller, 2009; Romani, Morgan, Gross, & McDonald, 2012). The research area is deceptively complex (e.g., cost-benefit versus cost-effective analyses, tangible versus intangible costs). Estimates of treatment's bang for the buck vary widely, ranging from savings of \$2,600,000 for each youth who ceases a criminal career by age 18 (Cohen & Piquero, 2009) to \$15,000 in life cycle benefits that are associated with small but meaningful reductions in recidivism (i.e., 7%; Drake et al., 2009; Farrington, Petrosino, & Welsh, 2001). Of the many estimates that have been proffered, that by Romani et al. (2012, see table 4, p. 158) best illustrates the cost-savings point, because they compared three standard treatment domains: traditional punishment approaches, services that adhere to RNR principles, and programs that do not follow RNR principles. Their analysis included programs that *successfully* reduced recidivism by at least 1%. The mean cost per 1% reduction was 20 times less for RNR versus punishment and 7 times less for RNR versus non-RNR programs. In addition, RNR effects sizes were, on average, much larger in magnitude than either of the comparators.

Notwithstanding the results presented by Smith and Gendreau (2012) and others, caution is warranted about concluding that the reasons low-risk offenders have been adversely affected by prison treatment programs are necessarily due to social learning or differential association explanations. The evidence, albeit on juvenile samples, has to be recognized (Handwerk, Field, & Friman, 2000; Weiss et al., 2005). For example, Weiss et al. (2005) drew attention to some serious flaws in studies that have provided data suggesting iatrogenic effects (e.g., Dishion, McCord, & Poulin, 1999) and, in response, produced a meta-analysis to confirm their supposition. Weiss et al. (2005) did hold out the possibility, however, that group treatment programs might provide inmates with an opportunity to make new friendships. This scenario is consistent with the prison studies that have demonstrated the risk principle among low-risk inmates. Furthermore, when prison programs are ineptly run—in other words, if therapists allow the antisocial values of higher-risk inmates to go unchecked—it is likely that low-risk members of treatment groups will be negatively affected.

SPECIAL CONDITIONS OF CONFINEMENT

Given the history of prison reform from its beginnings and the development of theories that are contrary in their predictions, differences of opinion about the effects of imprisonment have been commonplace. Arguably, a precedent of sorts had been set in this regard in the 1970s, when nothing-works proponents (i.e., Martinson, 1974) heaped scorn, including *ad hominem* attacks, on supporters of the rehabilitative ideal (see Gendreau & Ross, 1987; Palmer, Van Voorhis, Taxman, & MacKenzie, 2012). And as history is wont to do, we find ourselves in such a debate again, this time over the effects of severe forms of incarceration such as administrative segregation (AS), also known as solitary confinement. AS conditions typically include 23 hours per day of lockup with limited amenities and restricted sensory stimulation. The consensus among penologists, with few exceptions (see the review by Gendreau & Thériault, 2011), is that AS environments are psychologically destructive (Grassian, 2010; Haney, 2003). As Kupers (2008) so vividly put it, “[J]ust about for all prisoners, being held in isolated confinement for longer than three months causes lasting emotional damage if not full blown psychosis and functional disability” (p. 1006).

The recent publication of O’Keefe, Klebe, Strucker, Sturm, and Leggett’s (2010) study of the effects of AS within the Colorado state prison system has challenged the prevailing orthodoxy. First, some remarks on the study itself. Bearing in mind that no single study is “perfect,” the methodological standards set by O’Keefe et al.’s evaluation are of a caliber that has not been seen in segregation studies for approximately 40 years, a time when true experimental designs were used (e.g., Eccelstone, Gendreau, & Knox, 1974; Gendreau, Freedman, Wilde, & Scott, 1972). Consider the care that went into its design. Several external experts were assigned to the project to advise and oversee the research team. The study was

a quasi-experimental repeated-measures comparison group design. Inmates in AS were assessed five times over the course of a year, by far the lengthiest assessment period of which we are aware in the AS literature. The sample size was considerably larger ($n = 247$) than that used in other segregation studies of this kind. The measures used to chart inmates' psychopathology are among the field's most reliable and well validated (see O'Keefe et al., 2010, appendix B).

The results of the study were a surprise even to its investigators and were far removed from the predictions of those such as Kupers (2008). The Colorado team noted that, although a small percentage (i.e., 7%) of AS inmates may have been adversely affected, 20% showed improvements in level of functioning, and the remainder were stable throughout the period of confinement (Metzner & O'Keefe, 2011). Attempts to identify factors that were predictive of iatrogenic effects were unsuccessful. Reaction to these findings was swift and, from some quarters, rather uncharitable (Grassian, 2010; see also Immarigeon, 2011).

Space does not permit us, nor is it our intention, to conduct a box-score analysis of the various opinions on the Colorado findings. Readers who are interested in examining the minutiae of the debate are invited to peruse the above-cited materials to form their own opinions. Our preference is to move forward and suggest ways in which one may think more clearly about issues and propose some research and clinical tasks that forensic psychologists may find helpful in that regard.

THINKING META-ANALYTICALLY

In the course of our work in corrections, one strategy that has helped us in clarifying issues has been to adopt a meta-analytic mind-set (Gendreau & Smith, 2007). By this we mean not simply conducting meta-analytic studies, albeit that is a part of the process, but taking the broadest possible perspective on an issue and looking beyond the narrow confines of our respective disciplines and/or subspecialties. This has not been done often in corrections, in part because our subject matter has historically been a battleground for competing disciplines, each jockeying for supremacy. And how can it be otherwise, as there are few incentives to consider other sources of data, especially given the powerful professional reinforcers that serve to maintain a parochial focus on issues (i.e., professional perks and reinforcements within small circles of discipline-specific colleagues, journals, and granting agencies, etc.; Gendreau, 1996)? One result of this intellectual egocentrism among corrections-focused disciplines has been the growth of a silo model of knowledge generation that tends to inevitably stunt the knowledge cumulation process, something that should be of concern to correctional professionals of all stripes. Retreating to one's comfort zone is not a defensible intellectual position in the face of the correctional system's inherent challenges.

In addition to the examples provided in the theory section, the AS literature provides some telling illustrations. The Colorado group (O'Keefe et al., 2010) were not aware of the one theory (i.e., importation) that its members might have used in

their defense. Critics of the Colorado study have totally disregarded the voluminous restricted stimulation/sensory deprivation literature (Suedfeld, 1980; Zubek, 1969). This is a sad situation, because a thorough reading of the perceptual adaptation literature cautions one to be prudent in making dramatic proclamations about the negative effects of AS-type environments. For example, the initial findings of the early McGill studies, which put forth the notion that vivid perceptual distortions resulted from restricted environmental stimulation (e.g., Bexton, Heron, & Scott, 1954), were never replicated by those who followed up on the McGill work (see Zubek, 1969). The reasons for this failure to replicate can be found in the classic methodological critique of sensory deprivation research by Orne and Scheibe (1964). They found that response bias and demand characteristics were confounds that could well have contributed to the dramatic effects reported in the McGill studies. Gendreau and Thériault (2011) have provided evidence of the same confounds in recent AS studies.

Then there is the “theory,” or ways of thinking about knowledge cumulation, behind meta-analytic attitudes toward data collection (Cumming, 2011). All of the bloviation about the import of O’Keefe et al.’s (2010) study runs counter to the *raison d’être* of the knowledge cumulation process. As much as we may champion the Colorado study, we freely acknowledge that the information to be gained from a single study is finite, no matter how well it has been conducted (Schmidt, 1992). Additionally, given the sample sizes in most psychological research (e.g., $n < 100$; Schmidt, 1996), estimates of effect sizes are often embarrassingly imprecise based on the width of the associated CIs (Cohen, 1994). Of note, O’Keefe et al. (2010) did not report CIs.

Sound policy is best developed from repeated replications of a phenomenon (Hunter & Schmidt, 2004). It is absolutely imperative, therefore, that the Colorado study be replicated, particularly within other state systems, if for no other reason than that the level of care for AS inmates in Colorado may differ from that which we expect is the norm elsewhere (see Metzner & O’Keefe, 2011, p.12).

REDUCING THE USE OF AS

Regrettably, it is a dismal reality that the quintessential AS environment, the supermax prison—a primitive and unimaginative response to inmate management (Gendreau, 2012)—is likely with us for the foreseeable future, given the appeal of get-tough policies in political circles. In that light, we suggest four research/clinical agendas to mitigate against the use of AS (see Gendreau & Thériault, 2011).

First, since the literature on inmate tolerance of AS environments is almost nonexistent (Gendreau & Bonta, 1984; Zubek, 1969), jurisdictions that use AS should examine their historical file data and correlate adjustment to AS with whatever risk measures and clinical information they have available. This would be a first step toward developing a vulnerability-to-AS risk measure to identify which inmates

will react negatively to such conditions. Further, prison systems should implement the psychological measures that the Colorado system employed to assess the level of functioning of inmates in AS.

Second, with respect to mentally disordered inmates, prison authorities must use diagnostic measures with good predictive validities. One such protocol is the Camberwell Family Interview, a measure of expressed emotion, which has demonstrated very robust predictive validities with psychiatric relapse and rehospitalization (Smith, Gendreau, & Goggin, 2007; see Hooley & Parker, 2006, for a useful summary of the issues in administering this measure). Third, since AS usually contains a high proportion of high-risk acting-out inmates, programs that are RNR based should be used to deflect at least some among that group from AS (French & Gendreau, 2006).

Finally, we acknowledge the difficult circumstances that correctional officers work in at times. Working conditions can be less than ideal where programming options are limited, if not nonexistent, and some inmates (e.g., mentally disordered, violent) can be extremely challenging. Nevertheless, in the first author's experience, disruptive inmate behavior often results from the capricious manner in which they are treated. Some of the more common problems have been "simple" things, such as confusion regarding the criteria for being placed in AS or being released from it, changes in shift work patterns so that awareness of the psychological makeup of inmates and continuity of care is disrupted, a lack of even elementary activities to occupy inmates' time, and the imposition of petty rules that have little to do with security and lead to unnecessary provocations (Arbour, 2006; Gendreau & Thériault, 2011; Human Rights Watch, 2000).

CONCLUDING REMARKS

We opened this chapter with a brief history of prisons followed by a review of the three major theories of its effects on offender behavior. Our purpose in doing so was to lay the groundwork for a discussion of what works in addressing the overriding demands of penology; that is, the protection of the public through the safe and humane management of prisons. To that end, we reviewed the existing research record regarding the requisites for the accurate assessment and classification of inmates as well as the key components of effective correctional programming.

That said, the transfer of this body of knowledge from the lab to operational reality continues to be a challenge (i.e., Gendreau, Goggin, & Smith, 1999). As noted, the effects of real-world programs are markedly attenuated relative to the results of "demonstration" projects. This has prompted some researchers to examine aspects of program administration that may be linked to these differences (i.e., Magaletta, Morgan, Reitzel, & Innes, 2007; Wormith et al., 2009). It has also, however, contributed to a growing disenchantment with the role of rehabilitation in correctional policy, a trend that Cullen, Smith, Lowenkamp, and Latessa (2009) warn

is infiltrating the “marketplace of ideas” (p. 112). In keeping with this, Gendreau et al. (2009) have documented the disquieting popularity of offender programs founded upon common sense rather than empirical evidence (i.e., boot camps, pet therapy, acupuncture, etc.). They maintain that such correctional quackery has the potential to thwart knowledge transfer and hinder the practical application of effective offender treatment programs (Gendreau et al., 2009). Without a sustained commitment to advancing our knowledge in this area, we run the risk of having the results from the previous 30 years of research discounted by the vagaries of sociopolitical climate. The historical record is clear on just how readily this can happen in a propitious context (i.e., Cullen & Gendreau, 2000).

Despite these concerns, we are cautiously optimistic about the potential for prisons to be managed effectively, if correctional systems adhere to the evidence. The data is unequivocal: Among settings that have embraced the principles of RNR in their operational policies, reductions in rates of misconducts and postrelease recidivism are well documented. In contrast, based on the results of CPAI-2010[©] (Gendreau et al., 2010) evaluations, we have noted higher rates of both types of outcome among inmates serving time in correctional settings that do not incorporate such principles.

Certainly, additional research objectives remain. These include improving our knowledge of how to more successfully implement correctional programs, the development of specific treatment curricula, adding to the knowledge base on effective correctional programming for diverse inmate groups (i.e., females, Aborigines, youth, mentally disordered, etc.), and working to ensure that correctional policy is evidenced based. To that end, we reiterate our commitment to further elaboration and application of the principles of sound correctional practice in order to ensure the goal of public protection through the safe and humane management of prisons.

REFERENCES

- Abbott, E. (1927). The Civil War and the crime wave of 1865–1870. *Social Service Review*, 1, 212–234.
- Agnew, R. (2006). *Pressured into crime: An overview of general strain theory*. Los Angeles, CA: Roxbury.
- Akers, R. L. (1977). Type of leadership in prison: A structural approach to testing the functional and importation models. *Sociological Quarterly*, 18, 378–383.
- Allen, F. A. (1981). *The decline of the rehabilitative ideal: Penal policy and social purpose*. New Haven, CT: Yale University Press.
- American Correctional Association. (1972). Development of modern correctional concepts and standards. In R. M. Carter, D. Glaser, & L. T. Wilkins (Eds.), *Correctional institutions* (pp. 17–34). New York, NY: Lippincott.
- Andrews, D. A. (1980). Some experimental investigations of the principles of differential association through deliberate manipulations of the structure of service systems. *American Sociological Review*, 45, 448–462.

- Andrews, D. A., & Bonta, J. (1995). *The Level of Service Inventory–Revised*. Toronto, Canada: Multi-Health Systems.
- Andrews, D. A., & Bonta, J. (2003). *The psychology of criminal conduct* (3rd ed.). Cincinnati, OH: Anderson.
- Andrews, D. A., & Bonta, J. (2010a). *The psychology of criminal conduct* (5th ed.). Cincinnati, OH: Anderson.
- Andrews, D. A., & Bonta, J. (2010b). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law*, *16*, 39–55.
- Andrews, D. A., & Wormith, S. (1989). Personality and crime: Knowledge destruction and construction in criminology. *Justice Quarterly*, *6*, 289–310.
- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A psychologically informed meta-analysis. *Criminology*, *28*, 369–404.
- Arbour, L. (2006). *Commission of inquiry into certain events at the Prison for Women in Kingston*. Ottawa, ON, Canada: Public Works and Government Services of Canada.
- Armstrong, G. S., & MacKenzie, D. L. (2003). Private versus public juvenile correctional facilities: Do differences in environmental quality exist? *Crime and Delinquency*, *49*, 542–563.
- Bandura, A. (1973). *Aggression: A social learning analysis*. Oxford, England: Prentice-Hall.
- Banister, P. A., Smith, F. V., Heskin, K. J., & Bolton, N. (1973). Psychological correlates of long-term imprisonment: I. Cognitive variables. *British Journal of Criminology*, *13*, 312–323.
- Beccaria, C. (1986). *On crimes and punishments* (trans. D. Young). Indianapolis, IN: Hackett. (Original work published 1764)
- Bexton, W. H., Heron, W., & Scott, T. H. (1954). Effects of decreased variation in the sensory environment. *Canadian Journal of Psychology*, *8*, 70–77.
- Bolton, N., Smith, F. V., Heskin, K. J., & Banister, P. A. (1976). Psychological correlates of long-term imprisonment: IV. A longitudinal analysis. *British Journal of Criminology*, *16*, 38–47.
- Bonta, J., & Gendreau, P. (1990). Re-examining the cruel and unusual punishment of prison life. *Law and Human Behavior*, *14*, 347–372.
- Bonta, J., & Gendreau, P. (1992). Coping with prison. In P. Suedfeld & P. E. Tetlock (Eds.), *Psychology and social policy* (pp. 343–354). New York, NY: Hemisphere.
- Bonta, J., & Motiuk, L. (1992). Inmate classification. *Journal of Criminal Justice*, *20*, 343–353.
- Bonczar, T. (2003). *Prevalence of imprisonment in the U.S. population, 1974–2001*. Washington, DC: Bureau of Justice Statistics, U.S. Department of Justice.
- Bourgon, G., & Armstrong, B. (2005). Transferring the principles of effective treatment into a “real world” prison setting. *Criminal Justice and Behavior*, *32*, 3–25.
- Briggs, C. S., Sundt, J. L., & Castellano, T. C. (2003). The effect of supermaximum security prisons on aggregate levels of institutional violence. *Criminology*, *41*, 1341–1376.
- Brockway, Z. R. (1871). The ideal of a true prison system for a state. In E. C. Wines (Ed.), *Transactions of the National Congress on Penitentiary and Reformatory Discipline* (pp. 38–65). Albany, NY: Weed, Parsons.
- Buehler, R. E., Patterson, G. R., & Furniss, J. M. (1966). The reinforcement of behavior in institutional settings. *Behavior Research and Therapy*, *4*, 157–167.
- Bukstel, L. H., & Kilmann, P. R. (1980). Psychological effects of imprisonment on confined individuals. *Psychological Bulletin*, *88*, 469–493.

- Bush, J. (2005). Teaching self-risk management to violent offenders. In J. McGuire (Ed.), *What works: Reducing reoffending—Guidelines from research and practice* (pp. 140–154). Hoboken, NJ: Wiley.
- Camp, S., Gaes, G., Langan, N., & Saylor, W. (2003). The influence of prisons on inmate misconduct: A multilevel investigation. *Justice Quarterly*, *20*, 501–533.
- Campbell, M. A., French, S., & Gendreau, P. (2009). The prediction of violence in adult offenders: A meta-analytic comparison. *Criminal Justice and Behavior*, *36*, 567–590.
- Carrigan, D. O. (1991). *Crime and punishment in Canada, a history*. Toronto, Canada: McClelland & Stewart.
- Cellard, A. (2000). *Punishment, imprisonment and reform in Canada, from New France to the present* (Historical Booklet No. 60). Ottawa, Canada: Canadian Historical Association.
- Clements, C. B., Althouse, R., Ax, R. K., Magaletta, P. R., Fagan, T. J., & Wormith, J. S. (2007). Systemic issues and correctional outcomes: Expanding the scope of correctional psychology. *Criminal Justice and Behavior*, *34*, 919–932.
- Clemmer, D. (1940). *The prison community*. New York, NY: Holt, Rinehart, & Winston.
- Clemmer, D. (1950). Observations of imprisonment as a source of criminality. *Journal of Criminal Law and Criminology*, *41*, 311–319.
- Cochran, J. C., Mears, D. P., Bales, W. D., & Stewart, E. A. (2012). Does inmate behavior affect post-release offending? Investigating the misconduct-recidivism relationship among youth and adults. *Justice Quarterly*, 1–30.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, *49*, 997–1003.
- Cohen, M., & Piquero, A. (2009). New evidence on the monetary value of saving a high risk youth. *Journal of Quantitative Criminology*, *25*, 25–49.
- Colvin, M. (2000). *Crime and coercion: An integrated theory of chronic criminality*. New York, NY: St. Martin's Press.
- Cullen, F. T., Fisher, B. S., & Applegate, B. K. (2000). Public opinion about punishment and corrections. In M. Tonry (Ed.), *Crime and justice: A review of research* (Vol. 27, pp. 1–79). Chicago, IL: University of Chicago Press.
- Cullen, F. T., & Gendreau, P. (2000). Assessing correctional rehabilitation: Policy, practice, and prospects. In J. Horney (Ed.), *NIJ Criminal justice 2000: Changes in decision making and discretion in the criminal justice system*. Washington, DC: U.S. Department of Justice, National Institute of Justice.
- Cullen, F. T., & Gendreau, P. (2001). From nothing works to what works: Changing professional ideology in the 21st century. *Prison Journal*, *81*, 313–338.
- Cullen, F. T., Smith, P., Lowenkamp, C. T., & Latessa, E. J. (2009). Nothing works revisited: Deconstructing Farabee's Rethinking Rehabilitation. *Victims and Offenders*, *4*, 101–123.
- Cumming, G. (2011). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge Academic.
- Davidson, W., Gottschalk, R., Gensheimer, L., & Mayer, J. (1984). *Interventions with juvenile delinquents: A meta-analysis of treatment efficacy*. Washington, DC: National Institute of Juvenile Justice and Delinquency Prevention.
- Davidson, W. S., Jefferson, S. D., Legaspi, A., Lujan, J., & Wolf, A. M. (2001). Alternative interventions for juvenile offenders: History of the adolescent diversion project. In C. R. Hollin (Ed.), *Handbook of offender assessment* (pp. 221–236). Chichester, England: Wiley.

- Davidson, W. S., Redner, R., Blakely, C. H., Mitchell, C. M., & Emshoff, J. G. (1987). Diversion of juvenile offenders: An experimental comparison. *Journal of Consulting and Clinical Psychology, 55*, 68–75.
- de Beaumont, G., & de Tocqueville, A. (1970). *On the penitentiary system in the United States and its application in France* (trans. F. Lieber). New York, NY: A. M. Kelly. (Original work published 1833)
- DeJong, C. (1997). Survival analysis and specific deterrence: Integrating theoretical and empirical models of recidivism. *Criminology, 35*, 561–575.
- DeLisi, M., & Walters, G. D. (2011). Multiple homicide as a function of prisonization and concurrent instrumental violence: Testing an interactive model—A research note. *Crime and Delinquency, 57*, 147–161.
- Dickens, C. (1842). *American notes for general circulation*. Paris, France: Baudry's European Library.
- Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54*, 755–764.
- Dix, D. (1967). *Remarks on prisons and prison discipline in the United States*. Montclair, NJ: Patterson Smith. (Original work published 1845)
- Drake, E., Aos, S., & Miller, M. (2009). Evidence-based public policy options to reduce crime and criminal justice costs: Implications in Washington state. *Victims and Offenders, 4*, 170–196.
- Dvoskin, J. A., Skeem, J. L., Novaco, R. W., & Douglas, K. S. (2012). *Using social science to reduce violent offending*. New York, NY: Oxford University Press.
- Eagley, A., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.
- Eccelstone, C.E.J., Gendreau, P., & Knox, C. (1974). Solitary confinement of prisoners: An assessment of its effects on inmates' personal constructs and adrenocortical activity. *Canadian Journal of Behavioural Science, 6*, 178–191.
- Erdahl, A. (2001). History of corrections in Canada. In J. Winterdyk (Ed.), *Corrections in Canada: Social reactions to crime* (pp. 27–48). Toronto, Canada: Prentice Hall.
- Farrington, D. P., Petrosino, A., & Welsh, B. C. (2001). Systematic reviews and cost-benefit analysis of correctional interventions. *Prison Journal, 81*, 339–359.
- Finn, P. (1996). No-frills prison and jails: A movement in flux. *Federal Probation, 60*, 35–44.
- Fishbein, M. (1995). Developing effective behavior change interventions: Some lessons learned from behavioral research. In T. E. Backer, S. L. David, & G. Soucy (Eds.), *Reviewing the behavioral science knowledge base on technology transfer* (NIDA Research Monograph No. 155, pp. 246–261). Rockville, MD: National Institute on Drug Abuse.
- Flanagan, T. (1980). Time served and institutional misconduct patterns of involvement in disciplinary infractions among long-term and short-term inmates. *Journal of Criminal Justice, 8*, 357–367.
- French, S., & Gendreau, P. (2006). Reducing prison misconducts: What works! *Criminal Justice and Behavior, 33*, 185–218.
- Gaes, G. G., & Camp, S. D. (2009). Unintended consequences: Experimental evidence for the criminogenic effect of security level placement on post-release recidivism. *Journal of Experimental Criminology, 5*, 139–162.

- Garrett, C. J. (1985). Effects of residential treatment of adjudicated delinquents: A meta-analysis. *Journal of Research in Crime and Delinquency*, 22, 287–308.
- Gendreau, P. (1996). Offender rehabilitation: What we know and what needs to be done. *Criminal Justice and Behavior*, 23, 144–161.
- Gendreau, P. (2012, October 18). *Everything you wanted to know about prisons*. Keynote address to the Association for the Treatment of Sexual Abuse. Denver, CO.
- Gendreau, P., Andrews, D. A., & Thériault, Y. (2010). *Correctional Program Assessment Inventory-2010*[©] (CPAI-2010[©]). Beresford, New Brunswick, Canada: Author.
- Gendreau, P., & Bonta, J. (1984). Solitary confinement is not cruel and unusual punishment: Sometimes people are! *Canadian Journal of Criminology*, 26, 467–478.
- Gendreau, P., Freedman, N. L., Wilde, G. J., & Scott, G. D. (1972). Changes in EEG alpha frequency and evoked response latency during solitary confinement. *Journal of Abnormal Psychology*, 79, 54–59.
- Gendreau, P., Gibson, M., Surridge, C. T., & Hug, J. J. (1973). The application of self-esteem measures in corrections: A further report on the SEI. *Journal of Community Psychology*, 1, 423–425.
- Gendreau, P., Goggin, C., & Cullen, F. T. (1999). *The effects of prison sentences on recidivism*. Ottawa, Canada: Solicitor General Canada.
- Gendreau, P., Goggin, C., French, S., & Smith, P. (2006). Practicing psychology in correctional settings. In A. K. Hess & I. B. Weiner (Eds.), *The handbook of forensic psychology* (3rd ed., pp. 722–750). Hoboken, NJ: Wiley.
- Gendreau, P., Goggin, C., & Law, M. A. (1997). Predicting prison misconducts. *Criminal Justice and Behavior*, 24, 414–431.
- Gendreau, P., Goggin, C., & Smith, P. (1999). The forgotten issue in effective correctional treatment: Program implementation. *International Journal of Offender Therapy and Comparative Criminology*, 43, 180–187.
- Gendreau, P., Grant, B. A., & Leipziger, M. (1979). Self-esteem, incarceration and recidivism. *Criminal Justice and Behavior*, 6, 67–75.
- Gendreau, P., & Keyes, D. (2001). Making prisons safer and more humane environments. *Canadian Journal of Criminology*, 43, 123–130.
- Gendreau, P., Listwan, S. J., & Kuhns, J. B. (2013). *Making prisoners accountable: The potential of contingency management programs* (Unpublished manuscript).
- Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology*, 34, 575–607.
- Gendreau, P., & Ross, R. R. (1981). Correctional potency: Treatment and deterrence on trial. In R. Roesch & R. R. Corrado (Eds.), *Evaluation and Criminal Justice Policy* (pp. 29–57). Beverly Hills, CA: Sage.
- Gendreau, P., & Ross, R. R. (1987). Revivification of rehabilitation: Evidence from the 1980s. *Justice Quarterly*, 4, 349–408.
- Gendreau, P., & Smith, P. (2007). Influencing the “people who count”: Some perspectives on the reporting of meta-analytic results for prediction and treatment outcomes with offenders. *Criminal Justice and Behavior*, 34, 1536–1559.
- Gendreau, P., & Smith, P. (2012). Assessment and treatment strategies for correctional institutions. In J. A. Dvoskin, J. L. Skeem, & R. W. Novaco (Eds.), *Using social science to reduce violent offending* (pp. 157–177). New York, NY: Oxford University Press.

- Gendreau, P., Smith, P., & Thériault, Y. (2009). Chaos theory and correctional treatment: Common sense, correctional quackery, and the law of fartcatchers. *Journal of Contemporary Criminal Justice*, 25, 384–396.
- Gendreau, P., & Suboski, M. D. (1971). Classical discrimination eyelid conditioning in primary psychopaths. *Journal of Abnormal Psychology*, 77, 242–246.
- Gendreau, P., & Thériault, Y. (2011). Bibliotherapy for cynics revisited: Commentary on “one year long study of the psychological effects of administrative segregation.” *Correctional Mental Health*, 13, 1–11.
- Gibbons, J. J., & Katzenbach, N. B. (2006). *Confronting confinement: A report of the Commission on Safety and Abuse in America’s prisons*. New York, NY: Vera Institute of Justice.
- Glaser, D., & Stratton, J. R. (1972). Measuring inmate change in prison. In D. R. Cressey (Ed.), *Prison studies in institutional organization and change* (pp. 381–392). New York, NY: Holt, Rinehart & Winston.
- Glueck, S., & Glueck, E. T. (1965). *Five hundred criminal careers*. New York, NY: Kraus Reprint Corp. (Original work published 1939)
- Goggin, C. (2008). *Is prison “personality” associated with offender recidivism?* (Unpublished doctoral dissertation). University of New Brunswick, Fredericton, New Brunswick, Canada.
- Goggin, C., & Gendreau, P. (2006). The implementation and maintenance of quality services in offender rehabilitation programmes. In C. Hollin & E. Palmer (Eds.), *Offending behaviour programmes: Development, application, and controversies* (pp. 209–246). Chichester, England: Wiley.
- Goldstein, A. (1999). *The prepare curriculum: Teaching prosocial competencies*. Champaign, IL: Research Press.
- Grassian, S. (2010, November 15). Fatal flaws in the Colorado solitary confinement study. <http://solitarywatch.com/2010/11/15/fatal-flaws-in-the-colorado-solitary-confinement-study/>
- Guilford, J. P. (1946). *General psychology* (6th ed.). New York, NY: D. Van Nostrand.
- Handwerk, M. L., Field, C. E., & Friman, P. C. (2000). The iatrogenic effects of group intervention for antisocial youth: Premature extrapolations. *Journal of Behavioral Education*, 10, 223–238.
- Haney, C. (1997). Psychology and the limits to prison pain: Confronting the coming crisis in Eighth Amendment law. *Psychology, Public Policy, and Law*, 3, 499–588.
- Haney, C. (2003). Mental health issues in long-term solitary and “supermax” confinement. *Crime and Delinquency*, 49, 124–156.
- Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior*, 36, 865–891.
- Hanson, R. K., & Morton-Bourgon, K. E. (2009). Accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessments*, 21, 1–21.
- Hare, R. D. (1996). Psychopathy: A clinical construct whose time has come. *Criminal Justice and Behavior*, 23, 25–54.
- Hare, R. D. (2003). *The Hare Psychopathy Checklist: Revised (PCL-R)* (2nd ed.). Toronto, Canada: Multi-Health Systems.

- Hart, S. D., Cox, D. N., & Hare, R. D. (1995). *The Hare Psychopathy Checklist: Screening Version (PCL-SV)*. Toronto, Canada: Multi-Health Systems.
- Henning, K. R., & Frueh, B. C. (1996). Cognitive-behavioral treatment of incarcerated offenders: An evaluation of the Vermont Department of Corrections' Cognitive Self-Change program. *Criminal Justice and Behavior, 23*, 523–541.
- Heskin, K. J., Bolton, N., Smith, F. V., & Banister, P. A. (1974). Psychological correlates of long-term imprisonment: III. Attitudinal variables. *British Journal of Criminology, 14*, 150–157.
- Heskin, K. J., Smith, F. V., Banister, P. A., & Bolton, N. (1973). Psychological correlates of long-term imprisonment: II. Personality variables. *British Journal of Criminology, 13*, 323–330.
- Hirschi, T. (1969). *Causes of delinquency*. Berkeley: University of California Press.
- Hooley, J. M., & Parker, H. A. (2006). Measuring expressed emotion: An evaluation of the shortcuts. *Journal of Family Psychology, 20*, 386–396.
- Howard, J. (1991). The state of the prisons. In J. Muncie & R. Sparks (Eds.), *Imprisonment: European perspectives* (pp. 7–29). New York, NY: St. Martin's Press.
- Human Rights Watch. (2000). *Out of sight: Supermaximum security confinement in the United States*. New York, NY: Author.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.
- Immarigeon, R. (Ed.). (June 20, 2011). *Corrections and Mental Health, 20*. Retrieved from <http://community.nicic.gov/blogs/mentalhealth/archive/2011/06.aspx>
- Ives, G. (1970). *A history of penal methods: Criminals, witches, lunatics*. Montclair, NJ: Patterson Smith. (Original work published 1914)
- Jaman, D. R., Dickover, R. M., & Bennett, L. A. (1972). Parole outcome as a function of time served. *British Journal of Criminology, 12*, 5–34.
- Johnston, N. (2000). *Forms of constraint: A history of prison architecture*. Chicago: University of Illinois Press.
- Jonson, C. L. (2010). *The impact of imprisonment of reoffending: A meta-analysis* (Unpublished doctoral dissertation). Retrieved from http://cech.uc.edu/content/dam/cech/programs/criminaljustice/docs/phd_dissertations/Jonson-Cheryl-Lero.pdf
- Jonson, C. L. (2013). The effects of imprisonment. In F. T. Cullen & P. Wilcox (Eds.), *The Oxford handbook of criminological theory* (pp. 672–690). New York, NY: Oxford University Press.
- Keyes, D. (1996). *Preventing prison misconduct behaviour: A quantitative review of the literature* (Unpublished master's thesis). University of New Brunswick, Saint John, New Brunswick, Canada.
- Kupers, T. A. (2008). What to do with the survivors? Coping with the long-term effects of isolated confinement. *Criminal Justice and Behavior, 35*, 1005–1016.
- Lab, S. P., & Whitehead, J. T. (1990). From “nothing works” to “the appropriate works”: The latest stop on the search for the secular grail. *Criminology, 28*, 405–417.
- Latessa, E. J., Lovins, L. B., & Smith, P. (2010). *Final report: Follow-up evaluation of Ohio's community based correctional facility and halfway house programs—Outcome study*. Cincinnati, OH: University of Cincinnati, Division of Criminal Justice.

- Latessa, E. J., Lovins, L. B., Smith, P., & Makarios, M. (2010). *Follow-up evaluation of Ohio's community based correctional facility and halfway house programs: Program characteristics supplemental report*. Cincinnati, OH: University of Cincinnati, Division of Criminal Justice.
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York, NY: Springer.
- Lemert, E. M. (1951). *Social pathology: A systematic approach to the theory of sociopathic behavior*. New York, NY: McGraw-Hill.
- Lipsey, M. W. (1995). What do we learn from 400 research studies on the effectiveness of treatment with juvenile delinquents? In J. McGuire (Ed.), *What works: Reducing reoffending—Guidelines for research and practice* (pp. 63–78). Oxford, England: Wiley.
- Lipsey, M. W. (1999). Can rehabilitative programs reduce the recidivism of juvenile offenders? An inquiry into the effectiveness of practical programs. *Virginia Journal of Social Policy and Law*, 6, 611–641.
- Listwan, S. J., Colvin, M., Hanley, D., & Flannery, D. (2010). Victimization, social support and psychological well-being: A study of recently released prisoners. *Criminal Justice and Behavior*, 3, 1140–1159.
- Listwan, S. J., Sullivan, C. J., Agnew, R., Cullen, F. T., & Colvin, M. (2013). The pains of imprisonment revisited: The impact of strain on inmate recidivism. *Justice Quarterly*, 30, 1–25.
- Lovell, D., & Jemelka, R. (1996). When inmates misbehave: The costs of discipline. *Prison Journal*, 76, 165–179.
- Lowenkamp, C. T., Latessa, E. J., & Smith, P. (2006). Does correctional program quality really matter? The impact of adhering to the principles of effective intervention. *Criminology and Public Policy*, 5, 575–594.
- MacKenzie, D. L., & Goodstein, L. (1985). Long-term incarceration impacts and characteristics of long-term offenders: An empirical analysis. *Criminal Justice and Behavior*, 12, 395–414.
- Magaletta, P. R., Morgan, R. D., Reitzel, L. R., & Innes, C. A. (2007). Toward the one: Strengthening behavioural sciences research in corrections. *Criminal Justice and Behavior*, 34, 933–944.
- Marlatt, G. A., & Gordon, J. R. (1985). Relapse prevention: Theoretical rationale and overview of the model. In G. A. Marlatt & J. R. Gordon (Eds.), *Relapse prevention: Maintenance strategies in the treatment of addictive behaviors* (pp. 3–70). New York, NY: Guilford Press.
- Martinson, R. (1974). What works? Questions and answers about prison reform. *Public Interest*, 35, 22–54.
- Mason, P. (1998). The prison in cinema. *Images: A Journal of Film and Popular Culture*, 6. Retrieved from <http://www.imagesjournal.com/issue06/features/prison3.htm>
- Masters, J. C., Burish, T. G., Hollon, S. D., & Rimm, D. C. (1987). *Behavior therapy: Techniques and empirical findings*. San Diego, CA: Harcourt Brace Jovanovich.
- Matson, J. L., & Dilorenzo, T. M. (1984). *Punishment and its alternatives*. New York, NY: Springer.
- McGuire, J. (1995, November). The death of deterrence. In J. McGuire & B. Rowson (Eds.), *Does punishment work? Proceedings of a conference held at Westminster Central Hall, London, England*.

- Meskill, M. W. (1999). An American resolution: The history of prisons in the United States from 1777 to 1877. *Stanford Law Review*, *51*, 839–865.
- Metzner J. L., & O'Keefe, M. L. (2011). Psychological effects of administrative segregation: The Colorado study. *Corrections Mental Health Report*, *13*, 1–2, 13–14.
- Milan, M. A. (1987). Token economy programs in closed institutions. In E. K. Morris & C. J. Braukmann (Eds.), *Behavioral approaches to crime and delinquency: A handbook of application, research, and concepts* (pp. 195–222). New York, NY: Plenum Press.
- Moos, R. H. (1968). The assessment of social climate in correctional institutions. *Journal of Research in Crime and Delinquency*, *5*, 174–188.
- Morgan, R. D., & Flora, D. B. (2002). Group psychotherapy with incarcerated offenders: A research synthesis. *Group Dynamics: Theory, Research, and Practice*, *6*, 203–218.
- Motiuk, L. L. (1991). *Antecedents and consequences of prison adjustment: A systematic assessment and reassessment approach* (Unpublished doctoral dissertation). Carleton University, Ottawa, Ontario, Canada.
- Mowrer, O. H. (1960). *Learning theory and behavior*. New York, NY: Wiley.
- Nagin, D. S. (1998). Criminal deterrence research at the outset of the twenty-first century. In M. Tonry (Ed.), *Crime and justice: A review of research* (Vol. 23, pp. 1–42). Chicago, IL: University of Chicago Press.
- Nagin, D. S., Cullen, F. T., & Jonson, C. L. (2009). Imprisonment and reoffending. In M. Tonry (Ed.), *Crime and justice: A review of research* (Vol. 38, pp. 115–200). Chicago, IL: University of Chicago Press.
- Nossiter, A. (1994, September 17). Making hard time harder: States cut jail TV and sports. *New York Times*, pp. A1, A10.
- O'Keefe, M. L., Klebe, K. J., Stucker, A., Sturm, K., & Leggett, W. (2010). *One year longitudinal study of the psychological effects of administrative segregation*. Colorado Springs, CO: Colorado Department of Corrections. Retrieved from http://www.doc.state.co.us/sites/default/files/opa/AdSegReport_2010.pdf
- Oliver, P. (1998). *"Terror to evil-doers": Prisons and punishments in nineteenth-century Ontario*. Toronto, Canada: University of Toronto Press.
- Orne, M. T., & Scheibe, K. G. (1964). The contribution of non-deprivation factors in the production of sensory deprivation effects: The psychology of the "panic button." *Journal of Abnormal Psychology*, *68*, 3–12.
- Orsagh, T., & Chen, J.-R. (1988). The effect of time served on recidivism: An interdisciplinary theory. *Journal of Quantitative Criminology*, *4*, 155–171.
- Palmer, T. (1994). *A profile of correctional effectiveness and new directions for research*. Albany: State University of New York Press.
- Palmer, T., Van Voorhis, P., Taxman, F. S., & MacKenzie, D. L. (2012). Insights from Ted Palmer: Experimental criminology in a different era. Academy of Experimental Criminology 2011 Joan McCord Prize lecture. *Journal of Experimental Criminology*, *8*, 103–115.
- Phillipson, C. (1970). *Three criminal law reformers: Beccaria, Bentham, Romilly*. Montclair, NJ: Patterson Smith. (Original work published 1923)
- Piehl, A. M., & Useem, B. (2011). Prisons. In J. Q. Wilson & J. Petersilia (Eds.), *Crime and public policy* (pp. 532–558). New York, NY: Oxford University Press.
- Putney, S., & Putney, G. J. (1962). Origins of reformatory. *Journal of Criminal Law, Criminology, and Police Science*, *53*, 437–445.

- Rasch, W. (1981). The effects of indeterminate detention: A study of men sentenced to life imprisonment. *International Journal of Law and Psychiatry*, 4, 417–431.
- Richards, B. (1978). The experience of long-term imprisonment. *British Journal of Criminology*, 18, 162–169.
- Roberts, J. V., Crutcher, N., & Verbrugge, P. (2007). Public attitudes to sentencing in Canada: Exploring recent findings. *Canadian Journal of Criminology and Criminal Justice*, 49, 75–107.
- Romani, C. J., Morgan, R. D., Gross, N. R., & McDonald, B. R. (2012). Treating criminals behavior: Is the bang worth the buck? *Psychology, Public Policy, and Law*, 18, 144–165.
- Rothman, D. J. (1998a). The invention of the penitentiary. In T. J. Flanagan, J. W. Marquart, & K. G. Adams (Eds.), *Incarcerating criminals: Prisons and jails in social and organizational context* (pp. 15–23). London, England: Oxford University Press.
- Rothman, D. J. (1998b). Perfecting the prison: United States, 1789–1865. In N. Morris & D. J. Rothman (Eds.), *The Oxford history of the prison* (pp. 100–116). Oxford, England: Oxford University Press.
- Rotman, E. (1990). *Beyond punishment: A new view of rehabilitation of criminal offenders*. New York, NY: Greenwood Press.
- Rotman, E. (1998). The failure of reform: United States, 1865–1965. In N. Morris & D. J. Rothman (Eds.), *The Oxford history of the prison* (pp. 151–177), Oxford, U.K.: Oxford University Press.
- Sadler, C. M., & Powell, T. A. (2008, June). *The Vermont Cognitive Self-Change program: The case for risk-adjusted classification*. Proceedings of the North American Correctional and Criminal Justice Psychology Conference, Toronto, Ontario, Canada.
- Sapsford, R. J. (1978). Life-sentence prisoners: Psychological changes during sentence. *British Journal of Criminology*, 18, 128–145.
- Scalia, M. B. (1871). Historical sketch of national and international penitentiary conferences in Europe and America. In E. C. Wines (Ed.), *Transactions of the National Congress on Penitentiary and Reformatory Discipline* (pp. 267–277). Albany, NY: Weed, Parsons.
- Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist*, 47, 1173–1181.
- Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for the training of researchers. *Psychological Methods*, 1, 115–129.
- Schnur, A. C. (1949). Prison conduct and recidivism. *Journal of Criminal Law and Criminology*, 40, 36–42.
- Sellin, T. (1967). A look at prison history. *Federal Probation*, 31, 18–23.
- Simpson, J. (2012, May 18). “Time and punishment” now Canada’s way. *Globe and Mail*. Retrieved from <http://www.theglobeandmail.com/commentary/time-and-punishment-now-canadas-way/article4184777/>
- Smith, P., & Gendreau, P. (2012). *Treatment programs in prisons: Prison adjustment, recidivism and the risk hypothesis* (Unpublished manuscript).
- Smith, P., Gendreau, P., & Goggin, C. (2007). “What works” in predicting psychiatric hospitalization and relapse: The specific responsivity dimension of effective correctional treatment for mentally disordered offenders. In R. Ax & T. Fagan (Eds.), *Corrections, mental, and social policy: International perspectives* (pp. 209–233). Springfield, IL: Charles C Thomas.

- Smith, P., Gendreau, P., & Swartz, K. (2009). Validating the principles of effective intervention: A systematic review of the contributions of meta-analysis in the field of corrections. *Victims and Offenders, 4*, 148–169.
- Smith, P., Goggin, C., & Gendreau, P. (2002). *The effects of prison sentences and intermediate sanctions on recidivism: General effects and individual differences*. Ottawa, Canada: Solicitor General Canada.
- Spiegler, M. D., & Guevremont, D. C. (2010). *Contemporary behavior therapy* (5th ed.). Belmont, CA: Wadsworth.
- Steiner, B., & Wooldredge, J. (2008). Inmate versus environmental effects on prison rule violations. *Criminal Justice and Behavior, 35*, 438–456.
- Steiner, B., & Wooldredge, J. (2009). Rethinking the link between institutional crowding and inmate misconduct. *Prison Journal, 89*, 205–233.
- Suedfeld, P. (1980). *Restricted environmental stimulation: Research and clinical applications*. New York, NY: Wiley.
- Sundt, J. L., Castellano, T. C., & Briggs, C. S. (2008). The sociopolitical context of prison violence and its control: A case study of Supermax and its effect in Illinois. *Prison Journal, 88*, 94–122.
- Sykes, G. (1958). *The society of captives*. Princeton, NJ: Princeton University Press.
- Tanasichuk, C. L., Wormith, S. J., & Guzzo, L. (2009). *The predictive validity of the Level of Service Inventory–Ontario Revision (LSI-OR) with Aboriginal offenders*. Unpublished manuscript, Department of Psychology, University of Saskatchewan, Saskatoon, Canada.
- Taylor, I. (2003). *History of the alphabet: Semitic alphabets Part 1*. Whitefish, MT: Kessinger. (Original work published 1899)
- Thomas, C. W. (1977). Theoretical perspectives on prisonization: A comparison of the importation and deprivation models. *Journal of Criminal Law & Criminology, 68*, 135–145.
- Thomas, C., & Foster, S. (1973). The importation model perspective on inmate social roles: An empirical test. *Journal of Criminal Law and Criminology, 68*, 135–145.
- Tomlinson, M. H. (1978). “Prison palaces”: A re-appraisal of early Victorian prisons, 1835–77. *Historical Research, 51*, 60–71.
- Useem, B., & Kimball, P. (1989). *States of siege: U.S. prison riots 1971–1986*. New York, NY: Oxford University Press.
- Villetaz, P., Killias, M., & Zoder, I. (2006). *The effects of custodial vs. non-custodial sentences on re-offending: A systematic review of the state of knowledge*. Oslo, Norway: Campbell Collaboration.
- Webster, C. D., Douglas, K. S., Eaves, S. D., & Hart, S. D. (1997). Assessing risk of violence to others. In C. D. Webster & M. A. Jackson (Eds.), *Impulsivity: Theory, assessment, and treatment* (pp. 251–277). New York, NY: Guilford Press.
- Weiss, B., Caron, A., Ball, S., Tapp, J., Johnson, M., & Weisz, J. R. (2005). Iatrogenic effects of group treatment for antisocial youths. *Journal of Consulting and Clinical Psychology, 73*, 1036–1044.
- Whitehead, J. T., & Lab, S. P. (1989). A meta-analysis of juvenile correctional treatment. *Journal of Research in Crime and Delinquency, 26*, 276–295.
- Wines, E. C. (1968). *The state of prisons and child-saving institutions in the civilized world*. Montclair, NJ: Patterson Smith. (Original work published 1880)
- Wines, E. C., & Dwight, T. W. (1973). *Report on the prisons and reformatories of the United States and Canada, Made to the Legislature of New York, January, 1867*. New York, NY: AMS Press. (Original work published 1867)

- Wooldredge, J. D. (1998). Inmate lifestyles and opportunities for victimization. *Journal of Research in Crime and Delinquency*, 35, 480–502.
- Wormith, J. S. (1984). The controversy over the effects of long-term imprisonment. *Canadian Journal of Criminology*, 26, 423–437.
- Wormith, J. S. (1986, August). *The effects of incarceration: Myth-busting in criminal justice*. Paper presented at the 94th Annual Convention of the American Psychological Association, Washington, DC.
- Wormith, J. S., Althouse, R., Simpson, M., Reitzel, L. R., Fagan, T. J., & Morgan, R. D. (2009). The rehabilitation and reintegration of offenders: The current landscape and some future directions for correctional psychology. *Criminal Justice and Behavior*, 31, 879–892.
- Wortley, R. (2002). *Situational prison control: Crime prevention in correctional institutions*. Cambridge, UK: Cambridge University Press.
- Yang, M., Wong, S.C.P., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, 136, 740–767.
- Yochelson, S., & Samenow, S. E. (1976). *The criminal personality: Vol. 1. A profile for change*. New York, NY: Aronson.
- Zamble, E. (1992). Behavior and adaptation in long-term prison inmates: Descriptive longitudinal results. *Criminal Justice and Behavior*, 19, 409–425.
- Zamble, E., & Porporino, F. J. (1988). *Coping, behavior, and adaptation in prison inmates*. New York, NY: Springer-Verlag.
- Zamble, E., & Porporino, F. (1990). Coping, imprisonment, and rehabilitation: Some data and their implications. *Criminal Justice and Behavior*, 17, 53–70.
- Zubek, J.P. (1969). *Sensory deprivation: Fifteen years of research*. New York, NY: Appleton-Century-Crofts.
- Zubrycki, R. M. (1980). *The establishment of Canada's penitentiary system: Federal correctional policy 1867–1900*. Toronto, Canada: University of Toronto.

Treating Criminal Offenders

ROBERT D. MORGAN, DARYL G. KRONER, JEREMY F. MILLS, AND ASHLEY B. BATASTINI

OF the 7.1 million offenders under correctional supervision in the United States, 2,266,800 are in local jails or in the custody of state or federal prisons (Glaze, 2011), which makes effective and efficient treatments of paramount importance. The treatment of incarcerated offenders presents particular difficulties and challenges and typically encompasses one of two therapeutic goals: mental health stabilization or rehabilitation. Alternatively, these goals may be conceptualized as basic mental health services or rehabilitative mental health services (Dvoskin & Morgan, 2010; Morgan, 2003). As described by Morgan (2003), basic mental health services are essential for facilitating offender adjustment within the criminal justice setting and include services geared toward mental health stabilization with emphasis on symptom reduction and the development of effective coping skills. These services are legally mandated (e.g., *Ruiz v. Estelle*, 1980; see also Arndt, Turvey, & Flaum, 2002; Cohen & Dvoskin, 1992) and must be made available to all offenders. Rehabilitative services, in contrast, aim to increase desistance. Desistance is not analogous to reduced recidivism, which may not include commission of new criminal offenses (e.g., parole revocation due to not reporting for parole office visits); rather, it refers to offenders actively avoiding triggers and antecedents (i.e., patterns of behavior) of criminal activity (see Meisenhelder, 1977). That is, rehabilitative services aim to alter criminal propensity, tendencies, and lifestyles for a reduction in criminal behavior with concomitant increases in prosocial behavior. This chapter reviews the literature and reports on the effective and efficient provision of basic and rehabilitative services.

EFFECTIVENESS OF CORRECTIONAL INTERVENTIONS

Basic mental health services are effective in correctional settings. Morgan and colleagues conducted a meta-analysis of 26 studies and found that mental health

treatments resulted in improved mental health functioning (e.g., reduced symptom distress), improved coping skills, and improved institutional adjustment and behavioral functioning in offender populations (Morgan, Flora, Kroner, Mills, Varghese, & Steffan, 2012). In a similar meta-analytic study, Martin, Shannon, Wamboldt, & Wooten (2012) also found that mental health services resulted in decreased symptom distress and greater overall mental health functioning. Although conducted during a similar time frame, these two meta-analyses included minimal overlap. Morgan, Flora, et al. (2012) focused on psychosocial interventions primarily in correctional settings, whereas Martin et al. (2012) included system-oriented services (e.g., mental health court) for incarcerated and nonincarcerated offenders. The results, however, were remarkably similar, providing clear evidence for the effectiveness of basic mental health services for improving mental health functioning (e.g., symptom distress, improved behavioral functioning).

The evidence for the effectiveness of rehabilitative services is even more compelling. For example, McGuire (2002) noted that more than 2,000 studies have examined the effectiveness of correctional rehabilitation efforts aimed at identifying what works, and meta-analyses (see Andrews & Bonta, 2010, for a thorough review of this literature) have consistently advanced our knowledge of effective interventions. In fact, the principles of effective correctional interventions have been identified such that the question has changed from "Does correctional intervention work?" to "What works for whom and under what circumstances?" (Wormith, Althouse, Simpson, Fagan, & Morgan, 2007). Although it is increasingly clear that some approaches to correctional rehabilitation are superior to others (see Gendreau, Goggin, & Smith, Chapter 23 this volume), it is now generally and almost universally accepted that the most effective rehabilitative programs adhere to the principles of the risk-needs-responsivity paradigm (RNR; Andrews & Bonta, 2010).

RNR, as outlined by Andrews, Bonta, and Hoge (1990), presents three principles of effective interventions: risk, need, and responsivity. Higher-risk offenders should be the focus of correctional interventions (risk), interventions should target the changeable or dynamic risks associated with criminal behavior (need), and interventions should be tailored to offender characteristics that may influence intervention effectiveness, such as cognitive abilities, learning styles, and diversity issues among others (responsivity). Empirical evidence overwhelmingly supports the principles of RNR and, as summarized by Gendreau et al. (Chapter 23 this volume), shows that adherence to RNR produces typical reductions in recidivism of 10% to 30%. Although these studies employed recidivism as an outcome measure rather than desistance, it is clear that adhering to principles of RNR improves outcomes. In fact, the evidence is so overwhelming that failing to adhere to the principles of RNR may be considered not only professionally negligent but fiscally irresponsible as well.

Andrews, Zinger, et al. (1990) conducted a meta-analysis to examine the effectiveness of services that adhered to the principles of RNR. A variety of services were included in this review, including criminal sanctions (e.g., incarceration, probation)

and rehabilitation (psychosocial services aimed at increasing desistance measured in this meta-analytic review as reducing recidivism). As previously noted, services that adhered to the principles of RNR were more effective than correctional sanctions alone or programs that did not adhere to principles of RNR, with an average reduction in criminal recidivism of 30%. But what is the cost of these effective services? Romani, Morgan, Gross, and McDonald (2012) conducted a follow-up of all of the studies included in the Andrews et al. (1990) meta-analysis. Specifically, this follow-up study reevaluated the studies with a "maximum cost" procedure to determine cost-effectiveness of correctional services that adhered to the principles of RNR compared to services that either did not adhere to the principles of RNR or implemented traditional criminal sanctions. Notably, there were no significant differences in cost of delivery across service types. In other words, it appears no more expensive to provide services that adhere to principles known to reduce recidivism than to provide alternative or no services. Furthermore, results suggested that offenders receiving services that did not adhere to the principles of RNR spent considerably more time in services (i.e., they spent more time in services that were less effective than did offenders in effective treatments), which contributed to a cost *increase*. Most important, when cost was compared to effectiveness, the cost of services adhering to the principles of RNR was significantly less expensive (i.e., more cost effective) than sanctions or services that did not adhere to the principles of RNR. The authors concluded that, when considering the principles of RNR, the bang is clearly worth the buck.

BARRIERS TO TREATMENT

Treatments will be effective only if the offenders actually receive services; however, incarcerated offenders are hesitant, at best, to utilize mental health services (see Mathias & Sindberg, 1985; Morgan, Rozycki, & Wilson, 2004; Rappaport, 1971; Steffan & Morgan, 2005) and resistant to the therapeutic process (Milgram & Rubin, 1992), such that they actively avoid services (Riordan & Martin, 1993). Unfortunately, approximately one-half of the most disturbed prison inmates may go without mental health services during their time of need (Steadman, Holohean, & Dvoskin, 1991). Although the rate of underserved is a concern in state and federal prisons, the situation is more dire in local jails, where less than 10% of inmates receive mental health services (Steadman & Veysey, 2007) in spite of estimates ranging between 14% and 24% of jail inmates having a history of severe mental health problems (James & Glaze, 2006).

SERVICE UTILIZATION

Although correctional service utilization is a relatively understudied area of service delivery, research suggests that offender perceptions of treatment and treatment providers directly impacts service utilization. For example, inmates with positive

attitudes toward mental health services are more likely to seek services while incarcerated (Deane, Skogstad, & Williams, 1999), and offenders who receive mental health services during periods of nonincarceration are more likely to access services while incarcerated (Garrity, Hiller, Staton, Webster, & Leukefeld, 2002), especially if prior experiences are perceived as positive (Deane et al., 1999).

Offender characteristics also impact service utilization. Ethnic minority offenders tend to hold more negative attitudes about interventions than nonminority offenders and are less likely to access services as a result (Deane et al., 1999; Skogstad, Deane, & Spicer, 2006; Steadman et al., 1991). Male inmates and younger offenders are less likely to access mental health services (Reinsmith-Meyer, 2008; Steadman et al., 1991). Inmates recently incarcerated are also less likely to access services, as they are naive to service availability and how to access services, apprehensive about the quality of mental health treatment in prisons, and concerned about confidentiality and stigma associated with receiving mental health services (Morgan et al., 2004).

Beyond individual offender characteristics, there appear to be four over-arching barriers to service utilization in prisons (Morgan, Steffan, Shaw, & Wilson, 2007):

1. Self-preservation concerns (concerns regarding confidentiality and perceptions of weakness or colluding with staff)
2. Procedural concerns (lack of knowing how, when, and why to access services and anticipated length of services)
3. Self-reliance (reliance on self or close others for help)
4. Professional service provider concerns (questions of staff qualifications and dissatisfaction with previous mental health services)

Additional barriers to service utilization include concern that services in prison are not effective and offenders' lack of motivation (Reinsmith-Meyer, 2008). Importantly, interventions aimed at overcoming treatment barriers and improving offenders' attitudes toward treatment programs can have positive effects (Nedd & Shihadeh, 1974) and may appreciably impact offender service utilization.

Incarcerated offenders in greater psychological distress (Garrity et al., 2002; Williams, Skogstad, & Deane, 2001) are more likely to seek mental health services when experiencing behavioral dyscontrol (impulsive or harmful behaviors), physical health concerns, negative affect (stress, anxiety, depressed mood), interpersonal difficulties with nonincarcerated others, and problematic institutional relationships with staff or inmates (Morgan et al., 2007). Unfortunately, incarcerated offenders appear much less willing to seek mental health services when they are experiencing suicidal ideation due to fear of losing privileges or preferred housing or other reasons associated with stigma (Howerton et al., 2007; Skogstad et al., 2006; Steadman, McCarty, & Morrissey, 1989). When utilizing mental health services for these problems, incarcerated offenders overwhelmingly prefer individual therapy with psychologists or counselors to group therapy or services provided by other mental health professionals (Morgan et al., 2004).

How should correctional mental health professionals combat these barriers to service utilization? Education may be an effective strategy for inmates experiencing procedural barriers to accessing services. For example, during new inmate orientation, mental health staff should provide verbal and written instructions for how to access mental health services. General orientation services could also orient new inmates to the problems they are likely to encounter in prison and how mental health services can be of assistance. Possibly most important, mental health professionals should address perceived barriers (e.g., stigma, effectiveness, confidentiality concerns) at the outset. In addition to education, correctional mental health professionals may reduce inmate barriers to services by providing outreach programs comparable to the services counseling centers provide on college campuses. Outreach programs on correctional units/pods aimed at educating inmates about mental health issues, institutional stress, effective coping, and the like may increase trust and provide the impetus for inmates in distress to request services.

TREATMENT DROPOUTS—EVERY SESSION COUNTS

Given that appropriate interventions among offenders are effective at reducing recidivism, those who drop out miss the opportunity to be impacted by intervention. Similar to persons receiving mental health services in the community, where approximately 50% of clients do not complete the therapeutic programs they begin (Kazdin, 1994; Wierzbicki & Pekarik, 1993), offenders present high dropout rates from correctional interventions (Empey & Gordon, 1989; Hamberger, Lohr, & Gottlieb, 2000; Hunter & Figueredo, 1999). Notably, offenders who drop out typically reoffend at a higher rate than those who complete treatment (Hepburn, 2005; Hiller, Knight, Saum, & Simpson, 2006; Seager, Jellicoe, & Dhaliwal, 2004; Wexler, Falkin, & Lipton, 1990) and offenders not receiving treatment (McMurran & Theodosi, 2007), and they reoffend more quickly (Huebner & Cobbina, 2007; Prendergast, Hall, Wexler, Melnick, & Cao, 2004). Thus, reducing dropouts is an essential treatment consideration for correctional interventions and public safety.

In addition to loss of treatment opportunity is the cost associated with treatment dropouts. The costs of treatment attrition are more noticeable within a structured treatment context. For example, delivering treatment below maximum capacity increases the cost to the treatment provider or other participating clients. Also, attrition leaves a vacant treatment spot that may be difficult to fill immediately, given that many interventions are programmatically designed to be completed in sequential order (i.e., joining a program in progress may not be as beneficial as starting it from the beginning). Within corrections, especially local jails, this is of particular importance, as the window of opportunity for treatment for offenders can be limited (i.e., approaching release dates). These costs are compounded by the substantial percentage of offenders who drop out of treatment (27%–28%; Hepburn, 2005; Hiller et al., 2006), which can consume a substantial proportion of treatment budgets.

Few investigators have examined why offenders terminate therapeutic programs prematurely; however, McMurrin and McCulloch (2007) identified four factors related to treatment completion. Although noncompleters of therapeutic programs were motivated for treatment and to change their criminal behaviors, they viewed treatments as less relevant and less timely to their concerns or current situations. In addition, noncompleters tended to be resistant to group participation, and a small number of noncompleters reported that challenging or demanding work contributed to their decision to withdraw. As noted by McMurrin and McCulloch, their study included a very small sample size, and further research is needed to elucidate factors that contribute to premature termination and ultimately improve offender treatment retention.

It is possible to identify offenders at greatest risk for premature termination from therapeutic programs. Specifically, offenders who are young (Hambridge, 1990; Pelissier, Camp, & Motivans, 2003; Zanis et al., 2003), antisocial (Moore, Bergman, & Knox, 1999), less educated (Babcock & Steiner, 1999; Geer, Becker, Gray, & Krauss, 2001; Wormith & Olver, 2002), have more extensive criminal histories (Hiller, Knight, & Simpson, 2006) including histories of violence (Moore et al., 1999; Pelissier et al., 2003), have less stable community supports (Butzin, Saum, & Scarpitti, 2002; Wormith & Olver, 2002), have increased criminogenic needs (e.g., criminal thinking, antisocial personality; Pelissier et al., 2003; Richards, Casey, & Lecente, 2003; Walters, 2004), and show denial (Geer et al., 2001) are at the greatest risk for premature termination from correctional interventions. Because offenders at greatest risk for dropout can be identified (analogous to identifying high-risk offenders in the risk principle of RNR), service providers can implement interventions aimed at reducing the risk for premature termination.

MOTIVATION AND THERAPEUTIC RESISTANCE

Treatment preparation and readiness have been understudied with offender populations (Williamson, Day, Howells, Bubner, & Jauncey, 2003). In spite of the dearth of research examining the effect of offender motivation, readiness for change, and therapeutic resistance on outcomes of interest (e.g., desistance), we do know that interventions aimed at increasing offender motivation and decreasing therapeutic resistance can be achieved (Morgan et al., 2007; Newbern, Dansereau, & Pitre, 1999; Shearer, Myers, & Ogan, 2001). Notably, increasing offender motivation for change increases continuity of care via treatment follow-up (aftercare), which contributes to desistance (Burdon, Messina, & Prendergast, 2004; McGrath, Cumming, Livingston, & Hoke, 2003). Thus, all correctional interventions should address issues of motivation and resistance in the early stages of treatment, and motivational interviewing appears to be a particularly promising approach. Specifically, integrating motivational interviewing into existing therapeutic programs may go a long way toward increasing offender participation in treatment (Chambers, Eccleston,

Day, Ward, & Howells, 2008), and subsequently reducing premature therapeutic terminations (e.g., treatment dropouts).

EFFECTING CHANGE: EVIDENCE-BASED CORRECTIONAL PRACTICE

Treating offenders presents many challenges and difficulties; however, rehabilitative programs and interventions have proven effective for helping inmates achieve positive outcomes, including decreased distress, improved mental health functioning, reduced recidivism, and, possibly of greatest importance, desistance. Based on the evidence to date, we have preliminary support for evidence-based practices for intervening with incarcerated offenders. These include targeting factors associated with criminal risk, grounding interventions in cognitive-behavioral theory (CBT), using simple treatment heuristics, incorporating homework into the therapeutic process, using structure to facilitate learning, and intensifying services. One additional treatment issue that warrants discussion is the therapist's cultural competence and is included in the discussion below.

TREATMENT TARGETS

The risk principle, from RNR, provides a road map for targeted interventions aimed at reducing recidivism and increasing desistance. A thorough review of the RNR literature (see Andrews & Bonta, 2010) reveals eight primary factors that account for the greatest percentage of criminal risk (Andrews et al. referred to these risk factors as the "central eight" because of the strength of their predictability when compared to other risk variables). The risk factors are listed next.

1. History of antisocial behavior
2. Antisocial personality pattern
3. Antisocial cognitions
4. Antisocial associates
5. Family and/or relationship circumstances
6. School and/or work functioning
7. Leisure and/or recreational pursuits
8. Substance abuse

To reduce criminal risk, treatment providers must provide services that target (i.e., aim to reduce) these risk factors.

Not surprising given the broad acceptance of the importance of incorporating the principles of RNR into correctional interventions, treatment providers consider issues of criminal risk to be important treatment considerations (Bewley & Morgan, 2011), and they tend to provide interventions that target areas of prominent criminal risk (see Morgan et al., 2012). However, when intervening with incarcerated

offenders with mental illness, treatment providers consider issues of mental illness recovery (such as psychosocial rehabilitation) as more important than treatments targeting issues of criminal risk and needs (Bewley & Morgan, 2011). Given the overwhelming evidence (see Andrews & Bonta, 2010, and Gendreau et al., Chapter 23 this volume, for a thorough review of this evidence) demonstrating the benefits of RNR, we submit that interventions aimed at reducing criminal activity must be grounded in the primary risk factors of antisocial cognitions, antisocial associates, family and/or relationship circumstances, school and/or work functioning, leisure and/or recreational pursuits, and substance abuse. Failure to do so is analogous to a physician treating persons with heart disease at risk for cardiac arrest without prescribing medications to reduce blood pressure as well as providing or recommending interventions aimed at improving stress management, diet, exercise, and other positive lifestyle changes.

COGNITIVE-BEHAVIORAL THEORY

As reviewed by Gendreau, Goggin, and Smith (Chapter 23, this volume), correctional interventions are more effective when they are grounded in CBT. In fact, CBT-driven interventions have proven effective with violent offenders (Berry, 2003), including sex offenders (Hanson et al., 2002; Polizzi, MacKenzie, & Hickman, 1999; Valliant & Antonowicz, 1992) and nonviolent offenders (Bonta, Wallace-Capretta, & Rooney, 2000).

The most prominently researched cognitive-behavioral program for offenders is the reasoning and rehabilitation (R&R) program, which was developed and implemented with federal prisoners in Canada in the 1980s (Ross, Fabiano, & Ewles, 1988). R&R utilizes a cognitive-behavioral approach and is delivered in a structured format over 36 two-hour sessions with groups of 6 to 12 offenders. The program utilizes an interactive approach, with multiple techniques of presentation. Role plays, games, cognitive exercises, and discussion are used to stimulate participation and improve reasoning skills (Robinson & Proporino 2004). The program aims to decrease criminality (i.e., program focus on criminalness) by addressing the cognitive distortions and thinking styles associated with criminal behavior. The curriculum covers the specific topics of self-control, interpersonal problem-solving skills, social perspective taking, critical reasoning, cognitive style, and values (Robinson & Proporino, 2004). A significant amount of time is spent addressing cognitive rigidity and problem-solving skills. Offenders are taught to think more systematically and analyze various solutions and outcomes before acting. Research on R&R with offenders has shown the program to be an effective means of reducing recidivism (Pearson, Lipton, Cleland, & Yee, 2002; Proporino & Robinson, 1995), and offenders demonstrated improvement in their criminal attitudes, criminal identification, and cognitive reasoning (Fabiano, Robinson, & Proporino, 1990). Notably, a meta-analytic review by Landenberger and Lipsey (2005) found that R&R was as effective as any other cognitive-behavioral intervention for the treatment of criminal offenders. Further meta-analytic reviews of R&R demonstrated reductions

in recidivism, on average, of 14% when compared to control groups and it was equally as effective with incarcerated offenders as with offenders in the community (Tong & Farrington, 2006, 2008).

Beyond demonstrating the effectiveness of R&R, Landenberger and Lipsey's (2005) meta-analysis examined therapeutic elements of cognitive-behavioral interventions for adult and juvenile offenders. Treatment elements included in this review were cognitive restructuring, interpersonal problem-solving skills, social skills, anger control, moral reasoning, victim impact, substance abuse, behavioral modification, and relapse prevention skills. Of these CBT elements, interpersonal problem solving and anger management had a positively statistically significant effect on recidivism outcome. Two areas, victim impact and behavioral modification, had a negative effect on recidivism outcome. This finding is particularly notable given the large number of correctional interventions that aim to increase victim impact, which intuitively seems to be a reasonable therapeutic strategy, but evidence suggests otherwise.

SIMPLE TREATMENT HEURISTICS

Given inmates' below-average educational attainment (Harlow, 2003) and intellectual functioning (Birmingham, Mason, & Grubin, 1996; Herrnstein & Murray, 1994), it is important to present therapeutic constructs and information in as simple a manner as possible (Morgan, Kroner, & Mills, 2006). Specifically, Morgan et al. recommended that learning heuristics be developed in a manner that is consistent with the offenders' everyday behaviors. For example, when educating offenders about the negative influences and risks of having criminal associates (see Mills, Jones, & Kroner, 2005), treatment participants are instructed to rate relationships utilizing a common everyday metaphor: a stoplight. A red light is universally recognized as a stop signal, and offenders can label their criminal and nonproductive associates accordingly. A yellow light is recognized as a warning of a pending red light (or a caution signal), and offenders can label their associates who engage in some antisocial and nonproductive behavior accordingly. A green light is universally recognized as positive (a go), and offenders can label their prosocial associates accordingly. Although a very simple concept, this learning heuristic allows inmates to evaluate their life situation from a perspective that is common to them; thus, they spend their time on evaluating their situation (working to reduce risk) and not on learning a complicated learning strategy or technique.

HOMEWORK

The best interventions and treatment programs are more effective with incarcerated offenders when they are able to incorporate what they have learned into their everyday environment. Out-of-treatment homework is one process that allows this to happen, and, not surprisingly, meta-analytic reviews with offenders have found that homework outside of the treatment setting significantly improves outcomes

(Morgan & Flora, 2002; Morgan et al., 2012). In fact, the evidence supporting the use of homework in offender treatment is sufficiently strong to suggest that it is essential for achieving maximum therapeutic benefit with offenders (McDonald & Morgan, 2012; Morgan et al., 2006). Two guidelines should inform therapists' use of homework: (1) homework exercises should be simple and structured to facilitate use of learned skills and behaviors, not challenge offenders' learning of skills and knowledge (i.e., homework is not a test of retention or learning); and (2) homework should be applicable but stimulating enough to sustain offenders' interest. It is commonly recognized that offenders are impulsive and easily bored (see Zamble & Quinsey, 1997); consequently, homework needs to be simplified but interesting. For example, watching a predetermined movie to identify instances of criminal thinking that led to antisocial behavior will be much more stimulating to offenders than a bibliotherapy assignment designed to educate offenders about the negative impact of their thinking on their behavior. Obviously, homework will be effective only if it is completed. McDonald and Morgan (2012) identified two promising strategies for enhancing homework compliance. Although future research needs to examine these preliminary findings in greater detail, strategies of public commitment (i.e., having offenders publicly commit, in the treatment group, for example, to completing homework assignments; Freeman & Rosenfield, 2002) and task modeling (i.e., providing in-session modeling and/or rehearsal; Kazantzis & Lampropoulos, 2002) are promising strategies for increasing homework compliance.

STRUCTURE

Given that CBT is a structured therapy that has proven effective with offenders, it should be of no surprise that structured interventions produce more favorable outcomes for offenders than nonstructured interventions (Leak, 1980; Morgan & Flora, 2002). Structure most commonly takes one of two forms. Structure can be incorporated into the therapeutic process by the addition of specific (structured) learning activities. These are typically skill-based activities that help offenders develop specific skills or abilities. Alternatively, structure can be incorporated into the therapeutic process so that the treatment specifies who will do what and how. For example, psychoeducational processes can be regularly integrated into an intervention to facilitate learning and acquiring new information. Regardless of how structure is integrated into the therapeutic process, it is clear that a structured approach produces superior results when compared to nonstructured processes such as psychodynamic approaches (Andrews et al., 1990).

INTENSIVENESS

Services for offenders are most effective when they are intensive. Specifically, appropriately intensive services occupy a significant portion of the offender's time (between 40% and 70%) and are of significant duration (between 3 and 12 months; Gendreau, 1996b). The more intensive therapeutic services are (i.e., of longer

treatment duration), the better the outcomes when working with offenders (Lipsey, 1989, as cited in Andrews & Bonta, 2010). Along similar lines, the greater the treatment dosage, the better the outcomes (see Aytes, Olsen, Zakrajsek, Murray, & Ireson, 2001; Bourgon & Armstrong, 2005; Fisher, Beech, & Browne, 2000; Gossop, Marsden, Stewart, & Rolfe, 1999; Lipton, 1995; Westhuis, Gwaltney, & Hayaski, 2001; Wexler, Falkin, & Lipton, 1990). In fact, it has been found that, for every month an offender spends in treatment, a 4% decrease in recidivism can be expected (Burdon et al., 2004).

THERAPIST CULTURAL COMPETENCE

Although not yet an evidence-based practice in corrections, greater effort must be devoted to employing culturally competent therapists and developing culturally sensitive interventions. Programs such as that developed by Polaschek and Dixon (2001), which specifically integrates culturally based practices into a treatment program for violent offenders, are the exception rather than the rule. Given the racial disparity in prisons across the world, it is no longer acceptable for treatment providers to overlook issues of diversity and employ treatment materials that lack ethnic minority case material (e.g., images, vignettes). Cultural competence in correctional settings needs to extend beyond individual offender characteristics and include competence for working within the prison culture. Many correctional institutions maintain a machismo facade whereby everyone, including professional staff, is tough and tough on inmates. This is counterproductive for behavioral change and inconsistent with the principle of responsivity in RNR. On the contrary, the evidence is compelling (see Andrews & Bonta, 2010; Gendreau, 1996a, 1996b) that interpersonally sensitive therapists working from a service-oriented perspective produce better outcomes with offenders.

TREATING SPECIAL POPULATIONS

Our discussion to this point has focused on issues generic to all incarcerated offenders and effective therapeutic strategies for intervening with nonspecific offender groups. Notably, however, some groups of offenders present with special needs and require specific interventions as a result. Although we do not have the space to review all such groups, we review effective interventions and strategies for four groups of offenders: (1) offenders with mental illness, (2) offenders with learning disorders, (3) violent offenders, and (4) juvenile offenders. Comparable information is available for other specialty groups, including sex offenders, offenders with substance abuse problems, and psychopathic offenders.

OFFENDERS WITH MENTAL ILLNESS

When treating imprisoned offenders with mental illness, service providers tend to place more emphasis on basic mental health services (e.g., symptom management

and stabilization) than on rehabilitative efforts, such as risk-need or preparing inmates for release (Bewley & Morgan, 2011). This finding is in spite of the compelling evidence of the benefits of RNR, as previously discussed (and discussed by Gendreau et al., Chapter 23 this volume). It is likely that service providers opt for providing basic mental health services, at least in part, due to offender need. Local jails and prisons are responsible for helping inmates maintain stable mental health, and basic mental health services are a primary mechanism for that to occur; however, it also appears that service providers are likely driven to favor basic mental health services over rehabilitative services because of assumptions common in the mental health and criminal justice fields that offenders' criminal behavior is driven by an absence of adequate mental health services and by destabilization (Lamb & Bachrach, 2001; Lamb & Weinberger, 1998; Teplin, 1984). However, offenders with mental illness have some of the same risk factors, including primary criminal risk factors, as their offender peers who are not mentally ill. Specifically, research examining two independent samples of incarcerated offenders ($n = 414$ and $4,204$) showed that incarcerated offenders with mental illness produced levels of criminal thinking and antisocial attitudes consistent with nonmentally ill inmates (Morgan, Fisher, Duan, Mandracchia, & Murray, 2010; Wolff, Morgan, Shi, Huening, & Fisher, 2011). Offenders with mental illness are likely to face other primary criminal risk factors, including occupational limitation, impaired family relations, and substance abuse. In fact, the evidence is so compelling that it is now recognized that the offender incarcerated due to complications with mental illness is the exception (Skeem, Manchak, & Peterson, 2011); thus, when working with offenders with mental illness, service providers must target dual issues of mental illness and criminal propensity.

As noted previously, treatment providers in correctional facilities provide effective basic mental health services. Specifically, a comprehensive meta-analytic review of interventions for incarcerated offenders found significant improvements for general mental health outcomes, improved coping skills, and improved institutional adjustment with fewer behavioral problems (Morgan et al., 2012)—all goals of basic mental health services in jails and prisons. Similar outcomes were obtained in a separate meta-analysis, with reductions in symptom distress and improved functioning (Martin, Dorken, Wamboldt, & Wooten, 2012). Notably, however, interventions also appear to diminish criminal recidivism as well as psychiatric recidivism (return to hospitalization), which are goals of rehabilitative services. Morgan, Flora, et al. (2012) found evidence for reduced criminal and psychiatric recidivism, both notable accomplishments. The most significant treatment gains with respect to effect size were produced in the sole study that targeted both mental health needs and criminal behavior. Similarly, Martin et al. (2012a) found clear reductions in continued criminal justice involvement of any type (e.g., revoked, new arrest, etc.). These meta-analyses also provide important insights into effective therapeutic strategies. Morgan, Flora, et al. (2012) found that the use of homework (with emphasis on active homework exercises that required offender activity, such

as practicing learned skills or social interaction), behavioral practice of new behaviors, and an open treatment admission policy all contributed to more favorable outcomes. Martin et al. (2012) found that continuity of services between institutions and community, allowing for some level of voluntariness in the intervention, and nonspecified treatment duration as opposed to time-limited services all produced more favorable outcomes.

The findings just described—that offenders with mental illness present with criminal risk factors similar to non-mentally ill inmates and that treatments that integrate both mental health and correctional rehabilitative efforts produce favorable outcomes compared to interventions that do either alone—support the opinion of Hodgins et al. (2007), who noted that offenders with mental illness present unique challenges that require service providers to treat both psychiatric symptoms and criminal propensity and risk. That is, interventions for offenders with mental illness should aim to decrease psychiatric hospitalization days as well as time spent incarcerated while simultaneously working to improve quality of life (e.g., increased number of functional days). Desistance and mental health recovery (*recovery* here does not refer to remission but to achieving a return to independence as a result of self-management of illness producing improved quality of life; see Corrigan, Mueser, Bond, Drake, & Solomon, 2008) are the ultimate goals.

We developed Changing Lives and Changing Outcomes: A Treatment Program for Justice Involved Persons with Mental Illness (Morgan, Kroner, Mills, & Bauer, 2012), a comprehensive and holistic intervention designed for the specific needs of offenders with mental illness. Changing Lives and Changing Outcomes utilizes a bi-adaptive model of intervention by targeting dual (bi) issues of mental illness and criminal propensity to improve functional (adaptive) outcomes for offenders with mental illness. The aim of this treatment model is not to cure mental illness, but rather to maximize adaptive behaviors to optimize functioning. Changing Lives and Changing Outcomes includes a three-part treatment protocol. Part I of the protocol is targeted to issues of mental illness but integrates issues of criminogenic risk and includes three treatment modules: Mental Illness and Criminalness Awareness, Medication Adherence, and Coping with Mental Illness and Criminalness. Part II of the treatment protocol is targeted to issues of criminalness but integrates issues of mental illness and includes Problematic Thoughts and Attitudes, Emotions Management, and Problematic Associates. Part III of the treatment protocol includes three modules that are relevant to both mentally ill and offender populations. They are included as overlapping treatment targets and include Preparing for Change, Skill Development (i.e., problem-solving skills, social and recreational skills, and vocational/housing skill development), and Substance Abuse. The program consists of 77 sessions that last between 1.5 and 2 hours. We recommend a minimum of 3 sessions per week for an approximate treatment delivery time of 6 months.

Preliminary field testing of this program with five groups of incarcerated offenders with mental illness ($n = 50$) is promising (Morgan, Kroner, Mills, Bauer, & Serna, 2012). Of particular significance, given the length and intensity of the program and

typical correctional dropout rates previously discussed, preliminary fidelity results proved very encouraging, as 31 (66%) completed the program. In fact, this completion rate was almost identical to the correctional treatment completion rate of 64.2% found in a meta-analytic review of studies with 41,000 participants, when many of these studies included much shorter and less intense programs. Participants attended 94.08% of sessions and completed 83.48% of assigned homework (homework assigned for 89% of sessions). Participants were, on average, engaged and active in the treatment process and reported significant treatment satisfaction and positive therapeutic bond with treatment providers on standardized measures. Finally, significant improvements over time (pre–post) with small to moderate effect sizes were noted on measures of symptom distress, psychopathology, and some aspects of criminal thinking.

OFFENDERS WITH INTELLECTUAL DISABILITIES

Offenders with intellectual disabilities (IDs) can have a high occurrence of antisocial/aggressive disorders (Lund, 1990). Although these offenders have cognitive deficits, the literature has repeatedly shown interventions to be effective (Taylor, 2010). Specifically, research has shown that cognitive-behavioral techniques are particularly effective when addressing treatment targets with this population of offenders (Barron, Hassiotis, & Banes, 2002). As with other offenders, structured interventions delivered in a consistent and reliable fashion with appropriate staff-to-offender ratios are important for treatment success.

Effective understanding of offenders with IDs is essential to providing effective correctional interventions. A significant barrier to engaging these clients is the overlap between mental illness and ID. Issues of mental illness among offenders with IDs often go undetected (undiagnosed, untreated) for one of the following reasons (Taylor, Lindsay, & Willner, 2008):

1. Two distinct groups of professionals typically provide services for those with IDs and those with mental illness.
2. There is an absence of good tools for assessing mental health concerns among those with IDs.
3. Poor differential diagnosis may occur when symptoms of IDs are attributed to mental health concerns and mental health issues are attributed to intellectual deficits.

VIOLENT OFFENDERS

Current approaches to treating violent offenders are based on social learning and social information-processing theories (Cortoni, Nunes, & Latendresse, 2006; Polaschek & Dixon, 2001; Serin & Preston, 2001; Wong, Gordon, & Gu, 2007). The basic premise is that violent behaviors have been learned through direct experience;

modeling by family, peers, and cultural figures; and reinforcement and/or cognitive mediation. Interventions endeavor to reduce the risk of violent recidivism in high-risk male offenders through learning nonviolent alternatives. This involves providing the skills required to identify negative lifestyles and heighten awareness of violence, responsibility, and control. Such programs are designed to motivate participants to challenge their use of violence, change their antisocial and pro-violence attitudes and beliefs, and develop a prosocial lifestyle. Central components for effective interventions among violent offenders include: anger management, addressing antisocial attitudes and cognitive distortions, developing relationship enhancement and social problem-solving skills, developing conflict resolution skills, and developing self-management skills. Given that learning takes place with small increments, most programs treating violent offenders are of substantial length, usually lasting a minimum of 6 months. The delivery of programs for violent offenders will have a high level of structure, an emphasis modeling and rehearsing new skills in sessions, and having offenders practice the skill on their living units. New skills will include restructuring negative thoughts and behaviors associated with patterns of violence.

Programs designed to reduce violent behavior are effective. For example, a large (458 beds) prison-based therapeutic program designed for violent offenders and grounded in principles of RNR resulted in reduced disciplinary infractions and staff and inmate assaults (Wang, Owens, Long, Diamond, & Smith, 2000). A psychosocial intervention with particular promise is the Violence Reduction Program: A Treatment Program for Violence Prone Forensic Clients (VRP; Wong & Gordon, 2012). The VRP is also grounded in principles of RNR, is consistent with evidence-based practices described earlier (e.g., utilizes CBT), and includes three therapeutic phases: (1) learning about aggressive behaviors and readiness for change; (2) skill development to manage thoughts, feelings, and behaviors associated with violence; and (3) over-learning skills and relapse prevention. Because the intervention is not time limited, offenders can work through the program at their own pace and consistent with their responsivity needs. Evaluations of the VRP to date have produced positive outcomes that include successful transfer to lower-security facilities, fewer institutional behavior problems, and reduced community violence (Di Placido, Simon, Witte, Gu, & Wong, 2006; Wong et al., 2005, 2007). Although further research examining the effectiveness of the VRP program is warranted, these preliminary findings are very encouraging.

JUVENILE OFFENDERS

The general approach to youthful offenders within the justice system has been largely rehabilitative rather than punitive. As such, the primary goal of juvenile court is to address specific targets for treatment that are likely to reduce the risk of continued criminal conduct. This differs from the perspective of criminal court, which aims to demonstrate culpability and impose a punishment equal to the crime

committed, with less emphasis on rehabilitation (Batastini, Hunt, Present-Koller, & DeMatteo, 2011). Juvenile offenders are believed to be more malleable to behavior change than adults, given their vulnerable developmental stage. In other words, there is a greater hope that these offenders—whose early criminal career is more easily disrupted—will become productive members of society.

For juvenile offenders, the use of a comprehensive mental health assessment appears to be particularly important to the selection of effective treatment strategies—a step that is often overlooked by professionals (Mulvey & Iselin, 2008). For example, identifying individual responsivity factors, including learning style, cognitive ability, developmental level, and psychological functioning, is necessary to match youthful offenders with interventions that are delivered in a clear and understandable manner (DeMatteo, Hunt, Batastini, & LaDuke, 2010). Clinicians also tend to base programming for juveniles on factors that are less potent predictors of reoffending (Borum, 2003); however, we know that noncriminogenic factors, such as low self-esteem and motivation for success, have limited empirical support compared to criminogenic factors, such as antisocial attitudes and negative peer associates (Andrews & Bonta, 2010). Furthermore, treatments for youthful offenders are often selected on the basis of a single risk component despite the fact that juvenile offenders are a particularly complex population with diverse and multifaceted needs (DeMatteo et al., 2010).

One of the most promising treatment approaches that accounts for the many factors contributing to youth delinquency is multisystemic therapy (MST), a model of therapy based on ecological and systems theory that incorporates various interpersonal networks in which the youth is involved (e.g., school, community, family, social/peer). MST first assesses problems within each of these networks that contribute to the youth's behavioral and psychological functioning. Individualized intervention protocols are then developed to address known problem areas through a variety of mechanisms (e.g., family therapy, problem solving, behavioral modification, and psychopharmacology; Sheidow & Henggeler, 2005). Research has consistently demonstrated acceptability and efficacy of MST programs. Results of some studies suggest as high as 98% completion rate for youthful offenders and their families (e.g., Henggeler, Pickrel, Brondino, & Crouch, 1996). Furthermore, reduction rates of recidivism range from 26% to 69% across studies comparing MST to a comparison group (Sheidow & Henggeler, 2005). MST also increases school attendance and abstinence from substance use (Brown, Henggeler, Shoenwald, Brondino, & Pickrel, 1999; Henggeler, Clingempeel, Brondino, & Pickrel, 2002). Other similar community-based treatment programs for juvenile offenders (e.g., functional family therapy, parent management training, wraparound services, treatment foster care) have also demonstrated some effectiveness in reducing important risk factors that perpetuate criminal behavior (Sheidown & Henggeler, 2005; Tate & Redding, 2005).

Because MST is delivered primarily in outpatient settings, given its focus on the youth's family and social contexts, it may not be a viable option for juvenile offenders detained in residential care or detention centers.

As has occurred with many adult offenders, there has been an increased reliance on the justice system to provide services for juveniles with more pervasive mental health problems (Redding, Lexcen, & Ryan, 2005). Results of one study suggested that as many as 66% to 75% of youth who come in contact with the juvenile justice system meet diagnostic criteria for a psychiatric disorder (Teplin et al., 2006). When conduct disorder was excluded, this number decreased only slightly (Teplin et al. 2006). Fortunately, given the nature of the juvenile justice system, it is generally better equipped than adult correctional systems to address mental health needs. However, there are several disadvantages for youth who are isolated from their communities during treatment. These include possible academic, developmental, or social delays; limited family involvement that may weaken parent–child bonds; and negative peer attachments that are often formed when troubled youth live together (Redding et al., 2005).

Despite these barriers, some interventions can be effective for youth living in secure settings. To reduce acting out and promote prosocial behaviors, behavioral contingency plans that rely on consistent and immediate consequences are often employed in these settings. However, Redding et al. (2005) warned that appropriate staff training is necessary to prevent iatrogenic effects of too much punishment on mentally ill youth, as some disciplinary violations may be a reflection of disturbed symptomology. Cognitive-behavioral approaches have also shown promise, particularly for specific problem areas. For example, a treatment outcome study on juvenile sex offenders that combined elements of CBT and relapse prevention found that juveniles who completed the treatment program had significantly lower sexual and general recidivism rates than youth who failed to complete the program or who received treatment as usual (Worling & Curwen, 2000).

New programs have also emerged in an effort to address some of the gaps in treatment that residential and correctional-based programs can create. One such initiative is the establishment of intensive aftercare programs (IAPs). These programs are designed to ensure continuity of care and maintenance of positive treatment gains following institutionalization of juvenile delinquents. Like MST, IAPs emphasize prosocial connections for the youth in the community by involving families, schools, and correctional agencies. IAPs integrate therapy services with intensive supervision and case management (Tate & Redding, 2005).

TELEHEALTH: AN INCREASINGLY COMMON MODALITY FOR SERVICE DELIVERY IN CORRECTIONS

With recent technological advances that have expanded the ways in which communication can occur, it is not surprising that similar advances have also been observed within the healthcare system, including the practice of psychology and psychiatry. Telehealth—also referred to as telemedicine—is a method of service delivery that allows clients and providers to interact in real time over a distance (Ax, Fagan, & Holton, 2002). Typically this involves the use of audiovisual equipment, such as digital videoconferencing programs. Given the increase in psychological applications

of telehealth, the American Psychological Association (Palomares, 2012) has established a task force to create best practice guidelines. For the purpose of uniformity, this task force has designated the term *telepsychology* to refer to field-specific (e.g., counseling services, psychodiagnostic assessment) uses of this modality.

The use of telehealth in forensic and correctional settings appears to be gaining attention among policy makers, professionals, and the individuals served by the criminal justice system. It has been estimated that about one in every five telehealth applications involves correctional healthcare (Lowes, 2001), and interventions that target the needs of mentally ill inmates are one of the most frequently cited uses (Ax et al., 2007). Telehealth has been used in these settings for a variety of purposes, including legal consultation and court testimony, forensic mental health assessment (e.g., competency evaluations, sexually violent predator evaluations), juvenile rehabilitation, psychiatric medication management, group treatment for inmates in segregation, and training and continuing education of professionals (Larsen, Stamm, Davis, & Magaletta, 2004).

Several trends in the correctional system suggest that telehealth, and telepsychology in particular, may be promising approaches for the treatment of criminal offenders. For one, rising incarceration rates and economic hardships have prompted prison expansion throughout the United States (Hooks, Mosher, Rotolo, & Lobao, 2004). Correctional facilities not only provide a place to house offenders, but they are often seen as a quick-fix solution to unemployment and local debt (Glasmeier & Farrigan, 2007; King, Mauer, & Huling, 2004). Unfortunately, these facilities are often located in rural areas where access to quality treatment resources may be limited (Ax et al., 2002). The increasing costs associated with correctional healthcare (e.g., travel expenses, driver salaries, services rendered) and the scarcity of professionals willing to treat this population further reduces the number of incarcerated individuals who receive adequate treatment (Daniel, 2007; Magaletta, Fagan, & Ax, 1998). For those who receive services, their needs are often addressed in an untimely or inconsistent manner (Magaletta et al., 1998).

Telepsychology offers a number of benefits that may alleviate many of the issues faced by outside agencies and the institutions they serve. The most obvious benefit is the safety and security of providers, correctional staff, the public, and offenders. The ability to deliver services within a secure environment removes the risk of escape or harm to staff and civilians (Zaylor, Nelson, & Cook, 2001). Likewise, the convenience of treating offenders in the comfort of their own offices may entice more providers to include these individuals on their caseloads. Furthermore, telepsychology can remove some of the hurdles to initiating services such as transportation, security procedures, and travel time (National Institute of Justice, 1999), which may reduce the length of time offenders must wait for treatment. Benefits of telehealth also apply to psychiatric emergency care situations that can occur when onsite providers are not readily available (Magaletta, Fagan, & Peyrot, 2000). Proactively responding to offenders' needs likely will reduce agitation and increase trust in the justice system.

These effects may, in turn, help manage stress and anxiety among correctional staff, who will be able to focus more on maintaining a secure facility instead of attending to grievances or misconduct (Magaletta et al., 1998).

Cost effectiveness is another widely cited benefit of telepsychology. This is currently the most robust area of research to date (Ax et al., 2007). An in-person health consultation for an inmate is estimated to cost about \$173, compared to \$71 for a telehealth consultation (National Institute of Justice, 2002). States with substantial telehealth networks have reported saving between \$200 and \$1,000 for every consult (Kinsella, 2004). Although some correctional agencies—given their already limited budgets—may be reluctant to invest in telehealth or telepsychology equipment, start-up costs (including software, hardware, and signal transmission) have decreased considerably in recent years (Miller, Clark, Velkamp, Burton, & Swope, 2008). More access to lower-cost services can translate to more individuals who have both mental health and criminal justice problems receiving treatment.

However, before telepsychology is accepted as a new solution to the many problems in correctional mental health, it is important first to evaluate whether it can produce the desired treatment effects. One concern with telepsychology is the potential loss of connectedness between patient and provider. However, research consistently suggests that this is not the case. For example, Morgan, Patrick, and Magaletta (2008) evaluated perceptions of the treatment experience following psychiatric and psychological consultations conducted via telepsychology and in person. Results showed no significant differences across service modality for inmates' current mood, perceptions of the working alliance, and overall satisfaction with the provider and treatment modality. Similarly, a survey of inmates who received psychiatric services via videoconferencing revealed that 81% of respondents rated treatment positively, while 35% of respondents actually preferred videoconferencing to in-person sessions (Magaletta et al., 2000). The acceptability of this modality among offenders is necessary for ensuring other important aspects of the treatment process, such as the therapeutic alliance and compliance with recommendations.

Less is known about the ability of remotely delivered services to assess and subsequently reduce symptomology. The research that is available implies relative comparability with in-person services. One example is a study by Lexcen, Hawk, Herrick, and Blank (2006) that demonstrated modest to excellent interrater reliability ($r = .69$ to $.82$) between telepsychology and in-person forensic mental health assessments using the Brief Psychiatric Rating Scale-Anchored Version and the MacArthur Competence Assessment Tool—Criminal Adjudication. Likewise, in a study of jailed inmates receiving psychiatric services through videoconferencing, Nelson, Zaylor, and Cook (2004) compared patient symptom ratings completed prior to the consultation with provider evaluations of mental health functioning completed during the consultation. Significant positive correlations were found between patient and provider symptom impressions, particularly with regard to

suicidal ideation. However, this study did not include an in-person comparison group. Despite this evidence, the reliability and validity of clinical diagnoses and treatment recommendations derived from virtual assessments require further exploration, particularly when persons with severe cognitive or behavioral disturbances are being evaluated.

Data on treatment outcome are more limited, but preliminary research points to positive effects. One study examined whether telepsychology services were longitudinally effective in improving psychological functioning from both the provider and patient perspective (Zaylor et al., 2001). Adult inmates detained in a rural jail received either suicide consultation or long-term care for a mental illness through a telehealth clinic. Results indicated a significant time effect, such that patients reported experiencing less distress over time and psychiatrists rated patients as less "ill" over time. Fox, Connor, McCullers, & Waters (2008) also found positive treatment gains in a sample of juvenile offenders. Results showed significant improvements in goal attainment (e.g., education, social skills) from pre- to postintervention.

Treatment outcome research that includes baseline and follow-up data on relevant clinical markers (e.g., disciplinary infractions, crisis interventions), comparison groups (e.g., treatment as usual), larger sample sizes, and more rigorous statistical methods is necessary to draw firmer conclusions about the efficacy of telepsychology in treating offenders with mental health problems (Antonacci, Bloch, Saeed, Yildirim, & Talley, 2008; Ax et al., 2007; Monnier, Knapp, & Frueh, 2003; Morgan et al., 2008). Research should also investigate more intensive clinical applications of telepsychology including psychoeducation, cognitive-behavioral therapy, psychosocial interventions, and family counseling. Additionally, a better understanding of the strengths and limitations of telepsychology for serving various forensic subpopulations, including offenders with thought and personality disorders, neurologically or cognitively impaired offenders, suicidal offenders, offenders in administrative segregation, female offenders, and juvenile offenders (Antonacci et al. 2008; Magaletta et al., 2000), is needed.

Telepsychology appears to be a practical alternative to in-person service delivery. The evidence at present indicates that this modality may begin to close the gap between the high rates of offenders with mental health needs and the low availability of appropriate treatment resources in some settings. In fact, telepsychology has a variety of benefits over in-person consultations (e.g., increased safety, inclusion of family or multiple community providers in the treatment process, more timely service delivery, decreased costs). Although services provided over a distance likely have unique limitations (e.g., some lost behavioral data), something is better than nothing. With continued research efforts, the field of correctional mental health will be better informed about when and for whom telepsychology is most effective (Batastini, McDonald, & Morgan, 2012). These technological advances could slow the revolving door of the criminal justice system for disturbed offenders.

INTEGRATING ASSESSMENT INTO THE THERAPEUTIC PROCESS

Intervention with offenders has a different focus from traditional therapeutic intervention. For example, within corrections, the global focus is on improved public safety as opposed to the alleviation of human suffering. Similarly, the role of assessment in corrections is different from its role in noncorrectional settings. Specifically, assessment in corrections has two primary foci: (1) assessment of risk (see earlier discussion of RNR and Gendreau, Goggin, and Smith, Chapter 23 this volume) and (2) assessment of therapeutic change.

ASSESSING RISK

With offenders, assessment focuses on the estimation of risk and identification of criminal risk factors as compared to the diagnosis and estimation of symptom severity when intervening with nonoffender populations. As previously noted, Andrews and Bonta (1994) identified dynamic risk factors (also referred to as criminogenic needs; Andrews & Bonta, 2010) as targets for intervention that could potentially reduce reoffending. A meta-analysis demonstrated that dynamic risk factors were among the better predictors of recidivism (Gendreau, Little, & Goggin, 1996), including antisocial companions, antisocial attitudes, and antisocial personality. An advantage of focusing on dynamic risk factors is the potentially explanatory role over more static risk factors, such as age, gender, race, and criminal history. As noted in the risk principle, it is increasingly important that risk be routinely assessed prior to initiating treatments or interventions.

The routine assessment of risk and needs (combining static and dynamic risk factors) has contributed to the prediction of recidivism and has been operationalized through instruments such as the Level of Service Inventory–Revised (LSI-R; Andrews & Bonta, 1995) and the Level of Service/Case Management Inventory (LS/CMI; Andrews, Bonta, & Wormith, 2004). These instruments utilize both static and dynamic domains that include criminal history, education, employment, finances, interpersonal relationships, attitudes, companions, and substance abuse. These instruments identify not only an underlying level of criminal risk but also intervention targets, thus integrating risk assessment with intervention recommendations and risk management strategies. Essential to effective correctional interventions is the assessment of risk and risk factors and linking them to the risk assessment and risk management process.

Elsewhere we have articulated the integration of static and dynamic risk factors into what we refer to as an *integrated-actuarial* approach to risk assessment (Mills, Kroner, & Morgan, 2011). In brief, this approach to risk assessment does four things:

1. It identifies an actuarial estimate of criminal risk.
2. It identifies potentially dynamic risk factors (those risk factors that can be changed and lead to reductions in risk).

	RNR		
Integrated-Actuarial Risk Assessment	Risk	Needs	Responsivity
Identify Actuarial Risk			
Identify Dynamic Risk Factors			
Recommend Intervention			
Identify Risk Management Strategies			

Figure 24.1 Overlap of Effective Correctional Intervention and Risk Assessment

3. It provides recommendations for intervention/treatment.
4. It provides risk management strategies.

This approach complements the endeavor of offender treatment and the overall assessment of change that stem from the intervention (see Figure 24.1; shaded areas show overlap). For example, within the RNR framework, the assessment of actuarial risk identifies the level of intervention required. High-risk offenders require more intervention to effect change than lower risk offenders. The needs principle directs the identification of dynamic or changeable risk factors and recommends intervention to change (reduce) the potential of those factors. Finally, the responsivity principle looks to idiographic issues that might influence the acquisition of intervention benefit and are directly related to the recommendations concerning intervention and subsequent risk management strategies. We articulate these points to demonstrate that appropriate risk assessment and optimal interventions are not divergent but rather complementary processes.

ASSESSING OFFENDER CHANGE

As mentioned previously, effective correctional intervention calls for the accurate identification of dynamic or changeable risk factors: This could refer to substance abuse, antisocial attitudes, antisocial associates, or cognitions supporting violence,

among others. Once the criminogenic content areas are identified, appropriate interventions can be selected. This procedure is consistent with the hatched area of the overlap between risk assessment and effective correctional intervention in Figure 24.1. With respect to any criminogenic content area, the challenge for both the risk assessor and the program facilitator is the identification of meaningful offender change. To optimize the likelihood of success, we recommend a multimethod, multiconstruct, multi-time, and multi-outcome target approach.

Multimethod Approach. Within correctional populations, the term *multimethod* typically refers to interview-based measures, offender self-report, and/or behavioral markers. In some institutions, third-party data may also be available in the form of correctional officer reports, unit counselor ratings, and the like. One of the more frequently used interview-based measures that covers a broad range of criminogenic content areas is the LSI-R (Andrews & Bonta, 1995). The 54 items, which are scored following a record review and comprehensive psychosocial interview, cover a broad range of criminogenic content areas, such as education/employment, financial, family marital, accommodation, leisure/recreation, companions, alcohol/drug problem, emotional/personal, and attitudes/orientation. Interrater reliability coefficients as reported by the authors range from .80 to .96 (Andrews & Bonta, 1995), and subsequent research has supported these initial reliability findings (Kroner & Mills, 2001). The LSI-R is an effective predictor of recidivism; however, its use as a measure of treatment success has recently gained momentum (Gendreau & Smith, 2007; Girard & Wormith, 2004; Hollin & Palmer, 2006; Lowenkamp, Holsinger, & Latessa, 2001; Manchak, Skeem, & Douglas, 2007; Rooney & Hanson, 2001).

Another clinician-scored instrument that is accruing support for assessing criminogenic content areas and specifically within the context of treatment change is the Violence Risk Scale (VRS; Wong & Gordon, 2006). Among the criminogenic content areas assessed by the scale are criminal attitudes, criminal peers, emotional regulation/control, substance abuse, stability of relationships, impulsivity, and cognitive distortions. All of the variables assessed by the VRS are rated on a 4-point Likert type scale from 0 to 3. Like the LSI-R, VRS total scores are predictive of postrelease criminal behavior (Beggs & Grace, 2010). More will be said about the VRS measuring treatment change a little later. The benefit of both the LSI-R and VRS is that, in addition to assessing for the presence of specific criminogenic content areas, the scales also produce an index of overall risk for criminal behavior.

While there are a number of self-report instruments that measure various specific criminogenic content areas, few have been specifically designed to capture sufficient content areas to predict recidivism. One exception to this is the Self-Appraisal Questionnaire (SAQ; Loza, 2005). The SAQ is composed of 72 true or false items that form seven clinical subscales and a validity subscale. The clinical subscales include criminal tendencies (antisocial attitudes, beliefs, behaviors, and feelings), antisocial

personality problems (characteristics similar to those used to diagnose antisocial personality disorder), conduct problems (assesses childhood behavioral problems), criminal history, alcohol/drug abuse, antisocial associates (the offender's perception of the effect of his associates on his criminal activities), and anger (measures reaction to anger). The validity subscale was designed to detect careless responses or related problems associated with responding to self-report measures. The reliability, and construct and concurrent validity of the SAQ have been demonstrated (Loza, 2005), as has the predictive validity of the SAQ over a 2-year (Kroner & Loza, 2001; Loza & Loza-Fanous, 2001) and 5-year period (Loza & Loza-Fanous, 2003). The SAQ is at least as effective as four other well-established professionally rated and widely used measures for the prediction of recidivism (Loza & Loza-Fanous, 2001). Further, the SAQ has been cross-validated with Australian, British, and Singaporean samples (Loza et al., 2004). One limitation of the SAQ is that it has not been utilized as a treatment outcome measure, so, although it can be useful in predicting risk, it has not been established as a measure to show posttreatment reductions in risk. Research of this nature would be very beneficial to the field.

Criminal cognitions and attitudes have long been associated with criminal behavior and have been related to self-reported criminal activity in offenders (Healy & O'Donnell, 2006) and self-reported antisocial behavior in nonoffender samples (McCoy et al., 2006). Self-reported antisocial attitudes are associated with criminal activity (Mills & Kroner, 1997) and are predictive of both general and violent reoffending (Mills, Kroner, & Hemmati, 2004; Polaschek, Collie, & Walkey, 2004) and intimate partner violence (Henning, Martinsson, & Holdford, 2009). Among soccer fans in Europe, antisocial attitudes were strongly associated with self-reported verbal and physical aggression (van Hiel, Hautman, Cornelis, & de Clercq, 2007). Self-reported criminal attitudes add incrementally to the prediction of subsequent recidivism and disciplinary infractions over established risk assessment measures like the Psychopathy Checklist, LSI-R, and Statistical Information on Recidivism (Mills et al., 2004; Walters, 2009; Walters & Mandell, 2007; Walters & Schlauch, 2008).

Self-reported criminal attitudes can change, and those changes can be measured and related to outcome (Beggs & Grace, 2010). Measures of criminal thinking, such as the Psychological Inventory of Criminal Thinking Styles (PICTS; Walters, 2006) are sensitive to detecting treatment change (e.g., pre-post treatment measure) among offenders (Walters, 2002, 2009). Walters, Trgovac, Rychlec, Di Fazio, and Olson (2002) reported on a series of studies using the 13-item scale of Current Criminal Thinking items that contained items such as "I have trouble following through on good initial intentions." The results showed significant change following a 10-week criminal lifestyle program, whereas a wait-list control group did not evidence change during the same time frame. Elsewhere, Kroner and Yessine (in press) showed that criminal attitudes as measured by the Measures of Criminal Attitudes and Associates (MCAA; Mills, Kroner, & Forth, 2002) could change meaningfully between pre- and postintervention testing. Further, by using a more idiographic

methodology, Kroner and Yessine (in press) showed that change in attitudes toward criminal associates could be linked to changes in recidivism. Positive change in attitudes has been shown among offenders with mental disorders who participated in a cognitive skills program (Ashford, Wong, & Sternbach, 2008). Sex offender attitudes have also been shown to change for the better between pre- and postintervention testing (Keeling, Rose, & Beech, 2006; Witte, Di Palcido, Gu, & Wong, 2006).

Behavioral indicators provide a meaningful assessment of change. Common behavioral markers in corrections include frequency and severity of disciplinary infractions, work ratings, number of sick calls, and changes in security classification. In the community, treatment providers may look at sick days, number of days to employment or tardy for work, parole or probation visits kept/missed, and results of urinalysis. Analyses of these behavioral markers provide clinicians valuable information beyond interview or self-report, as they are indicative of functioning in the real world. Offenders may report improvement, but if those improvements are not being realized in the real-world setting, any change that is happening is less meaningful.

Multiconstruct Approach. In addition to the criminogenic construct of interest, other treatment related constructs are important to consider within a framework of intervention assessment. For example, the work of Wong et al. related to the VRS and Violence Risk Scale: Sex Offender Version (VRS:SO; Olver & Wong, 2011; Wong, Olver, & Stockdale, 2009) clearly demonstrates the importance of measuring an offender's progress along the continuum of the stages of change (precontemplation, contemplation, preparation, action, and maintenance) of the transtheoretical model of change (TTM; Prochaska, DiClemente, & Norcross, 1992). In fact, research on the VRS and VRS:SO has demonstrated that clinician-rated change on the criminogenic content (e.g., criminal attitudes) is related to reductions in recidivism only if the offender has meaningfully progressed along the stages of change continuum (e.g., moved from preparation to action stage). The important point here is that readiness to change is related to meaningful risk-reducing change resulting from therapeutic interventions. Motivational interventions are viewed as an essential component of interventions with patients who suffer from co-occurring disorders (Drake et al., 2001).

Readiness to change is related to offender program performance whether measured by self-report or by treatment facilitator ratings. As an example, participant rating or facilitator ratings of readiness to change were recorded in two separate samples of offenders provided correctional interventions to address intimate partner violence (Connors, Mills, & Gray, 2012, 2013). In both samples, these ratings were related to postprogram performance and demonstrated an interaction effect such that more motivated offenders acquired knowledge and skills at a significantly faster rate than less motivated offenders. This would suggest that readiness

to change is important and may be effectively measured using anchored ratings by either the participant or the therapist.

Another construct of importance when assessing change is therapeutic alliance. Although it is not consistently measured within correctional intervention research, it is related to outcomes in other areas of therapeutic intervention. An early meta-analysis by Horvath and Symonds (1991) revealed a modest but significant relationship between working alliance and outcomes in therapy. Brocato and Wagner (2008) found that therapeutic alliance was related to improvement in readiness to change but not offender retention in a community-based mandated drug treatment program. Nonetheless, offenders who scored higher on therapist bonding were more likely to experience increases in motivation to change, and higher motivation to change was associated with the number of days in the program. This finding suggests a possible indirect effect of therapeutic alliance on treatment retention. Elsewhere, in noncorrectional community mental health settings, therapeutic alliance has been found to be positively associated with outcomes (Howgego, Yellowlees, Owen, Meldrum, & Dark, 2003).

One of the better-known measures of working alliance is the Working Alliance Inventory (WAI; Horvath & Greenberg, 1989). Scores on the WAI are associated with client reported satisfaction with and change during therapy (Horvath & Greenberg, 1989). Further, relationships between the scores on the task subscale and subsequent skill acquisition during therapy were also observed. There are substantial intercorrelations among the subscales of tasks, goals, and bonds (r from .69 to .92), and scale internal consistency is very good ($\alpha = .93$). The WAI has been used with offenders with personality disorders receiving mental health treatment and measured at baseline, 6 months, and 24 months (Fortune et al., 2011).

Multi-time Approach. Pre–post-treatment testing or assessment is a standard within many treatment paradigms, and multi-time assessment of changeable risk factors is also becoming more commonplace (Hanson, Harris, Scott, & Helmus, 2007). Although pre–post testing/assessment is informative, there are two meaningful limitations. First, pre–post testing cannot identify when meaningful change occurred. For example, at what point in the treatment or intervention did motivation begin to improve? Second, as discussed previously, because interventions typically target changeable risk factors, not measuring or assessing those risk factors well after the intervention may fail to capture the dynamic nature of the risk factor and presumes that gains made during the intervention period continue unchanged from the point of postintervention testing. For example, an intervention targeting associates or attitudes may raise awareness and intent to change among offenders who have completed the intervention. The improvement between the pre- and posttesting is presumed static. However, if the offender is released to the same criminal subculture in which he or she had been previously involved, why would there be an expectation of stasis in the dynamic gains made? This problem

may explain the relatively few studies that have associated a change in pre–post intervention change with outcome (Kroner & Yessine, in press).

With these limitations in mind, we recommend that interventions, when possible, measure central constructs at multiple points during the intervention. Caution must be used when administering the same self-report instruments so as to avoid frequent exposures to the items; however, this could be overcome through the use of anchored facilitator ratings or assessments (Connors et al., 2012). In their assessment of an intimate partner violence program, Connors et al. assessed offender progress at pre-, mid-, and postintervention points. This permitted an examination of the trajectory of change over the course of the intervention. Additionally, postintervention assessment of these dynamic factors may provide a measure of the ongoing state of the dynamic change. However, based on previous experience, repeated measurement of dynamic risk factors over time can result in participant attrition if it employs methods that require much time or effort from the offender (Morgan, Kroner, & Mills, 2012).

Multi-Outcomes Approach. Traditional outcomes employed in correctional intervention research are the dichotomous measures of recidivism or institutional misconduct. More recently, other measures of recidivism have begun to be included as potential outcomes, including time before reoffense (Mills, Kroner, & Hemmati, 2007), number of offenses (Bushway, Paternoster, & Brame, 2003; Savolainen, 2009), and severity of reoffending. These latter outcomes stem from the introduction of desistance as a directional, gradual process of rehabilitation (Fagan, 1989; Sampson & Laub, 1993) wherein for some, desistance is a relative reduction in criminal activity as opposed to a complete cessation (Bushway et al., 2003; Laub, Nagin, & Sampson, 1998). Increased time prior to a less severe offense may be viewed by some as an improved outcome. Nonetheless, apart from a growing agreement that desistance is a process, there is little agreement on its measurement (Kazemian, 2007).

TESTING FOR OFFENDER CHANGE

We recommend a four-step strategy to test for offender change:

1. Pre–post significance testing
2. Examining the magnitude of effects
3. Clinical significance testing
4. Examining the reliability of change

(See Morgan et al., 2012b for an example of this strategy.)

Pre–Post Significance Testing. The use of pre–post treatment change is a staple in treatment outcome assessments. Although it has many benefits for identifying

if treatments work, this method of assessing outcomes is of limited utility for informing practitioners about what works and for whom. Although we continue to advocate for (and use in our own work) pre–post significance testing, additional limitations are noted—results are unduly influenced by sample size and are not amenable for examining individual changes. Thus, in addition to examining pre–post outcomes, we encourage the use of mediators and moderators as additional methods for assessing change (see, e.g., Kazdin, 2007).

Examining the Magnitude of Effect (With Confidence Intervals). Examining effect sizes offers advantages over traditional pre–post testing methods. Specifically, effect sizes have the benefit of being less influenced by sample size and allow for an examination of the size (magnitude) of treatment effects. The larger the effect size (i.e., numerical value), the greater the treatment impact. Cohen's d and the partial eta squared statistic are commonly used effect size procedures, and partial eta squared is easily obtained in most statistical packages.

Clinical Significance Testing. Clinical significance testing involves identifying the cases that have returned to a subclinical level of functioning as measured on standardized, psychometrically sound tests. Several methods can be used to examine clinical changes. One way is to compare the treated offender's score to the scores of the normative and/or clinical samples provided by the test. If the offender's initial (pretreatment) score is in the clinically significant range but the posttreatment score is in the nonclinically significant range, then treatment is deemed effective (i.e., clinical change has occurred). This procedure uses a cut-off score for the normative sample and a cut-off score for the clinical sample. Significant clinical change can also be assessed by examining the variance in posttreatment test scores compared to pretreatment test scores. Specifically, meaningful change is assumed if the posttreatment scores are 2 standard deviations from the pretreatment mean. The third calculation requires the practitioner to compute two cut-off calculations from data on normative and clinical samples of a test. One calculation involves identifying the overlap between the normative group and clinical group, and the other involves identifying where there is no overlap between these two groups. Thus, clinical significance can be calculated for each scale, and the number of clients who reach clinical significance can be computed. The aim of clinical significance testing is to identify clinically meaningful change that is beyond chance and also is of value to the practitioner.

An interesting and innovative method for examining clinical change is the use of percent of maximum possible (POMP) scores. Although the POMP is not a specific statistic, computation of POMP scores can help clinicians to understand the amount of change across individuals. Specifically, POMP scores can determine the percent of reduction in scores on the pre- and postmeasures. With a POMP score, a 0 represents the minimum possible score and 100 represents the maximum

possible score (see Cohen, Cohen, Aiken, & West, 1999, for the computation formula for POMP scores). Using POMP scores allows for interpreting the scale differences according to easily observed percentages and making direct comparisons among the scales (e.g., Gerend, Aiken, West, & Erchull, 2004; Srivastava, John, Gosling, & Potter, 2003).

Reliability of Change. The previous three methods of assessing change focus on an overall effectiveness of an intervention. Examining the reliability of change allows for an estimate of idiographic change. One drawback of using pre-intervention and postintervention scores (i.e., POMP scores) is not being able to calculate the confidence levels of the change scores associated with the intervention. In order to address this issue, Jacobson and Truax (1991) developed the Reliable Change Index (RCI), which allows for the assessment of change beyond what could be attributed to measurement variability or error. The RCI incorporates scale reliabilities into the calculation of reliable change, and the formula allows for a more precise measure of therapeutic change. Specifically, this calculation allows the identification of how much change has occurred as a result of the intervention and if this change is statistically reliable.

Using a RCI cut-off score of 1.96 or greater, the difference between pre- and post-scores is considered indicative of statistically significant (95% confidence interval) and clinically meaningful change (Wise, 2004). With this calculation, the percentage of participants of the sample who experienced “real” change can be derived. These calculations occur independently for each pre- and post- measure used. A noted assumption made by Jacobson et al. (1984) is that all the pretesting scores would be in a dysfunctional range.

CONCLUSIONS AND FUTURE DIRECTIONS

In this chapter, we reviewed the effectiveness of treatment with criminal offenders, including the barriers to effective interventions, evidence-based practice strategies, and treatment approaches for four specialized offender populations. In addition, we summarized the role of assessment in identifying relevant risk factors prior to initiating treatment as well as evaluating offender change during the course of treatment. This review highlighted the remarkable accomplishments in the field of offender treatment and correctional mental health over the last 40 years; however, important advances remain to be made.

Meaningful information is missing from important areas that are significant considerations in offender treatment. For example, we did not discuss issues of the iatrogenic effects of criminal sanctions, incarceration, and provision of mental health services to incarcerated offenders, as this issue has been severely understudied. We also focused minimally on individual characteristics of offenders that contribute to or inhibit therapeutic progress, because very little is known about these phenomena.

Last, almost no research has examined the assessment change process with offenders. In fact, research examining the relationship between therapeutic change and long-term outcomes (e.g., the relationship of change resulting from prison treatment program to community outcomes) is almost nonexistent. Although this research is costly (longitudinal research of this nature is time intensive), it is necessary, as it is no longer acceptable simply to examine change following treatment without examining long-term outcomes. Research examining the relationship between therapeutic change and outcomes is important from a mental health, humanistic, public safety, and policy perspective.

Despite the need to improve our knowledge in these areas, there is much to be positive about. As this chapter shows, correctional interventions are not only effective but are effective with some of the most difficult correctional clients, such as mentally ill and violent offenders. Evidence-based practices that produce positive outcomes have been identified, and research continues to shed light on how best to treat offenders. The future of correctional treatment is bright. There is a plethora of work already being conducted across the world as well as an unprecedented number of undergraduate and graduate students eager to carry the field forward. We predict that the interface of criminal justice and mental health treatment will advance significantly over the next 20 years.

REFERENCES

- Andrews, D. A., & Bonta, J. (1994). *The psychology of criminal conduct*. Cincinnati, OH: Anderson.
- Andrews, D. A., & Bonta, J. (1995). *The Level of Service Inventory-Revised (manual)*. Toronto, Canada: Multi-Health Systems.
- Andrews, D. A., & Bonta, J. (2010). *The psychology of criminal conduct* (5th ed.). New Providence, NJ: Matthew Bender.
- Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, *17*, 19–52.
- Andrews, D. A., Bonta, J., & Wormith, J. S. (2004). *Manual for the Level of Service/Case Management Inventory (LS/CMI)*. Toronto, Canada: Multi-Health Systems.
- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta J., Gendreau P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, *28*, 369–404.
- Antonacci, D. J., Bloch, R. M., Saeed, S. A., Yildirim, Y., & Talley, J. (2008). Empirical evidence on the use and effectiveness of telepsychiatry via videoconferencing: Implications for forensic and correctional psychiatry. *Behavioral Sciences and the Law*, *26*, 253–269. doi:10.1002/bsl.812
- Arndt, S., Turvey, C. L., & Flaum, M. (2002). Older offenders, substance use, and treatment. *American Journal of Geriatric Psychology*, *10*, 733–739.
- Ashford, J. B., Wong, K. W., & Sternbach, K. O. (2008). Generic correctional programming for mentally ill offenders: A pilot study. *Criminal Justice and Behavior*, *35*, 457–473.
- Ax, R. K., Fagan, T. J., & Holton, S.M.B. (2002). Individuals with serious mental illnesses in prison: Rural perspectives and issues. In S. B. Hudnall (Ed.) *Rural behavioral health*

- care: *An interdisciplinary guide* (pp. 203–215). Washington, DC: American Psychological Association.
- Ax, R. K., Fagan, T. J., Magaletta, P. R., Morgan, R. D., Nussbaum, D., & White, T. W. (2007). Innovations in correctional assessment and treatment. *Criminal Justice and Behavior*, *34*, 893–905. doi: 10.1177/0093854807301555
- Aytes, K. E., Olsen, S. S., Zakrajsek, T., Murray, P., & Ireson, R. (2001). Cognitive/behavioral treatment for sexual offenders: An examination of recidivism. *Sexual Abuse: Journal of Research and Treatment*, *13*, 223–231.
- Babcock, J. C., & Steiner, R. (1999). The relationship between treatment, incarceration, and recidivism of battering: A program evaluation of Seattle's coordinated community response to domestic violence. *Journal of Family Psychology*, *13*, 46–59.
- Barron, P., Hassiotis, A., & Banes, J. (2002). Offenders with intellectual disability: The size of the problem and therapeutic outcomes. *Journal of Intellectual Disability Research*, *46*(6), 454–463. doi:10.1046/j.1365-2788.2002.00432.x
- Batastini, A. B., Hunt, E., Present-Koller, J., & DeMatteo, D. (2011). Federal standards for community registration of juvenile sex offenders: An evaluation of risk prediction and future implications. *Psychology, Public Policy, and Law*, *17*, 451–474. doi: 10.1037/a0023637
- Batastini, A. B., McDonald, B., & Morgan, R. D. (2013). Videoteleconferencing in forensic and correctional practice. In K. Myers & C. Turvey (Eds.), *Telemental health: Clinical, technological, and administrative foundations of evidence-based practice*, pp. 251–272. Waltham, MA: Elsevier.
- Beggs, S. M., & Grace, R. C. (2010). Assessment of dynamic risk factors: An independent validation study of the Violence Risk Scale: Sexual Offender Version. *Sexual Abuse: A Journal of Research and Treatment*, *22*, 234–251.
- Berry, S. (2003). Stopping violent offending in New Zealand: Is treatment an option? *New Zealand Journal of Psychology*, *32*, 92–100.
- Bewley, M. T., & Morgan, R. D. (2011). A national survey of mental health services available to offenders with mental illness: Who is doing what? *Law & Human Behavior*, *35*, 351–363.
- Birmingham, L., Mason, D., & Grubin, D. (1996). Prevalence of mental disorder in remand prisoners: Consecutive case study. *British Medical Journal*, *313*, 1521–1524.
- Bonta, J., Wallace-Capretta, S., & Rooney, J. (2000). A quasi-experimental evaluation of an intensive rehabilitation supervision program. *Criminal Justice and Behavior*, *27*, 312–329.
- Borum, R. (2003). Managing at-risk juvenile offenders in the community: Putting evidence-based principles into practice. *Journal of Contemporary Criminal Justice*, *19*, 114–137. doi:10.1177/1043986202239745
- Bourgon, G., & Armstrong, B. (2005). Transferring the principles of effective treatment into a “real world” prison setting. *Criminal Justice and Behavior*, *32*, 3–25.
- Brocato, J., & Wagner, E. F. (2008). Predictors of retention in an alternative-to-prison substance abuse treatment program. *Criminal Justice and Behavior*, *35*, 99–119.
- Brown, T. L., Henggeler, S. W., Schoenwald, S. K., Brondino, M. J., & Pickrel, S. G. (1999). Multisystemic treatment of substance abusing and dependent juvenile delinquents: Effects on school attendance at posttreatment and 6-month follow-up. *Children's Services: Social Policy, Research, and Practice*, *2*, 81–93.
- Burdon, W. M., Messina, N. P., & Prendergast, M. L. (2004). The California treatment expansion initiative: Aftercare participation, recidivism, and predictors of outcomes. *Prison Journal*, *84*, 61–80.

- Bushway, S. D., Paternoster, R., & Brame, R. (2003). Examining the prevalence of criminal desistance. *Criminology*, *41*, 423–448.
- Butzin, C. A., Saum, C. A., & Scarpitti, F. R. (2002). Factors associated with completion of a drug treatment court diversion programme. *Substance Use and Misuse*, *37*, 1615–1633.
- Chambers, J. C., Eccleston, L., Day, A., Ward, T., & Howells, K. (2008). Treatment readiness in violent offenders: The influence of cognitive factors on engagement in violence programs. *Aggression and Violent Behavior*, *13*, 276–284.
- Cohen, F., & Dvoskin, J. A. (1992). Inmates with mental disorders: A guide to law and practice. *Mental & Physical Disability Law Reporter*, *16*(3–4), 339–346.
- Cohen, P., Cohen, J., Aiken, L. S., & West, S. G. (1999). The problem of units and the circumstance for POMP. *Multivariate Behavioral Research*, *34*(3), 315–346. doi:10.1207/S15327906MBR3403_2
- Connors, A. D., Mills, J. F., & Gray, A. L. (2012). An evaluation of intimate partner violence intervention with incarcerated offenders. *Journal of Interpersonal Violence*, *27*, 1176–1196. doi:10/1177/0886260511424499
- Connors, A. D., Mills, J. F., & Gray, A. L. (2013). Intimate partner violence intervention for high-risk offenders. *Psychological Services*, *10*, 12–23.
- Corrigan, P. W., Mueser, K. T., Bond, G. R., Drake, R. E., & Solomon, P. (2008). *Principles and practice of psychiatric rehabilitation: An empirical approach*. New York, NY: Guilford Press.
- Cortoni, F., Nunes, K., & Latendresse, M. (2006). *An examination of the effectiveness of the violence prevention program (R-178)*. Ottawa, Canada: Correctional Service of Canada.
- Daniel, A. E. (2007). Care of the mentally ill in prisons: Challenges and solutions. *Journal of the American Academy of Psychiatry and the Law*, *35*, 406–410.
- Deane, F. P., Skogstad, P., & Williams, M. W. (1999). Impact of attitudes, ethnicity and quality of prior therapy on New Zealand male prisoners' intentions to seek professional psychological help. *International Journal for the Advancement of Counseling*, *21*, 55–67.
- DeMatteo, D., Hunt, E., Batastini, A. B., & LaDuke, C. (2010). The disconnect between assessment and intervention in the risk management of criminal offenders. *Open Access Journal of Forensic Psychology*, *2*. Retrieved from http://www.forensicpsychologyunbound.ws/OAJFP/Volume_2_2010_files/DeMatteo%202010.pdf
- Di Placido, C., Simon, T., Witte, T., Gu, D., & Wong, S. C. P. (2006). Treatment of gang members can reduce recidivism and institutional misconduct. *Law and Human Behavior*, *30*(1), 93–114.
- Drake, R. E., Essock, S. M., Shaner, A., Carey, K. B., Minkoff, K., Kola, L., . . . Rickards, L. (2001). Implementing dual diagnosis services for clients with severe mental illness. *Psychiatric Services*, *52*, 469–476.
- Dvoskin, J. A., & Morgan, R. D. (2010). Correctional psychology. In I. Weiner & W. E. Craighead (Eds.), *Corsini encyclopedia of psychology* (Vol. 1, pp. 417–420). Hoboken, NJ: Wiley.
- Empey, D. D., & Gordon, D. A. (1989). Attrition from court-referred family therapy in three rural community mental health centers. *Juvenile and Family Court Journal*, *40*, 45–54.
- Fabiano, R., Robinson, D., & Proporino, F. (1990). *A preliminary assessment of the cognitive skills training programme: A component of living skills program. Programme description, research findings and implementation strategy*. Ottawa, Canada: Correctional Service Canada.
- Fagan, J. (1989). *Cessation of family violence: Deterrence and dissuasion*. In L. Ohlin & M. Tonry (Eds.), *Crime and justice: A review of research*. Chicago, IL: University of Chicago Press.

- Fisher, D., Beech, A., & Browne, K. (2000). The effectiveness of relapse prevention training in a group of incarcerated child molesters. *Psychology, Crime & Law*, *6*, 181–195.
- Fortune, Z., Barrett, B., Armstrong, D., Coid, J., Crawford, M., Mudd, D.,... Moran, P. (2011). Clinical and economic outcomes from the UK pilot psychiatric services for personality-disordered offenders. *International Review of Psychiatry*, *23*, 61–69.
- Fox, K. C., Connor, P., McCullers, E., & Waters, T. (2008). Effect of a behavioural health and specialty care telemedicine programme on goal attainment for youth in juvenile detention. *Journal of Telemedicine and Telecare*, *14*, 227–230. doi: 10.1258/jtt.2008.071102
- Freeman, A., & Rosenfield, B. (2002). Modifying therapeutic homework for patients with personality disorders. *Journal of Clinical Psychology*, *58*, 513–524. doi:10.1002/jclp.10029
- Garrity, T. F., Hiller, M. L., Staton, M., Webster, J. M., & Leukefeld, C. G. (2002). Factors predicting illness and health services use among male Kentucky prisoners with a history of drug abuse. *Prison Journal*, *82*, 295–313.
- Geer, T. M., Becker, J. V., Gray, S. R., & Krauss, D. (2001). Predictors of treatment completion in a correctional sex offender treatment program. *International Journal of Offender Therapy and Comparative Criminology*, *45*, 302–313.
- Gendreau, P. (1996a). Offender rehabilitation: What we know and what needs to be done. *Criminal Justice and Behavior*, *23*, 144–161.
- Gendreau, P. (1996b). *The principles of effective intervention with offenders*. In A. T. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 117–130). Thousand Oaks, CA: Sage.
- Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology*, *34*, 575–607.
- Gendreau, P., & Smith, P. (2007). Influencing the “people who count”: Some perspectives on the reporting of meta-analytic results for prediction and treatment outcomes with offenders. *Criminal Justice and Behavior*, *34*, 1536–1559. doi:10.1177/0093854807307025
- Gerend, M. A., Aiken, L. S., West, S. G., & Erchull, M. J. (2004). Beyond medical risk: Investigating the psychological factors underlying women’s perceptions of susceptibility to breast cancer, heart disease, and osteoporosis. *Health Psychology*, *23*(3), 247–258. doi:10.1037/0278–6133.23.3.247
- Girard, L., & Wormith, J. S. (2004). The predictive validity of the Level of Service Inventory–Ontario Revision on general and violent recidivism among various offender groups. *Criminal Justice and Behavior*, *31*, 150–181. doi:10.1177/0093854803261335
- Glasmeier, A. K., & Farrigan, T. (2007). The economic impacts of the prison development boom on persistently poor rural places. *International Regional Science Review*, *30*, 274–299. doi: 10.1177/0160017607301608
- Glaze, L. E. (2011). *Correctional population in the United States, 2010* (Bureau of Justice Statistics Report No. NCJ 236319). Washington DC: Department of Justice.
- Gossop, M., Marsden, J., Stewart, D., & Rolfe, A. (1999). Treatment retention and 1 year outcomes for residential programs in England. *Drug and Alcohol Dependence*, *57*, 89–98.
- Hamberger, L. K., Lohr, J. M., & Gottlieb, M. (2000). Predictors of treatment dropout from a spouse abuse program. *Behavior Modification*, *24*, 528–552.
- Hambridge, J. A. (1990). Use of an outpatient forensic psychology service: The non-attenders. *International Journal of Offender Therapy and Comparative Criminology*, *34*, 165–169.
- Hanson, R. K., Gordon, A., Harris A.J.R., Marques, J. K., Murphy, W., Quinsey, V. L., & Seto, M. C. (2002). First report of the collaborative outcome data project on the effectiveness

- of psychological treatment for sex offenders. *Sexual Abuse: A Journal of Research and Treatment*, 14, 169–197.
- Hanson, R. K., Harris, A. J. R., Scott, T., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project* (Corrections User Report No. 2007–05). Ottawa, Canada: Public Safety Canada.
- Harlow, C. W. (2003). *Education and correctional populations*. Bureau of Justice Statistics, Special Report (NCJ 195670). Washington DC: U.S. Department of Justice, Office of Justice Programs.
- Healy, D., & O'Donnell, I. (2006). Criminal thinking on probation: A perspective from Ireland. *Criminal Justice and Behavior*, 33, 782–802.
- Henggeler, S. W., Pickrel, S. G., Brondino, M. J., & Crouch, J. L. (1996). Eliminating (almost) treatment dropout of substance abusing or dependent delinquents through home-based multisystemic therapy. *American Journal of Psychiatry*, 153, 427–428.
- Henggeler, S. W., Clingempeel, W. G., Brondino, M. J., & Pickrel, S. G. (2002). Four-year follow-up of multisystemic therapy with substance-abusing and substance-dependent juvenile offenders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41, 868–874. doi: 0.1097/00004583-200207000-00021
- Henning, K., Martinsson, R., & Holdford, R. (2009). Gender differences in risk factors for intimate partner violence recidivism. *Journal of Aggression, Maltreatment & Trauma*, 18, 623–645.
- Hepburn, J. R. (2005). Recidivism among drug offenders following exposure to treatment. *Criminal Justice Policy Review*, 16, 237–259. doi:10.1177/0887403404270604
- Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York, NY: Free Press.
- Hiller, M. L., Knight, K., Saum, C. A., & Simpson, D. D. (2006). Social functioning, treatment dropout, and recidivism of probationers mandated to a modified therapeutic community. *Criminal Justice and Behavior*, 33, 738–759. doi:10.1177/0093854806288242
- Hodgins, S., Müller-Isberner, R., Freese, R., Tiihonen, J., Repo-Tiihonen, E., Eronen, M., . . . Kronstrand, R. (2007). A comparison of general adult and forensic patients with schizophrenia living in the community. *International Journal of Forensic Mental Health*, 6, 63–75. doi:10.1080/14999013.2007.10471250
- Hollin, C. R., & Palmer, E. J. (2006). The Level of Service Inventory–Revised profile of English prisoners: Risk and reconviction analysis. *Criminal Justice and Behavior*, 33, 347–366. doi:10.1177/0093854806286195
- Hooks, G., Mosher, C., Rotolo, T., & Lobao, L. (2004). The prison industry: Carceral expansion and employment in U.S. counties, 1969–1994. *Social Science Quarterly*, 85, 37–56.
- Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36, 223–233.
- Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*, 38, 139–149.
- Howerton, A., Byng, R., Campbell, J., Hess, D., Owens, C., & Aitken, P. (2007). Understanding help seeking behaviour among male offenders: Qualitative interview study. *British Medical Journal*, 334, 303.
- Howgego, I. M., Yellowlees, P., Owen, C., Meldrum, L., & Dark, F. (2003). The therapeutic alliance: The key to effective patient outcome? A descriptive review of the evidence

- in community mental health case management. *Australian and New Zealand Journal of Psychiatry*, 37, 169–183.
- Huebner, B. M., & Cobbina, J. (2007). The effect of drug use, drug treatment participation, and treatment completion on probationer recidivism. *Journal of Drug Issues*, 37(3), 619–641.
- Hunter, J. A., & Figueredo, A. J. (1999). Factors associated with treatment compliance in a population of juvenile sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 11, 49–67.
- Jacobson, N. S., Follette, W. C., Revenstorf, D., Hahlweg, K., Baucom, D. H., & Margolin, G. (1984). Variability in outcome and clinical significance of behavioral marital therapy: A reanalysis of outcome data. *Journal of Consulting and Clinical Psychology*, 52(4), 497–504. doi:10.1037/0022-006X.52.4.497
- Jacobson, N. S., & Traux, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12–19.
- James, D. J., & Glaze, L. E. (2006). *Mental health problems of prison and jail inmates* (Bureau of Justice Statistics Special Report, NCJ 213600). Washington, DC: Department of Justice.
- Kazantzis, N., & Lampropoulos, G. K. (2002). The use of homework in psychotherapy: An introduction. *Journal of Clinical Psychology*, 58, 487–488.
- Kazdin, A. E. (1994). Methodology, design, and evaluation in psychotherapy research. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy behavior change* (pp. 19–71). New York, NY: Wiley.
- Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology*, 3, 1–27.
- Kazemian, L. (2007). Desistance from crime: Theoretical, empirical, methodological, and policy considerations. *Journal of Contemporary Criminal Justice*, 23, 5–27. doi:10.1177/1043986206298940
- Keeling, J. A., Rose, J. L., & Beech, A. R. (2006). An investigation into the effectiveness of a custody-based cognitive-behavioural treatment for special needs sexual offenders. *Journal of Forensic Psychiatry and Psychology*, 17, 372–392.
- King, R. S., Mauer, M., & Hurling, T. (2004). An analysis of the economics of prison siting in rural communities. *Criminology & Public Policy*, 3, 453–480. doi: 10.1111/j.1745-9133.2004.tb00054.x
- Kinsella, C. (2004). *Corrections health care costs*. Retrieved from www.csg.org/knowledge-center/docs/TA0401CorrHealth.pdf
- Kroner, D. G., & Loza, W. (2001). Evidence for the efficacy of self-report in predicting violent and nonviolent criminal recidivism. *Journal of Interpersonal Violence*, 16, 168–177.
- Kroner, D. G., & Mills, J. F. (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior*, 28, 471–489.
- Kroner, D. G., & Yessine, A. K. (in press). Changing risk factors that impact recidivism: In search of mechanisms of change. *Law and Human Behavior*.
- Lamb, H. R., & Bachrach, L. L. (2001). Some perspectives on deinstitutionalization. *Psychiatric Services*, 52, 1039–1045. doi:10.1176/appi.ps.52.8.1039
- Lamb, H. R., & Weinberger, L. E. (1998). Persons with severe mental illness in jails and prisons: A review. *Psychiatric Services*, 49, 483–492.

- Landenberger, N., & Lipsey, M. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology*, 1(4), 451-476. doi:10.1007/s11292-005-3541-7
- Larsen, D., Stamm, B. H., Davis, K., & Magaletta, P. R. (2004). Prison telemedicine and TMH utilization in the United States: State and federal perceptions of benefits and barriers. *Telemedicine and e-Health*, 10, 81-89. doi:10.1089/1530562042632056
- Laub, J. H., Nagin, D. S., & Sampson, R. J. (1998). Trajectories of change in criminal offending: Good marriages and the desistance process. *American Sociological Review*, 63, 225-238.
- Leak, G. K. (1980). Effects of highly structured versus nondirective group counseling approaches on personality and behavioral measures of adjustment in incarcerated felons. *Journal of Counseling Psychology*, 27, 520-523.
- Lexcen, F. J., Hawk, G. L., Herrick, S., & Blank, M. B. (2006). Use of video conferencing for psychiatric and forensic evaluations. *Psychiatric Services*, 57, 713-715. doi:10.1176/appi.ps.57.5.713
- Lipsey, M. W. (1989, November). *The efficacy of intervention for juvenile delinquency: Results from 400 studies* Paper presented at the 41st annual meeting of the American Society of Criminology. Reno, NV.
- Lipton, D. S. (1995). *The effectiveness of treatment for drug abusers under criminal justice supervision* (NCJ 157642). Washington, DC: National Institute of Justice.
- Lowenkamp, C. T., Holsinger, A. M., & Latessa, E. J. (2001). Risk/need assessment, offender classification, and the role of childhood abuse. *Criminal Justice and Behavior*, 28, 543-563. doi:10.1177/009385480102800501
- Lowes, R. (2001). Telemedicine. *Medical Economics*, 78, 24.
- Loza, W. (2005). *Self-Appraisal Questionnaire (manual)*. Toronto, Canada: Multi-Health Systems.
- Loza, L., Cumbleton, A., Shahinfar, A., Neo, L. H., Evens, M., Conley, M., & Summers, R. (2004). Cross-validation of the Self-Appraisal Questionnaire (SAQ): An offender risk and need assessment measure on Australian, British, Canadian, Singaporean, and American Offenders. *Journal of Interpersonal Violence*, 19, 1172-1190.
- Loza, W., & Loza-Fanous, A. (2001). The effectiveness of the Self-Appraisal Questionnaire (SAQ) in predicting offenders' post-release outcome: A comparison study. *Criminal Justice and Behavior*, 28, 105-121.
- Loza, W., & Loza-Fanous, A. (2003). More evidence for the validity of the Self-Appraisal Questionnaire for predicting violent and non-violent recidivism: A 5-year follow-up study. *Criminal Justice and Behavior*, 30, 709-721.
- Lund, J. (1990). Mentally retarded criminal offenders in Denmark. *British Journal of Psychiatry*, 156, 726-731. doi:10.1192/bjp.156.5.726
- Magaletta, P. R., Fagan, T. J., & Ax, R. K. (1998). Advancing psychology services through TMH in the Federal Bureau of Prisons. *Professional Psychology: Research and Practice*, 29, 543-548.
- Magaletta, P. R., Fagan, T. J., & Peyrot, M. F. (2000). TMH in the federal bureau of prisons: Inmates' perceptions. *Professional Psychology: Research and Practice*, 31, 497-502. doi:10.1037/0735-7028.31.5.497
- Manchak, S. M., Skeem, J. L., & Douglas, K. S. (2007). Utility of the Revised Level of Service Inventory (LSI-R) in predicting recidivism after long-term incarceration. *Law and Human Behavior*, 32, 477-488. doi:10.1007/s10979-007-9118-4

- Martin, M. S., Dorken, S. K., Wamboldt, A. D., & Wootten, S. E. (2012). Stopping the revolving door: A meta-analysis on the effectiveness of interventions for criminally involved individuals with major mental disorders. *Law and Human Behavior, 36*, 1–12. doi: 10.1007/s10979-011-9274-4
- Mathias, R. E., & Sindberg, R. (1985). Psychotherapy in correctional settings. *International Journal of Offender Therapy and Comparative Criminology, 29*, 265–275.
- McCoy, K., Fremouw, W., Tyner, E., Clegg, C., Johansson-Love, J., & Strunk, J. (2006). Criminal thinking styles and illegal behavior among college students: Validation of the PICTS. *Journal of Forensic Science, 51*, 1174–1177.
- McDonald, B. R., & Morgan, R. D. (2012). *Enhancing homework compliance in correctional psychotherapy* (Unpublished manuscript).
- McGrath, R. J., Cumming, G., Livingston, J. A., & Hoke, S. E. (2003). Outcome of a treatment program for adult sex offenders. *Journal of Interpersonal Violence, 18*, 3–17.
- McGuire, J. (2002). Integrating findings from research reviews. In J. McGuire (Ed.), *Offender rehabilitation and treatment: Effective programmes and policies to reduce re-offending* (pp. 3–38). Chichester, England: Wiley.
- McMurran, M., & McCulloch, A. (2007). Why don't offenders complete treatment? Prisoners' reasons for non-completion of a cognitive skills programme. *Psychology, Crime & Law, 13*, 345–354.
- McMurran, M., & Theodosi, E. (2007). Is treatment non-completion associated with increased reconviction over no treatment? *Psychology, Crime & Law, 13*, 333–343.
- Meisenhelder, T. (1977). An explanatory study of exiting from criminal careers. *Criminology, 15*, 319–334.
- Milgram, D., & Rubin, J. S. (1992). Resisting the resistance: Involuntary substance abuse group therapy. *Social Work with Groups, 15*, 95–110.
- Miller, T. W., Clark, J., Veltkamp, L. J., Burton, D. C., & Swope, M. (2008). Teleconferencing model for forensic consultation, court testimony, and continuing education. *Behavioral Sciences and the Law, 26*, 301–313. doi:10.1002/bsl.809
- Mills, J. F., Jones, M. N., & Kroner, D. G. (2005). An examination of the generalizability of the LSI-R and VRAG probability bins. *Criminal Justice and Behavior, 32*, 565–585. doi:10.1177/0093854805278417
- Mills, J. F., & Kroner, D. G. (1997). The Criminal Sentiments Scale: Predictive validity in a sample of violent and sex offenders. *Journal of Clinical Psychology, 53*, 399–404.
- Mills, J. F., Kroner, D. G., & Forth, A. E. (2002). Measures of criminal attitudes and associates (MCAA): Development, factor structure, reliability, and validity. *Assessment, 9*, 240–253.
- Mills, J. F., Kroner, D. G., & Hemmati, T. (2004). The Measures of Criminal Attitudes and Associates (MCAA): The prediction of general and violent recidivism. *Criminal Justice and Behavior, 31*, 717–733.
- Mills, J. F., Kroner, D. G., & Hemmati, T. (2007). The validity of violence risk estimates: An issue of item performance. *Psychological Services, 4*, 1–12.
- Mills, J. F., Kroner, D. G., & Morgan, R. D. (2011). *The clinician's guide to violence risk assessment*. New York, NY: Guilford Press.
- Monnier, J., Knapp, R. G., & Frueh, B. C. (2003). Recent advances in telepsychiatry: An updated review. *Psychiatric Services, 54*, 1604–1609.
- Moore, D. L., Bergman, B. A., & Knox, P. L. (1999). Predictors of sex offender treatment completion. *Journal of Child Sexual Abuse, 7*, 73–88.

- Morgan, R. D. (2003). Basic mental health services: Services and issues. In T. Fagan & R. K. Ax (Eds.), *Correctional mental health handbook* (pp. 59–71). Thousand Oaks, CA: Sage.
- Morgan, R. D., Fisher, W. H., Duan, N., Mandracchia, J. T., & Murray, D. (2010). Prevalence of criminal thinking among state prison inmates with serious mental illness. *Law & Human Behavior, 34*, 324–336. doi:10.1007/s10979-009-9182-z
- Morgan, R. D., & Flora, D. B. (2002). Group psychotherapy with incarcerated offenders: A research synthesis. *Group Dynamics: Theory, Research, and Practice, 6*, 203–218. doi:10.1037//1089-2699.6.3.203
- Morgan, R. D., Flora, D. B., Kroner, D. G., Mills, J. F., Varghese, F., & Steffan, J. S. (2012). Treating offenders with mental illness: A research synthesis. *Law and Human Behavior, 36*, 37–50. doi:10.1007/s10979-011-9271-7
- Morgan, R. D., Kroner, D. G., & Mills, J. F. (2006). Group psychotherapy in prison: Facilitating change inside the walls. *Journal of Contemporary Psychotherapy, 36*, 137–144. doi:10.1007/BF02729057
- Morgan, R. D., Kroner, D. G., & Mills, J. F. (2012). *Re-entry: Dynamic risk assessment* (Technical Report). Washington, DC: National Institute of Justice.
- Morgan, R. D., Kroner, D. G., Mills, J. F., & Bauer, R. (2012). *Changing lives and changing outcomes: A treatment guide for offenders with mental illness* (Unpublished treatment manual).
- Morgan, R. D., Kroner, D. G., Mills, J. F., Bauer, R., & Serna, C. (2012). *Treating justice involved persons with mental illness: Preliminary evaluation of a comprehensive treatment program* (Unpublished manuscript).
- Morgan, R. D., Patrick, A. R., & Magaletta, P. R. (2008). Does the use of telemental health alter the treatment experience? Inmates' perceptions of telemental health versus face-to-face treatment modalities. *Journal of Consulting and Clinical Psychology, 76*, 158–162. doi:10.1037/0022-006X.76.1.158
- Morgan, R. D., Rozycki, A. T., & Wilson, S. (2004). Inmate perceptions of mental health services. *Professional Psychology: Research and Practice, 35*, 389–396.
- Morgan, R. D., Steffan, J. S., Shaw, L., & Wilson, S. (2007). Needs for and barriers to correctional mental health services: Inmate perceptions. *Psychiatric Services, 58*, 1181–1186. doi:10.1176/appi.ps.58.9.1181
- Mulvey, E., & Iselin, A. (2008). Improving professional judgments of risk and amenability in juvenile justice. *Future of Children, 18*, 35–57.
- National Institute of Justice. (1999, March). *Telemedicine can reduce correctional health care costs: An evaluation of a prison telemedicine network*. Retrieved from www.ncjrs.gov/pdffiles1/175040.pdf
- National Institute of Justice. (2002, May). *Implementing telemedicine in correctional facilities*. Retrieved April 5, 2012 from www.ncjrs.gov/pdffiles1/nij/190310.pdf
- Nedd, A.N.B., & Shilhadeh, E. S. (1974). The impact of group therapy on penitentiary vocational and academic education programs. *Journal of Psychology, 88*, 215–221.
- Nelson, E., Zaylor, C., & Cook, D. (2004). A comparison of psychiatrist evaluation and patient symptom report in a jail telepsychiatry clinic. *Telemedicine Journal and e-Health, 10*, 54–59. doi:10.1089/1530562042631967
- Newbern, D., Dansereau, D. F., & Pitre, U. (1999). Positive effects on life skills, motivation, and self-efficacy: Node-link maps in a modified therapeutic community. *American Journal of Drug and Alcohol Abuse, 25*, 407–423.

- Olver, M. E., & Wong, S. C. P. (2011). A comparison of static and dynamic assessment of sexual offender risk and need in a treatment context. *Criminal Justice and Behavior, 38*, 113–126.
- Palomares, R. S., (2012, August). Relevant laws, regulations, & policies relating to telepsychology practice. In L. Campbell, F. Millián, and R. S. Palomares, *Draft guidelines for the practice of telepsychology*. Symposium conducted at the meeting of the American Psychological Association, Orlando, FL.
- Pearson, F. S., Lipton, D. S., Cleland, C. M., & Yee, D. S. (2002). The effects of behavioral/cognitive-behavioral programs on recidivism. *Crime and Delinquency, 48*, 479–496.
- Pelissier, B., Camp, S. D., & Motivans, M. (2003). Staying in treatment: How much difference is there from prison to prison? *Psychology of Addictive Behaviors, 17*, 134–141.
- Polaschek, D.L.L., & Dixon, B. G. (2001). The violence prevention project: The development and evaluation of a treatment programme for violent offenders. *Psychology, Crime & Law, 7*(1), 1–23. doi:10.1080/10683160108401779
- Polaschek, D.L.L., Collie, R. M., & Walkey, F. H. (2004). Criminal attitudes to violence: Development and preliminary validation of a scale for male prisoners. *Aggressive Behavior, 30*, 484–503.
- Polizzi, D. M., MacKenzie, D. L., & Hickman, L. J. (1999). What works in adult sex offender treatment? A review of prison- and non-prison-based treatment programs. *International Journal of Offender Therapy and Comparative Criminology, 43*, 357–374.
- Prendergast, M. L., Hall, E. A., Wexler, H. K., Melnick, G., & Cao, Y. (2004). Amity prison-based therapeutic community: 5-year outcomes. *Prison Journal, 84*, 36–60.
- Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to the addictive behaviors. *American Psychologist, 47*, 1102–1114.
- Proporino, F. J., & Robinson, D. (1995). An evaluation of the reasoning and rehabilitation programme with Canadian federal offenders. In R. R. Ross & R. D. Ross (Eds.), *Thinking straight: The reasoning and rehabilitation programme for delinquency prevention and offender rehabilitation* (pp. 155–191). Ottawa, Canada: AIR.
- Rappaport, R. G. (1971). Group psychotherapy in prison. *International Journal of Group Psychotherapy, 21*, 265–274.
- Redding, R. E., Lexcen, F. J., & Ryan, E. P. (2005). Mental health treatment for juvenile offenders in residential psychiatric and juvenile justice settings. In K. Heilbrun, N.E.S. Goldstein, & R. E. Redding (Eds.), *Juvenile delinquency: Prevention, assessment, and intervention* (pp. 282–309). New York, NY: Oxford University Press.
- Reinsmith-Meyer, C. L. (2008). Barriers to mental health and substance abuse treatment among incarcerated offenders. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 69*, 1970.
- Richards, H. J., Casey, J. O., & Lecente, S.W. (2003). Psychopathy and treatment response in incarcerated female substance abusers. *Criminal Justice and Behaviour, 30*, 251–276.
- Riordan, R. J. & Martin, M. H. (1993). Mental health counseling and the mandated client. *Journal of Mental Health Counseling, 15*, 373–383.
- Robinson, D., & Proporino, F. J. (2004). Programming in cognitive skills: The reasoning and rehabilitation programme. In C. R. Hollin (Ed.), *The essential handbook of offender assessment and treatment* (pp. 63–77). Hoboken, NJ: Wiley.

- Romani, C. J., Morgan, R. D., Gross, N. R., & McDonald, B. R. (2012). Treating criminal behavior: Is the bang worth the buck? *Psychology, Public Policy and Law*, *18*(1), 144–165. doi:10.1037/a0024714
- Rooney, J., & Hanson, R. K. (2001). Predicting attrition from treatment programs for abusive men. *Journal of Family Violence*, *16*, 131–149. doi:10.1023/A:1011106902465
- Ross, R. R., Fabiano, E. A., & Ewles, C. D. (1988). Reasoning and rehabilitation. *International Journal of Offender Therapy and Comparative Criminology*, *32*, 29–36.
- Ruiz v. Estelle, 503 F. Supp. 1265 (S.D. Texas 1980).
- Sampson, R., and Laub, J. (1993). *Crime in the making: Pathways and turning points through life*. Cambridge, MA: Harvard University Press.
- Savolainen, J. (2009). Work, family and criminal desistance. *British Journal of Criminology*, *49*, 285–304. doi:10.1093/bjc/azn084
- Seager, J. A., Jellicoe, D., & Dhaliwal, G. K. (2004). Refusers, dropouts, and completers: Measuring sex offender treatment efficacy. *International Journal of Offender Therapy and Comparative Criminology*, *48*, 600–612. doi:10.1177/0306624X04263885
- Serin, R. C., & Preston, D. L. (2001). Managing and treating violent offenders. In J. B. Ashford, B. D. Sales, & W. H. Reid (Eds.), *Treating adult and juvenile offenders with special needs* (pp. 249–271). Washington, DC: American Psychological Association.
- Shearer, R. A., Myers, L. B., & Ogan, G. D. (2001). Treatment resistance and ethnicity among female offenders in substance abuse treatment programs. *Prison Journal*, *81*, 55–72.
- Sheidow, A. J., & Henggeler, S. W. (2005). Community-based treatments. In K. Heilbrun, N. E. S. Goldstein, & R. E. Redding (Eds.), *Juvenile delinquency: Prevention, assessment, and intervention* (pp. 257–281). New York, NY: Oxford University Press.
- Skeem, J. L., Manchak, S., & Peterson, J. K. (2011). Correctional policy for offenders with mental illness: Creating a new paradigm for recidivism reduction. *Law and Human Behavior*, *35*, 110–126.
- Skogstad, P., Deane, F. P., & Spicer, J. (2006). Social-cognitive determinants of help-seeking for mental health problems among prison inmates. *Criminal Behaviour and Mental Health*, *16*, 43–59.
- Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology*, *84*(5), 1041–1053. doi:10.1037/0022-3514.84.5.1041
- Steadman, H. J., Holohean, E. J., Jr., & Dvoskin, J. (1991). Estimating mental health needs and service utilization among prison inmates. *Bulletin of the American Academy of Psychiatry and the Law*, *19*, 297–307.
- Steadman, H. J., McCarty, D. W., & Morrissey, J. P. (1989). *The mentally ill in jail: Planning for essential services*. New York, NY: Guilford Press.
- Steadman, H. J., & Veysey, B. M. (2007). *Research in brief: National Institute of Justice: Providing services for jail inmates with mental disorders*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, National Institute of Justice. Retrieved from <https://www.ncjrs.gov/pdffiles/162207.pdf>
- Steffan, J. S., & Morgan, R. D. (2005). Meeting the needs of mentally ill offenders: Inmate service utilization. *Corrections Today*, *67*, 38–41, 53.
- Tate, D. C., & Redding, R. E. (2005). Mental health and rehabilitative services in juvenile justice: System reforms and innovative approaches. In K. Heilbrun, N. E. S. Goldstein,

- & R. E. Redding (Eds.), *Juvenile delinquency: Prevention, assessment, and intervention* (pp. 134–160). New York, NY: Oxford University Press.
- Taylor, J. (2010). Psychotherapy for people with learning disabilities: Creating possibilities and opportunities. A review of the literature. *Journal of Learning Disabilities and Offending Behaviour*, 1(3), 15–25. doi:10.5042/jldob.2010.0625
- Taylor, J. L., Lindsay, W. R., & Willner, P. (2008). CBT for people with intellectual disabilities: Emerging evidence, cognitive ability and IQ effects. *Behavioural and Cognitive Psychotherapy*, 36, 723–733. doi:10.1017/S1352465808004906
- Teplin, L. A. (1984). Criminalizing mental disorder: The comparative arrest rate of the mentally ill. *American Psychologist*, 39, 794–803.
- Teplin, L. A., Abram, K. M., McClelland, G. M., Mericle, A. A., Dulcan, M. K., & Washburn, J. J. (2006). *Psychiatric disorders of youth in detention* (NCJ 210331). Office of Juvenile Justice and Delinquency Prevention Bulletin. Retrieved from <https://www.ncjrs.gov/pdffiles1/ojdp/210331.pdf>
- Tong, L.S.J., & Farrington, D. P. (2006). How effective is the “Reasoning and Rehabilitation” programme in reducing reoffending? A meta-analysis of evaluations in four countries. *Psychology, Crime & Law*, 12(1), 3–24. doi:10.1080/10683160512331316253
- Tong, L.S.J., & Farrington, D. P. (2008). Effectiveness of “Reasoning and Rehabilitation” in reducing reoffending. *Psicothema*, 20(1), 20–28.
- Valliant, P. M., & Antonowicz, D. H. (1992). Rapists, incest offenders, and child molesters in treatment: Cognitive and social skills training. *International Journal of Offender Therapy and Comparative Criminology*, 36, 221–230.
- van Hiel, A., Hautman, L., Cornelis, I., & de Clercq, B. (2007). Football hooliganism: Comparing self-awareness and social identity theory explanations. *Journal of Community & Applied Social Psychology*, 17, 169–186.
- Walters, G. D. (2002). Current and historical content scales for the Psychological Inventory of Criminal Thinking Styles (PICTS). *Legal and Criminological Psychology*, 7, 73–86.
- Walters, G. D. (2004). Predictors of early termination in a prison-based program of psychoeducation. *Prison Journal*, 84, 171–183. doi:10.1177/0032885504265076
- Walters, G. D. (2006). *The Psychological Inventory of Criminal Thinking Styles (PICTS) professional manual*. Allentown, PA: Center for Lifestyle Studies.
- Walters, G. D. (2009). The Psychological Inventory of Criminal Thinking Styles and Psychopathy Checklist: Screening Version as incrementally valid predictors of recidivism. *Law and Human Behavior*, 33, 497–505.
- Walters, G. D., & Mandell, W. (2007). Incremental validity of the Psychological Inventory of Criminal Thinking Styles and Psychopathy Checklist: Screening Version in predicting disciplinary outcome. *Law and Human Behavior*, 31, 141–157.
- Walters, G. D., & Schlauch, C. (2008). The Psychological Inventory of Criminal Thinking Styles and Level of Service Inventory-Revised: Screening Version as predictors of official and self-reported disciplinary infractions. *Law and Human Behavior*, 32, 454–462.
- Walters, G. D., Trgovac, M., Rychlec, M., Di Fazio, R., & Olson, J. R. (2002). Assessing change with the Psychological Inventory of Criminal Thinking Styles: A controlled analysis and multisite cross-validation. *Criminal Justice and Behavior*, 29, 308–331.
- Wang, E. W., Owens, R. M., Long, S. A., Diamond, P. M., & Smith, J. L. (2000). The effectiveness of rehabilitation with persistently violent male prisoners. *International Journal of Offender Therapy and Comparative Criminology*, 44, 505–514.

- Westhuis, D. J., Gwaltney, L., & Hayaski, R. (2001). Outpatient cocaine abuse treatment: Predictors of success. *Journal of Drug Education, 31*, 171–183.
- Wexler, H. K., Falkin, G. P., & Lipton, D. S. (1990). Outcome evaluation of a prison therapeutic community for substance abuse treatment. *Criminal Justice and Behavior, 17*, 71–92.
- Wexler, H. K., De Leoan, G., Thomas, G., Kressle, D., & Peters, J. (1999). The Amity Prison TC Evaluation. *Criminal Justice and Behavior, 26*, 147–167. doi:10.1177/0093854899026002001
- Wierzbicki, M., & Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. *Professional Psychology: Research and Practice, 24*, 190–195.
- Williams, M. W., Skogstad, P., & Deane, F. P. (2001). Attitudes of male prisoners toward seeking professional psychological help. *Journal of Offender Rehabilitation, 34*(2), 49–61.
- Williamson, P., Day, A., Howells, K., Bubner, S., & Jauncey, S. (2003). Assessing offender readiness to change problems with anger. *Psychology, Crime & Law, 9*, 295–307.
- Wise, E. A. (2004). Methods for analyzing psychotherapy outcomes: A review of clinical significance, reliable change, and recommendations for future directions. *Journal of Personality Assessment, 82*(1), 50–59. doi:10.1207/s15327752jpa8201_10
- Witte, T. D., Di Placido, C., Gu, D., & Wong, S. C. P. (2006). An investigation of the validity and reliability of the Criminal Sentiments Scale in a sample of treated sex offenders. *Sex Abuse, 18*, 249–258.
- Wolff, N., Morgan, R. D., Shi, J., Fisher, W., & Huening, J. (2011). Comparative analysis of thinking styles and emotional states of male and female inmates with and without mental disorders. *Psychiatric Services, 62*, 1485–1493.
- Wong, S., & Gordon, A. (2012). *The violence reduction program: A treatment program for violence-prone forensic clients*. Unpublished treatment program.
- Wong, S. C. P., & Gordon, A. (2006). The validity and reliability of the Violence Risk Scale: A treatment friendly violence risk assessment tool. *Psychology, Public Policy, and Law, 12*, 279–309.
- Wong, S. C. P., Gordon, A., & Gu, D. (2007). Assessment and treatment of violence-prone forensic clients: An integrated approach. *British Journal of Psychiatry, 190*(49), S66–S74. doi:10.1192/bjp.190.5.s66
- Wong, S. C. P., Olver, M. E., & Stockdale, K. C. (2009). The utility of dynamic and static factors in risk assessment, prediction, and treatment. In J. T. Andrade (Ed.), *Handbook of violence risk assessment and treatment: New approaches for mental health professionals* (pp. 83–120). New York, NY: Springer.
- Wong, S. C. P., Van der Veen, S., Leis, T., Denkhaus, H., Gu, D., Liber, E., & Middleton, H. (2005). Reintegrating seriously violent and personality disordered offenders from a super-maximum security institution into the general offender population. *International Journal of Offender Therapy and Comparative Criminology, 49*(4), 362–375.
- Worling, J., & Curwen, T. (2000). Adolescent sexual offender recidivism: Success of specialized treatment and implications for risk prediction. *Child Abuse & Neglect, 24*, 965–982.
- Wormith, J. S., & Olver, M. E. (2002). Offender treatment attrition and its relationship with risk, responsivity, and recidivism. *Criminal Justice and Behavior, 29*, 447–471.
- Wormith, J. S., Althouse, R.L.R., Simpson, M., Fagan, T. J., & Morgan, R. D. (2007). The rehabilitation and reintegration of offenders: The current landscape and some future directions for correctional psychology [Special issue]. *Criminal Justice and Behavior, 34*, 879–892.

- Zamble, E., & Quinsey, V. L. (1997). *The criminal recidivism process*. Cambridge, England: Cambridge University Press.
- Zanis, D. A., Mulvaney, F., Coviello, D., Alterman, A. I., Savitz, B., & Thompson, W. (2003). The effectiveness of early parole to substance abuse treatment facilities on 24-month criminal recidivism. *Journal of Drug Issues*, 33, 223–236.
- Zaylor, C., Nelson, E., & Cook, D. J. (2001). Clinical outcomes in a prison telepsychiatry clinic. *Journal of Telemedicine and Telecare*, 7, 47–49.

Assessing and Treating Sex Offenders

W. L. MARSHALL, DOUGLAS BOER, AND LIAM E. MARSHALL

SEX offending is a serious social problem affecting the lives of many innocent men, women, and children (Calhoun & Wilson, 2000; Paolucci, Genuis, & Vidato, 2001). Although sex offending has a long history (Licht, 1932; Taylor, 1954), only relatively recently has a concerted effort been made to appropriately assess and treat these offenders (Laws & Marshall, 2003; W. L. Marshall & Laws, 2003). Depending on the source of the data, up to 50% of adult women will report having been sexually abused on at least one occasion (Di Vasto et al., 1984), and this is true across all countries (van Dijk & Mayhew, 1992). A similar number of children also appear to have been abused (Briere & Elliot, 2003), and adult males are also victimized. Even when quite restrictive definitions of abuse are employed, the number of people victimized by sex offenders remains alarmingly high, and many appear to suffer immediate and long-term consequences (Browne & Finkelhor, 1986; Resick, 1993). Although it appears that the majority of the offenders are male, there are also reports of female sex offenders (Ford, 2009). Because our work has been almost exclusively with male offenders, we will, throughout this chapter, refer to the offenders as males.

Sex offending is, then, an issue calling for a systematic and comprehensive response. Of course, such a response is most effective when it is based on carefully collected data. This chapter addresses some of the pertinent available evidence. The focus is on what is known about the features of sex offenders, what has been established concerning their threat to reoffend, and the methods that have been developed to reduce that threat.

Because diagnosis and assessment of these offenders is the first step in dealing with them, we take up that issue first. We then discuss the ways in which risk has been assessed, and finally we provide an overview of treatment, focusing in particular on our own treatment program. Although theories of sexual offending have proven valuable, we do not address them in this chapter. The interested reader

is referred to two excellent books on such theories (Stinson, Sales, & Becker, 2008; Ward, Polaschek, & Beech, 2006).

DIAGNOSTIC AND ASSESSMENT ISSUES

The first, and most fundamentally essential, step in considering the value of a diagnosis is to establish its reliability (Nelson-Gray, 1991). In the present context, this means agreement between two or more diagnosticians on the diagnosis to be applied to an individual. Without establishing reliability, questions concerning the validity of a diagnosis are moot. Studies examining the reliability of the various paraphilias have not provided satisfactory evidence of consistency across diagnosticians. The best studies to date have examined the diagnostic processes involved in evaluating sex offenders referred for civil commitment under sexually violent predator laws. In these processes, the diagnostic evaluations are required to meet criteria established by the state, which then qualifies these evaluators as experts in conducting these assessments. Two studies (Levenson, 2004; Perillo, Mercado, & Jeglic, 2011) found far less than satisfactory agreement between assessors, and this was true for pedophilia, sexual sadism, and paraphilia not otherwise specified (NOS). In a later report by Packard and Levenson (2006), which employed alternative analyses of the data in Levenson's (2004) earlier study, the results were better but nevertheless fell short of the established criteria for an important decision (Cicchetti & Sparrow, 1981; Cohen, 1969).

CHILD MOLESTERS

The *Diagnostic and Statistical Manual (DSM)* of the American Psychiatric Association first used the term *pedophilia* in its third edition (*DSM-III*; American Psychiatric Association [APA], 1980) to describe a specific subset of child molesters. Until the publication of *DSM-IV* (APA, 1994), the manual defined pedophilia in a way that excluded a substantial number of child molesters. This was not an unreasonable position, given that many men who molest children do not appear to have fixated or preferential sexual interests in children (W. L. Marshall & Fernandez, 2003). Unfortunately, however, many clinicians and researchers used the term *pedophilia* more generically to include all child molesters, thereby causing considerable confusion. Despite this inconsistent use of terms to describe these offenders, it appears that clinicians agree that all child molesters require treatment, regardless of whether they meet diagnostic criteria for a mental disorder. In the settings where we work (i.e., federal prisons, secure psychiatric settings, institutional and community programs for adult and juvenile offenders), administrators demand that we treat all sex offenders.

DSM III-R (APA, 1987, p. 285) defined pedophilia as involving "recurrent intense sexual urges and sexually arousing fantasies involving sexual activity with a prepubescent child or children." Because these criteria did not include actually

engaging in sexual activities with a child, a *DSM-III-R* diagnosis of pedophilia could be applied to persons who had never committed an offense, which seems reasonable if these people are plagued by distress over their deviant sexual interests. However, the diagnosis of pedophilia could not be applied to those persons who had molested a child but who were *not* plagued by deviant urges and fantasies. An examination of our extensive clinical files revealed no clear evidence of *recurrent* urges or fantasies in almost 60% of nonfamilial child molesters or in over 75% of incest offenders (W. L. Marshall, 1998). Apparently, recurrent urges and fantasies are not evident in all child molesters. Obviously, then, a *DSM-III-R* diagnosis of pedophilia did not have differential implications for treatment, given that clinicians have been treating all child molesters in the same way, regardless of their diagnostic status.

At the time of writing the most recent available version of *DSM*, *DSM-IV-TR* (APA, 2000) modified the criteria for pedophilia. The criteria now read "recurrent, intense sexually arousing fantasies, sexual urges, or behaviors involving sexual activity with a prepubescent child or children" (*DSM-IV-TR*, 2000, p. 572, emphasis added). The addition of "behaviors" along with the amendment to criterion B to read "the person has acted on these sexual urges, or the sexual urges or fantasies cause marked distress or personal difficulty" (p. 572), clearly suggest that all child molesters are now to be considered pedophiles. Whether clinicians will change their diagnostic practices in this way remains to be seen, but so far most still seem to distinguish pedophiles from other child molesters.

Of course, it may be that pedophiles (as a distinct subgroup of child molesters) are more or less responsive to treatment than are other child molesters. A detailed examination of our records on treated child molesters revealed no differences in the reoffense rates for those who could be classified as pedophiles versus those who could not (W. L. Marshall, 2008). Others have likewise found that diagnosing someone as a pedophile appears to have no implications for future risk (Kingston, Firestone, Moulden, & Bradford, 2007; Moulden, Firestone, Kingston, & Bradford, 2009; Wilson, Abracen, Looman, Picheca, & Ferguson, 2011). Perhaps it is better to use the term *child molesters* rather than *pedophiles*, since it appears to be the view of the majority of clinicians that all child molesters need treatment. We share this view partly because many child molesters who are deemed not to be pedophiles have earnestly sought treatment as they are disturbed by their behavior and interests. Also many of these offenders have significant deficits in areas of functioning that predict reoffending.

The addition of "behaviors" to the diagnostic criteria of *DSM-IV* and *DSM-IV-TR* represents a sensible change and allows for all child molesters (at least those who molest prepubescent children) to be diagnosed as pedophiles. If the application of this diagnosis allows the offenders to access financial support for treatment, then this change in diagnostic criteria serves a useful purpose. However, this and other changes over the years have caused serious problems in integrating research. In addition to the habit that some researchers have of using the term *pedophilia* as a

generic descriptor, these changes make it hard to compare different studies over different time periods.

There are two particular problems facing the clinician who wishes to rely on *DSM-IV-TR* to diagnose child molesters. According to *DSM-IV-TR*, pedophilia can be diagnosed if “the sexual urges or fantasies cause marked distress or interpersonal difficulty” (APA, 2000, p. 572). Presumably this is to allow the diagnosis to be applied to someone who complains of having unwanted sexual fantasies about children but has not yet acted on them. However, this would seem to exclude men with many victims over many years who are quite content with their deviant sexual interests. The criterion that specifies the child must be prepubescent, with the indication that this typically means under age 13 years, seems arbitrary and may be difficult to determine. In many cases, clinicians have no information independent of the offender as to the victim’s age, and child molesters characteristically report the child to be older at the time of the offending than the victim claims. Even with available official information, when the victim reports the offending some time after it commenced, clinicians are often faced with discrepant claims from the offender and the victim about the child’s age at the onset of abuse. More detailed and thorough criticisms of *DSM* criteria for pedophilia are provided by W. L. Marshall (1997, 2007) and by O’Donohue, Regev, and Hagstrom (2000).

RAPISTS

The *DSM* does not deem rapists (or, for that matter, child molesters who offend against postpubescent children) to have a diagnosable problem. Just why it is that a homosexual who feels uncomfortable about his or her sexual orientation should be said to have a disorder, whereas a man who repeatedly rapes women does not, is difficult to understand. The only rapists who would meet diagnostic criteria for a disorder according to *DSM-IV-TR* are those who are sexual sadists, although clinicians completing assessments for civil commitment purposes characteristically diagnose rapists as having a “paraphilia NOS” (see Doren, 2002, for a discussion). This latter decision seems arbitrary at best and contradicts the decisions of the *DSM* committee. Nevertheless, at all international conferences we have attended over the past 45 years, clinicians have clearly indicated that they believe that all rapists are in need of treatment.

In a series of articles, we have examined the meaning, application, and reliability of the diagnosis of sexual sadism. Our review of the literature (W. L. Marshall & Kennedy, 2003) revealed remarkably variable criteria employed by researchers and quite variable evidence used to determine whether an individual met these criteria. When we examined the information available to experienced diagnosticians applying *DSM* criteria within a prison setting, we found serious problems. Those offenders who were diagnosed as sexually sadistic had engaged in less brutal, less cruel, and less torturous acts than was the case for those to whom the diagnosis was not applied (W. L. Marshall, Kennedy, & Yates, 2002). We then

asked 15 internationally renowned experts on sadism to identify sexual sadists from detailed accounts (including life history, offense history, offense details, psychological and phallometric test results, and offenders' self-reports) extracted from the information made available to the diagnosticians in the earlier study. The classifications made by these experts (is or is not a sexual sadist) appeared almost random (W. L. Marshall, Kennedy, Yates, & Serran, 2002). Using the kappa statistic as an index, considerable disagreement was evident ($\kappa = .14$). These three reports clearly suggest that sexual sadism is, at best, a diagnosis in desperate need of clarification. We (W. L. Marshall & Hucker, 2006; Nitschke, Mokros, Osterheider, & Marshall, in press) have attempted to improve the diagnostic reliability for sadism by producing a scale that allows for both a categorical and dimensional estimate of sadism.

EXHIBITIONISTS

The *DSM* is quite clear that to meet criteria for exhibitionism, a person (typically a male) must have "recurrent, intense sexually arousing fantasies, sexual urges, or behaviors involving the exposure of one's genitals to an unsuspecting stranger" (APA, 2000, p. 569). Because all exhibitionists, by definition, engage in behaviors involving the exposure of their genitals to unsuspecting strangers, they all appear to meet the diagnostic criteria of a paraphilia. However, few exhibitionists report fantasies involving exposure (W. L. Marshall, Payne, Barbaree, & Eccles, 1991).

Clinicians have always considered such offenders to need treatment, particularly because this behavior is persistent. Marshall, Eccles, and Barbaree (1991), for example, reported that 57% of their untreated exhibitionists reoffended within 4 years of initial identification, and Maletzky (1991) found that his 770 exhibitionists averaged almost three exposures per week over an average period of 7.5 years. Unlike the problems raised in this chapter about pedophilia and rape, the diagnosis of exhibitionism appears uncomplicated. With exhibitionism, then, diagnosis leads to treatment, since all such offenders meet diagnostic criteria and all are deemed to need treatment.

ALTERNATIVES TO *DSM* CRITERIA

DSM criteria aside, most clinicians look to more objective methods of evaluating whether a child molester is a pedophile. Typically they search for indications that the offender has a persistent sexual interest or preference for children. Most child molesters are understandably reluctant to admit having persistent deviant fantasies even when they admit to having committed an offense. Faced with this fact, clinicians and researchers have employed phallometry. Phallometry involves the measurement (metric) of changes in the penis (phallus) in response to viewing or listening to depictions of various sexual stimuli. In the case of child molesters, this involves sexualized depictions of children or descriptions of sex between an

adult and a child; for rapists, these depictions portray nonconsenting forced sex with an adult. For sadists, the deviant themes describe a person committing severe sexual or nonsexual violence against an adult or a child; for exhibitionists, the scenes would depict acts of genital exposure to an unwilling person. In all these cases, the arousal displayed in response to the deviant scenes is compared with the arousal evoked by adult consenting sex. It is assumed that, if a man displays arousal to children that is equal to or greater than his responses to adult consenting sex, he is said to have deviant urges or fantasies. In the view of some authors (Freund & Blanchard, 1989; Freund & Watson, 1991), such a profile means that the man meets criteria for pedophilia.

The results of phallometric assessments are considered relevant not only to diagnoses but also to the identification of treatment targets and risk assessment (Seto, 2008). The evidence bearing on the value of these uses of phallometrics has been reviewed by numerous authors, with most coming to positive conclusions (Abel & Blanchard, 1976; Freund, 1981; W. D. Murphy & Barbaree, 1994; O'Donohue & Letourneau, 1992; Rosen & Beck, 1988). However, in our comprehensive review (W. L. Marshall & Fernandez, 2003), we were unable to find convincing evidence that the procedures were reliable, which is an essential first step in establishing the utility of any measure. Returning to the W. L. Marshall and Fernandez review, authorities on test development point out that one of the first steps to ensure the validity of the scores generated by a test is to establish reliability. K. R. Murphy and Davidshofer (1998) noted that "lack of reliability places a limit on the validity of the inferences drawn from test scores" and an "unreliable test . . . cannot possibly be valid" (p. 129). Nelson-Gray (1991) applied these principles to diagnostic practices and came to essentially the same conclusion, noting that, unless different clinicians can consistently agree on the application of a set of criteria, the diagnosis will be of no use in managing and treating clients. We would add that if a diagnosis is not reliable, different researchers applying the same diagnosis will likely generate inconsistent findings. We found that among child molesters, for example, only those who admitted to their problems and had multiple victims appeared deviant on phallometric assessments. Child molesters who were in denial or who had only a single victim displayed normative sexual preferences (Freund, Chan, & Coulthard, 1979; Freund & Watson, 1991), and yet these are the examinees for whom diagnostic issues are the most problematic. Of course, these offenders may have been deliberately controlling their responses or their responses may reflect their true dispositions; we simply do not know. Diagnosticians certainly appear to be reluctant to label a child molester as a pedophile unless he has multiple victims (W. L. Marshall, 1997; Seto, 2008).

Phallometric assessments are also employed in the assessments of rapists, sadists, and exhibitionists, at least in part to determine diagnostic status. However, the data on phallometric evaluations of rapists has led to conflicting interpretations. Lalumière and Quinsey (1994), for example, claim that, when done properly, such evaluations distinguish rapists from other men and thereby allow conclusions

about the degree of deviance. From our comprehensive reviews of the literature, we (W. L. Marshall & Fernandez, 2000a, 2003) concluded that such assessments do not reliably identify rapists and that, where differences occur between these offenders and others, the differences in the percentages in each group who appear deviant is quite small. Unfortunately, there are few studies evaluating sadists using these methods, but in these reports, it appears that some sadists do respond to sexualized violence (Barbaree, 1990). The literature on phallometric responses among exhibitionists is confusing, but there is no evidence that these men are aroused by depictions of exposing (W. L. Marshall et al., 1991). Phallometry, then, has not served as an adequate alternative to the *DSM* for diagnostic purposes. However, it is clear that, if an identified sexual offender displays deviant arousal at testing, then he has a problem that needs to be addressed in treatment. Such an index of deviance for those who molest male victims predicts reoffending (Hanson & Bussière, 1998).

In recent years, alternatives to phallometry have been developed that purport to identify deviance among sex offenders (Abel, Huffman, Warberg, & Holland, 1998; Thornton & Laws, 2009). Although there are promising results with these more cognitively based assessments, at present the data are not sufficiently convincing to recommend their widespread adoption. Similarly, the use of large batteries of self-report measures have not produced results that always distinguish sex offenders from others, presumably in part because sex offenders have a vested interest in portraying themselves positively, and in these tests the prosocial answers are obvious. Indeed, we have suggested that such assessments have no value for case formulations that might facilitate differential treatment (W. L. Marshall, Marshall, Serran, & O'Brien, 2011). Finally, the scientific bases of polygraphy have been seriously called into question for any purpose (Iacono & Patrick, 1999; National Research Council, 2003), despite its widespread popularity in sex offender programs.

ASSESSMENTS

As noted earlier, most clinicians working with sex offenders consider them all to be in need of treatment, regardless of whether they meet diagnostic criteria for a disorder. Given the remarkably damaging effects of these offending behaviors and the limitations of the *DSM*, this is a sensible strategy. In their work with sex offenders, clinicians are faced with a number of issues they either must address or are asked to address. To meet these demands, clinicians must come to some conclusion regarding the nature of the problem the offender presents, and to do this, typically a thorough evaluation is completed.

It is not possible in this chapter to cover all the questions asked of those who deal with sex offenders, but one issue can be dismissed right away. A clinician cannot offer assistance in the determination of the guilt or innocence of an alleged sex offender. Several reviews have identified problems that arise when such an appraisal

is attempted (Barbaree & Peacock, 1995; W. L. Marshall, 1996; W. L. Marshall & Fernandez, 2000b; Peters & Murphy, 1992; Simon & Schouten, 1992). These reviews concluded that the empirical bases of the procedures employed for these purposes are inadequately suited to the onerous task of determining whether an accused did or did not commit an offense. Responsible clinicians will, therefore, refuse to conduct appraisals aimed at determining culpability of accused sex offenders.

The three most important reasons to conduct a thorough assessment of sex offenders are to (1) determine risk to reoffend, (2) identify treatment needs, and (3) evaluate whether treatment has produced the desired changes. We consider each of these issues.

RISK ASSESSMENT

In his book *Predicting Violent Behavior: An Assessment of Clinical Techniques*, Monahan (1981) noted that (at that time) the “‘best’ clinical research currently in existence indicates that psychiatrists and psychologists are accurate in no more than one out of three predictions of violent behavior over a several-year period among institutionalized populations that had both committed violence in the past . . . and who were diagnosed as mentally ill” (p. 77). Nonetheless, in the same book, Monahan opined, “There may be circumstances in which prediction is both empirically possible and ethically appropriate” (p. 19). At the time of Monahan’s publication, the majority of risk assessments of violent and sexually violent offenders were completed using unstructured clinical judgment. Given that Meehl (1986) came to the conclusion that unstructured clinical judgment is significantly less accurate than an actuarial approach, we will set aside any further consideration of this approach. Although these authors were writing about the assessment of risk regarding violent offenders, usually mentally ill violent offenders, it has become apparent that the same principles apply to the evaluation of risk with sex offenders.

The field of actuarial risk assessment with violent offenders and sex offenders had its roots in the work of the Penetanguishene research group. This group first described its “Violence Prediction Scheme” in 1994 (Webster, Harris, Rice, Cormier, & Quinsey, 1994). This book outlined the “Risk Assessment Guide,” which later came to be described as the Violence Risk Appraisal Guide (VRAG), a measure that was developed on a mixed group of mentally disordered offenders. Later, Quinsey, Harris, Rice, and Cormier (1998) published another book describing the VRAG and a variation designed specifically for sex offenders: *Sex Offender Risk Appraisal Guide (SORAG)*.

Shortly after the advent of the VRAG, Hanson (1997) developed a brief actuarial scale for assessing sex offender recidivism, known as the Rapid Risk Assessment for Sex Offender Recidivism (RRASOR). At around the same time Thornton was developing the Structured Anchored Clinical Judgement scale (SACJ; see Grubin, 1998, for a description). Shortly thereafter, Hanson and Thornton merged their two

scales to form the Static-99 (Hanson & Thornton, 1999), which has become the most popular risk scale for sex offenders.

The Static-99 has routinely been shown to have moderate to good predictive accuracy, and the VRAG/SORAG similarly generate moderate predictive accuracy for violent and sexually violent behaviors (Hanson & Morton-Bourgon, 2004, 2007, 2009). Interestingly, the initial meta-analysis by Hanson and Morton-Bourgon (2004) found that the SORAG had poor discriminant validity, in that it had better predictive validity for nonsexual recidivism than for sexual recidivism, while the opposite was true for the Static-99. Research by Hanson and others has resulted in the Static-99R (www.static99.org), which better accounts for the effect of age (basically risk goes down as age goes up). There are more recently developed actuarial measures that show promise, including the Multisample Age-stratified Table of Sexual Recidivism Rates-1 (MATS-1) (Wollert, Cramer, Waggoner, Skelton, & Vess, 2010) and the Minnesota Sex Offender Screening Test-3 (MnSOST-3) (Duwe & Freske, 2012).

Actuarial tests used to estimate risk for sexual offending are not without their critics. A study of group and individual risk levels found that the confidence intervals of the Static-99 and the VRAG at the individual level were so large that the risk estimates were essentially meaningless (Hart, Michie, & Cooke, 2007). Blair, Marcus, and Boccaccini (2008) found that the effect sizes of the VRAG, SORAG, and Static-99 were significantly larger in studies conducted by the originators of these tests than in studies conducted by independent researchers.

These risk assessment instruments, and all those that follow this tradition, have assessed reoffense data on many subjects over many years and examined the relationship between reoffending and a wide variety of potential predictors. Following the example of actuarial tables employed by insurance companies, these risk assessment instruments categorize offenders in groups of various risk levels on the basis of the features that were found to predict risk. Like insurance actuarial tables, however, the features of a particular risk group do not mean that any individual sex offender is at this risk level. Some within each risk group will reoffend while some will not, so the problem for clinicians is to make inferences from these group estimates about the risk of *specific* individuals. This concern has led to the recommendation that clinical information (Mann, Hanson, & Thornton, 2011) be used to adjust actuarial estimates, a practice known as anchoring, which is not considered an appropriate practice by some (Quinsey et al., 1998).

Structured professional judgments (SPJs) for sexual offender risk assessment represent attempts to deal with this problem. These measures include the Sexual Violence Risk-20 (SVR-20; Boer, Hart, Kropp, & Webster, 1997), and the Risk for Sexual Violence Protocol (RSVP; Hart et al., 2003). SPJs for sexual offenders are composed of risk factors for sexually violent recidivism derived from the clinical and research literature. These risk factors have varying levels of predictive validity on their own. While Hanson and Morton-Bourgon (2004) found the SVR-20 had

better predictive validity than the Static-99, this was not the case in a subsequent meta-analysis (e.g., Hanson & Morton-Bourgon, 2007).

In addition, Storey, Watt, Jackson, and Hart (2012) found that the use of a clinical override or adjustment to the Static-99 ratings showed poorer predictive validity than the Static-99 ratings on their own. Thus, anchoring or adjusting actuarial risk estimates does not seem to be a good way to incorporate clinical risk issues into a risk analysis. Tests that utilize both SPJ items and actuarial items in an additive fashion would leave such instruments open to the criticism of anchoring.

The question of how best to utilize the risk assessment literature has been considered by Boer (2006) in a proposal for a “convergent” approach to risk assessment. Boer proposed not merging actuarial and SPJ findings or anchoring one to the other but utilizing the best tests for the case at hand. This approach provides information for decision makers about the baseline risk (the actuarial group-based risk level) and which individual risk issues need to be managed in each individual case. Boer recommended that the actuarial data and the SVR-20 information be reported independently and discussed in combination for the purposes of risk management. In this sense the actuarial data serve to establish a risk baseline with the SVR-20 guiding the selection of targets to address in treatment. When actuarial risk is high, there should be a greater urgency attached to monitoring and managing the SPJ risk factors. At the present time, it would seem that the separate reporting of risk assessment test findings and the subsequent integration of these findings in the risk management section of risk reports would be the best practice to help maximize public safety.

ASSESSING TREATMENT NEEDS AND TREATMENT-INDUCED CHANGES

The targets in these assessments should, of course, match the targets addressed in treatment. In early programs, the targets addressed in treatment either relied on the particular clinician’s choices or were derived from evidence indicating the differences between sex offenders and other males. Fortunately in recent years, following the observations of Andrews and his colleagues (Andrews, Bonta, & Hoge, 1990; Andrews et al., 1990; Dowden & Andrews, 2004) that treatment for offenders must address those factors that predict reoffending (i.e., so-called criminogenic factors), Hanson and colleagues (Hanson & Bussière, 1998; Hanson & Morton-Bourgon, 2005) have applied meta-analyses to a large number of studies of sex offenders. From these studies, they (see Hanson, 2006; Mann et al., 2011, for summaries) have identified a range of characteristics that are potentially modifiable and that they have called either “stable” dynamic factors or “acute” dynamic factors. Stable factors are those chronically present problems that can be effectively changed in treatment, whereas acute factors are more relevant to community supervision after discharge. It is these stable or criminogenic factors that must be addressed in treatment. Table 25.1 describes these treatment targets.

Table 25.1
Criminogenic Targets

Cognitions
Functional congruence with children
Hostility toward women
Lack of concern for others
Offense supportive attitudes
Relationship problems
Lack of intimacy
Insecure attachments
Emotional loneliness
Self-regulation deficits
Emotional and behavioral dysregulation
Poor coping skills
Sexual factors
Sexual preoccupation
Sexual violence
Sexual interest in children

Source: Adapted from Mann et al. (2011).

However, although it is not advisable to address too many noncriminogenic features, as this can reduce treatment effectiveness (Gendreau, French, & Gionet, 2004), there are sound reasons for modifying some of these features (Dowden & Andrews, 2003; W. L. Marshall & Marshall, 2012). For example, low self-esteem blocks effective engagement in treatment (Baumeister, 1993), and the majority of sex offenders suffer from this problem (W. L. Marshall, Anderson, & Champagne, 1997). As a consequence, this issue must be addressed in the early stage of treatment if effective progress is to be achieved.

Assessments characteristically target each of the criminogenic areas, although other rather obvious features are also appraised, such as whether the offenders have another serious disorder (e.g., brain damage, psychosis, or depression) that may affect their ability to participate in treatment, as might the offender's intellectual ability and educational attainment. Several test procedures are available that can be used to evaluate these various issues (Mussack & Carich, 2001), although additional information derived from interviews is typically deemed essential.

Cognitions. Measures of distorted cognitions that justify or excuse sexual offending have been described by Abel et al. (1989) and by Bumby (1996). However, most of the items in these measures address noncriminogenic issues. It is interesting that most sex offender programs continue to address an array of cognitions that are not in fact criminogenic (McGrath, Cumming, Burchard, Zeoli, & Ellerby, 2010).

For example, almost all programs attempt to have offenders take full responsibility for their offenses and try to eliminate all attempts at excuse making. Not only are excuses not criminogenic, they are in fact both healthy (Schlenker, Pontari, & Christopher, 2001) and predictive of desisting from future offending (Maruna, 2001). We (W. L. Marshall, Marshall, & Kingston, 2011; W. L. Marshall, Marshall, & Ware, 2009) have shown that the majority of so-called cognitive distortions addressed in treatment programs for sex offenders are not criminogenic and therefore should not be treatment targets. However, as shown in Table 25.1, some cognitions are problematic.

Howells (1979) described the tendency of child molesters to feel emotionally at ease with children and uncomfortable with adults, but he did not employ standardized tests. At present, there is no standard measure of emotional congruence with children. Hostility toward women is typically assessed by Burt's (1980) scales, and measures of empathy (e.g., Serran, 2002) serve to evaluate a lack of concern for others. Empathy, however, is more complex than just a lack of concern for others (L. E. Marshall & Marshall, 2011); it includes aspects of the ability to read other people's thoughts and feelings that have been identified as deficits in sexual offenders' "theory of mind" (Castellino, Bosco, Marshall, Marshall, & Veglia, 2011).

A Justifications Scale was developed by W. L. Marshall (1991) to evaluate offense supportive attitudes among child molesters. This scale was later appraised by Mann, Webster, Wakeling, and Marshall (2007) using large samples of sex offenders and nonoffenders. They demonstrated the scale's reliability and criterion validity showing that it distinguished child molesters from the others.

Relationship Problems. Although concerns about sex offenders' capacity for intimacy have been noted by several authors, it was not formally addressed until W. L. Marshall (1989) suggested its role in the etiology of sexual offending. The suggestion in Marshall's paper was that a lack of intimacy might be due to deficient relationship skills, which might in turn encourage the person to seek sexual contacts under circumstances that do not demand these skills (e.g., by raping a woman or molesting a child). Subsequent research has confirmed that sex offenders do indeed lack intimacy and that they develop insecure attachment styles as a result (W. L. Marshall & Marshall, 2010). Lack of satisfactory intimate relationships is one of the strongest predictors of reoffending among sex offenders (Hanson & Morton-Bourgon, 2005). Fortunately, good measures of intimacy (Social Intimacy Scale, R. S. Miller & Lefcourt, 1982), loneliness (Revised UCLA Loneliness Scale, Russell, Peplau, & Cutrona, 1980), and attachment (see Rich, 2006, for a description of various measures) are available.

Self-Regulation. The term *self-regulation* refers to a variety of processes that humans employ, or fail to employ, to moderate the expression of their behavior and internal states (Vohs & Baumeister, 2004). Thus, a person who displays sound

self-regulation would not typically behave impulsively or display behavior or emotions that are out of keeping with social standards. While impulsivity, or submitting to momentary temptations, might be seen as strictly a failure to exercise behavioral self-control, in fact, emotional regulation appears to be the basis for all forms of regulatory control (Gross & Thompson, 2007). In this understanding, emotions are seen as multifaceted phenomena involving changes in subjective experience and behavior as well as control over peripheral physiology (Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005). Indeed, there is “a natural link between affect and action... [such that] the affect loop has a direct influence on what occurs in the action loop” (Carver, 2004, p. 18). The primary target in the regulation of behavior then appears to be the regulation of affect (Gross, 2002; Larsen & Prizmic, 2004). There can be no doubt, however, that strategies for developing impulse control, such as the Reasoning and Rehabilitation Program developed by Ross (1995), can also be useful in developing effective behavioral regulation.

We (Hudson et al., 1993) demonstrated that sex offenders, compared to non-sex offenders and nonoffenders, have significant problems in accurately identifying emotions in other people. The strategy employed in our 1993 study, which required participants to identify facial expressions of emotions, also serves as a measure of emotional recognition skills. Clearly a failure to recognize emotions in others will impede the ability of sex offenders to empathically relate to others, which is a critical feature of all social relations, including the establishment and maintenance of intimate relations.

In addition to these more general features of self-regulation, sex offenders typically fail to cope effectively with the numerous problems that arise in the lives of all people (Serran, Firestone, Marshall, & Moulden, 2007). Research (Cortoni & Marshall, 2001; W. L. Marshall, Serran, & Cortoni, 2000) has also shown that sex offenders, particularly child molesters, characteristically adopt maladaptive coping styles by either becoming absorbed in their own emotional distress or avoiding attending to the problem. Endler and Parker (1990) have developed a sound measure of coping styles that allows a determination of the type of response people consistently make to the problems life presents to them. This permits the identification of persistent maladaptive coping so that treatment can be directed at instilling more adaptive strategies for dealing with life's problems.

Sexual Factors. In a series of studies, L. E. Marshall and his colleagues (L. E. Marshall & Marshall, 2001, 2006; L. E. Marshall, Marshall, Moulden, & Serran, 2008) demonstrated that as many as 40% of sex offenders meet criteria for sexual preoccupation. Hanson and Morton-Bourgon (2005) showed that this characteristic was the most powerful predictor of reoffending among a large sample of sex offenders. L. E. Marshall and Marshall (2010) demonstrated that the most accurate and reliable measure of this propensity was Carne's (1989) Sexual Addiction Screening Test.

We have already considered different strategies (i.e., phallometry and cognitive attentional measures) for assessing sexual deviance, so we will not repeat our concerns here, except to say that the assessment of deviant tendencies is essential and there is a clear need to develop more accurate evaluation procedures. Some features of deviant sexual interests predict reoffending and so must be addressed in treatment.

No doubt, there are many idiosyncratic features or factors related to sexual issues that may be relevant to the full assessment of a sex offender. Two that frequently seem to be salient are the offenders' own history of abuse and his hormonal functioning. A history of sexual, physical, or emotional abuse may have left the offender with many personal deficits and unresolved emotional conflicts, or it may have persuaded him that sexual abuse is normative and is not really problematic (Starzyk & Marshall, 2003). It appears that the best way to determine the incidence of abuse is by interview. Although a disproportionate number of sex offenders report being sexually abused as children (see Hanson & Slater, 1988, for a review of this literature), there is no way to independently verify their reports, and there are obvious self-serving reasons sex offenders may exaggerate or fabricate incidents of abuse.

Hormonal evaluations are costly. Unless the facilities are readily available, it would be impractical and likely not useful to assess every sex offender to determine whether his sex steroid system is functioning normally. There are reports (Bradford, 1990; Hucker & Bain, 1990; Land, 1995) clearly indicating that some, but few, sex offenders have elevated levels of sex steroids. Well-controlled studies have demonstrated that reducing these abnormal levels has a positive therapeutic effect (Bradford, 1990, 1993), so these problems cannot be dismissed.

TREATMENT

Sex offenders are, first and foremost, human beings. More to the point, they are human beings who have themselves, more often than not, suffered physical, sexual, and emotional abuse in their childhood (W. L. Marshall & Marshall, 2000; Starzyk & Marshall, 2003). These observations should encourage therapists to adopt a compassionate but challenging style in dealing with these clients. In any case, there is clear evidence that such a style is crucial to achieving the treatment goal of reduced reoffending (W. L. Marshall & Burton, 2010; W. L. Marshall, Marshall, & Burton 2013). Recent developments in clinical psychology have provided a model for this type of approach. In particular, the Positive Psychology movement (Linley & Joseph, 2004; Peterson, 2006; Seligman & Csikszentmihalyi, 2000; Snyder & Lopez, 2005) and Motivational Interviewing (W. R. Miller & Rollnick, 2002) encourage therapists to attend to the clients' strengths as well as their deficits and adopt a style that fully engages clients. We have assimilated these approaches and framed our treatment program as "strength-based" (W. L. Marshall, Marshall, Serran, et al., 2011), within which is embedded Ward's (2002) Good Lives Model. This

latter model serves as a framework for guiding our clients toward developing the skills, attitudes, and emotional regulatory processes necessary to achieving a more satisfying and offense-free lifestyle.

As our understanding of the range of problems that characterize sexual offenders has increased, so have the targets in treatment been expanded. In the late 1960s, behavior therapists, for example, assumed that sexual offending was motivated solely by deviant sexual preferences. Accordingly, the modification of deviant arousal was the prime, if not the only, focus of treatment (Bond & Evans, 1967). From the beginning of the 1970s, several authors suggested that social skills also needed to be improved (Barlow, 1973; W. L. Marshall, 1971); shortly thereafter, treatment providers added cognitive distortions (Abel, Blanchard, & Becker, 1978) and a broad range of other targets to the treatment of sex offenders (W. L. Marshall, Earls, Segal, & Darke, 1983). This expansion of treatment targets was sometimes based on clinical impressions and sometimes derived from studies showing differences between sex offenders and other men.

It was not until the early 2000s that researchers began to identify potentially modifiable features that predicted reoffending. These “criminogenic targets” met the requirements of Andrews and Bonta’s (2006) *Principles of Effective Offender Treatment*. An earlier meta-analysis (Andrews et al., 1990) had shown that in the treatment of various types of offenders, three principles defined effective programs: (1) the Risk Principle, which directed treatment at the highest risk offenders; (2) the Needs Principle, indicating that treatment must address criminogenic factors; and (3) the Responsivity Principle, requiring skilled therapists to deliver treatment in warm, empathic, supportive, and respectful ways while modeling prosocial attitudes and behavior and reinforcing their occurrence in the offenders. These three principles are collectively referred to as the RNR principles. Hanson, Bourgon, Helmus, and Hodgson (2009) later showed that these same principles apply to sex offender treatment. Unfortunately, the majority of sex offender programs have yet to systematically apply these principles. As a result of the limited applications of these sound bases of treatment, we will limit our description of treatment to our own program, which is based on the RNR principles but is expanded to include motivational and positive psychology features as well as using the Good Lives Model as a framework for treatment.

ROCKWOOD PROGRAM FOR SEXUAL OFFENDERS

The Rockwood Program for Sexual Offenders exclusively addresses the problems of male sex offenders. All treatment is conducted in groups, partly due to limited resources and far too many clients in need of help, but also because evidence appears to suggest that group therapy is the most effective strategy (W. L. Marshall & O’Brien, in press). Treatment involves one (or two if available) therapists guiding 8 to 10 offenders in two 2.5-hour sessions per week. Groups are open-ended (or rolling), meaning that, when one client reaches the goals of treatment, he graduates

and is replaced by a new client. Each client is able to progress at a rate that is personally suited to his learning style and ability. We employ the Therapist Rating Scale (W. L. Marshall et al., 2011) to evaluate each client's progress toward the goals of the program. We have shown that this rating scale displays good interrater reliability and predicts long-term treatment outcome (W. L. Marshall et al., 2011).

In all aspects of our program, the therapists consciously strive to present themselves in ways that are consistent with the evidence on the effective delivery of treatment area. From our review of the general literature on effective therapist characteristics (W. L. Marshall, Fernandez, et al., 2003), we developed a research project examining the features of therapists that predicted positive benefits in sex offender treatment. In two studies (W. L. Marshall, Serran, Fernandez, et al., 2003; Marshall, Serran, Moulden, et al., 2002), we demonstrated that therapists (both male and female) who displayed warmth and empathy and were rewarding and somewhat directive produced significant positive changes in the clients. These observations are consistent with sex offenders' reports of what they saw as the effective elements of treatment (Drapeau, 2005) and are also consistent with Andrews and Bonta's (2006) Core Correctional Practices (CCPs). The CCPs identified by Andrews and Bonta require therapists to deliver treatment in a warm, supportive, and respectful way while modeling and reinforcing prosocial attitudes and behavior. In addition, we also attempt to create cohesive and expressive groups because Beech (Beech & Fordham, 1997; Beech & Hamilton-Giachritsis, 2005) has shown that sex offender groups having these features are by far the most effective.

We conceptualize our program as consisting of three phases: (1) engagement; (2) targeting criminogenic features; and (3) development of postdischarge self-management plans. Since the majority of sex offenders come somewhat reluctantly to treatment, our first phase aims at motivating and engaging them. As a result, we begin Phase 1 by describing treatment in terms that we hope are more appealing than the usual approach. We point out to our clients that our aim is to provide them with the skills, attitudes, and emotional competence necessary to live a fulfilling and satisfying life. We tell clients that this will result in greater happiness than was true of their former lives and will, as a consequence, reduce the risk of a reoffense. We also advise the clients that they will not be required to describe the details of their offense(s), which characteristically puts them at ease. As we noted earlier, the evidence indicates that taking responsibility for their past offending is not a criminogenic factor, so there is no need to address this in treatment. What we want our clients to do is to take responsibility for their future.

In the first phase of treatment, we also address nonthreatening targets, including low self-esteem and shame, both of which present blocks to treatment engagement. In addition, we begin the process involved in helping our clients develop better strategies for coping with difficult situations or events. This serves as one of the effective ways to develop emotional regulation, which, as we have seen, underpins behavioral regulation.

Once the client has become fully engaged and cooperative, he moves to Phase 2 where we target all the identified criminogenic needs (see Table 25.1). The procedures we use to address these targets, as well as those in Phase 1, have been either adopted from procedures that in the general clinical literature have been shown to be effective or have been developed by us. In the latter case, we have demonstrated the effectiveness of these procedures. The evidence for the effectiveness of all the procedures we employ is described in our earliest book on treatment (W. L. Marshall, Anderson, & Fernandez, 1999).

Phase 3 integrates all that has been learned to date in a set of self-management plans that draw on Ward's (2002) Good Lives Model (GLM). The GLM identifies nine areas of human functioning in which all people strive to achieve some degree of fulfillment. This model was derived by Ward from the extensive body of research that was initiated by Maslow's (1968) perspective on self-actualization. This literature identifies the goals that people characteristically seek and the processes by which these goals are achieved. Each client is asked to choose two or three areas of functioning and identify strategies he needs to adopt to achieve higher functioning in each area. We typically encourage the development of occupational and academic skills and the pursuit of leisure activities with which to occupy the client's idle time. Clients are also asked to identify where they intend to live and with whom they will reside, and we ensure that none of the client's intended activities or accommodation will put him in unsupervised contact with potential victims. While we encourage clients to become aware of potential risks in their postrelease life, we do not ask them to generate elaborate relapse prevention plans.

OUTCOME EVALUATIONS

There have been two large-scale meta-analyses of treatment outcome with sex offenders (Hanson et al., 2002; Lösel & Schmucker, 2005) showing that treatment significantly reduces offending. In Hanson et al.'s (2002) study ($N = 9,454$) 16.8% of untreated sex offenders recidivated, whereas only 9.9% of treated offenders did. Lösel and Schmucker reported almost identical rates in an even larger group ($N = 22,181+$) of treated and untreated sex offenders. Many of the studies included in these two analyses did not involve random assignment to the treatment or no-treatment groups. As a result, they have been criticized as failing to produce convincing data (Rice & Harris, 2003). We (W. L. Marshall & Marshall, 2007) have offered an alternative perspective on the value of the random controlled trial, particularly on the practical limitations to doing such a study. Not surprisingly, others (Seto et al., 2008) have vigorously disagreed with our view on this.

We (W. L. Marshall, Marshall, Serran, et al., 2011) have reported outcome data on the Rockwood program that is conducted in a Canadian federal penitentiary. A cohort of 535 sex offenders treated between 1991 and 2001 were followed in the community after discharge from prison. We conducted two appraisals by accessing

the Canadian national database that identifies all charges (even if withdrawn) and convictions across the country. At the 5.4-year average follow-up, the sexual offense rate among treated clients was 3.2% against an expected rate (derived from actuarial risk assessment instruments) of 16.8%. Since we were able to recruit into treatment all but 3.6% of available offenders, there is not a sufficient number of untreated clients to provide a comparison group, thus the need to estimate the likely reoffense rate. Elsewhere, we have provided justifications for this approach (W. L. Marshall & Marshall, 2007), but we realize not all readers will agree with this strategy, so we are in the process of generating a comparison group of untreated offenders released from a similar federal prison. At an 8.4-year follow-up of this same group of 535 sex offenders, we found a slight increase in recidivism in the treated group (5.6%), but the expected rate (derived again from actuarial risk instruments) had also increased (23.8%). In both follow-up studies, treatment was shown to have generalization effects on reducing both nonsexual offending (13.6% against 40.0%) and nonsexual violent offending (8.4% against 34.8%) among our sexual offenders.

These findings provide, along with those reported by Hanson et al. (2002) and Lösel and Schmucker (2005), at least encouraging support for the idea that sex offenders can be effectively treated. It is important to note that the two meta-analyses—Hanson et al. and Lösel and Schmucker—revealed that not all programs were effective, with cognitive-behavioral programs tending to produce the best results. Indeed, recent studies have demonstrated that both multisystemic therapy (Bourduin, Schaeffer, & Heiblum, 2009) and psychodynamic therapy (Kreigman, 2006) effectively reduce sexual offending. It is important to note that both these latter programs addressed appropriate criminogenic features and delivered treatment in a warm, empathic, supportive, and rewarding way. We believe that these qualities in therapists are the critical features of effective treatment along with empirically sound procedures. In fact, the presence of these features (i.e., positive therapist qualities and effective procedures) exerts such a strong influence on outcome that the particular theoretical orientation of programs cannot exert more than a minor benefit, if anything.

CONCLUSIONS

Our review of the application of diagnostic criteria relevant to sex offenders suggests problems. Reliability across diagnosticians seems unsatisfactory, presumably due to the vagueness of the criteria. Alternatives to *DSM* diagnostic criteria, while having some value or offering some promise, also have limitations. Researchers and clinicians working with sex offenders have developed assessment procedures that effectively identify risk to reoffend. Some of these procedures are incorporated into what are referred to as actuarial risk instruments that describe essentially unchangeable features of the offenders. Although these are clearly valuable for various purposes, they offer little direction with respect to treatment planning. Fortunately, there are also instruments that identify potentially modifiable features

that also predict reoffending; these features are referred to as either criminogenic features or stable dynamic factors.

Batteries of tests have been generated to identify treatment needs. These same tests provide methods for evaluating the changes induced by treatment. In both cases, these assessment batteries are most beneficial when they target the evaluation of known criminogenic needs.

Finally, treatment procedures have been developed that, in their best form, generate reductions in reoffense rates. However, it appears that treatment is effective only when it addresses known criminogenic features and is delivered in a way consistent with what is known about positive therapist qualities and effective group climate. When sex offender treatment adheres to these principles, the stated therapeutic orientation of the program is largely irrelevant to effectiveness. We encourage those who offer treatment to sex offenders, and those who treat other types of offenders, to adhere to the principles of effective treatment that are so clearly stated by Andrews and Bonta (2006).

REFERENCES

- Abel, G. G., & Blanchard, E. B. (1976). The measurement and generation of sexual arousal in male sexual deviates. In M. Hersen, R. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification* (Vol. 11, pp. 99–136). New York, NY: Academic Press.
- Abel, G. G., Blanchard, E. B., & Becker, J.V. (1978). An integrated treatment program for rapists. In R.T. Rada (Ed.), *Clinical aspects of the rapist* (pp. 161–214). New York, NY: Grune & Stratton.
- Abel, G. G., Gore, D. K., Holland, C. L., Camp, N., Becker, J. V., & Rathner, J. (1989). The measurement of the cognitive distortions of child molesters. *Annals of Sex Research, 2*, 135–152.
- Abel, G. G., Huffman, J., Warberg, B., & Holland, C. L. (1998). Visual reaction time and plethysmography as measures of sexual interest in child molesters. *Sexual Abuse: A Journal of Research and Treatment, 10*, 81–95.
- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.
- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- Andrews, D. A., & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Cincinnati, OH: Anderson.
- Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior, 17*, 19–52.
- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology, 28*, 369–404.

- Barbaree, H. E. (1990). Stimulus control of sexual arousal: Its role in sexual assault. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 115–142). New York, NY: Plenum Press.
- Barbaree, H. E., & Peacock, E. J. (1995). Phallometric assessment of sexual preferences as an investigative tool in cases of alleged child sexual abuse. In T. Ney (Ed.), *Allegations of child sexual abuse: Assessment and case management* (pp. 242–259). New York, NY: Brunner/Mazel.
- Barlow, D. H. (1973). Increasing heterosexual responsiveness in the treatment of sexual deviation: A review of the clinical and experimental evidence. *Behavior Therapy*, 4, 655–671.
- Baumeister, R. F. (Ed.). (1993). *Self-esteem: The puzzle of low self-regard*. New York, NY: Plenum Press.
- Beech, A. R., & Fordham, A. S. (1997). Therapeutic climate of sexual offender treatment programs. *Sexual Abuse: A Journal of Research and Treatment*, 9, 219–237.
- Beech, A. R., & Hamilton-Giachritsis, C. E. (2005). Relationship between therapeutic climate and treatment outcome in group-based sexual offender treatment programs. *Sexual Abuse: A Journal of Research and Treatment*, 17, 127–140.
- Blair, P. R., Marcus, D. K., & Boccaccini, M. T. (2008). Is there an allegiance effect for assessment instruments? Actuarial risk assessment as an exemplar. *Clinical Psychology: Science and Practice*, 15, 346–360.
- Boer, D. P. (2006). Sexual offender risk assessment strategies: Is there a convergence of opinion yet? *Sexual Offender Treatment*, 1, 1–4.
- Boer, D. P., Hart, S. D., Kropp, P. R., & Webster, C. D. (1997). *Manual for the Sexual Violence Risk –20: Professional guidelines for assessing risk of sexual violence*. Vancouver, Canada: Mental Health, Law, and Policy Institute.
- Bond, I., & Evans, D. (1967). Avoidance therapy: Its use in two cases of underwear fetishism. *Canadian Medical Association Journal*, 96, 1160–1162.
- Borduin, C., Schaeffer, C. M., & Heiblum, N. (2009). A randomised clinical trial of Multisystemic Therapy with juvenile sexual offenders: Effect in youth social ecology and criminal activity. *Journal of Consulting and Clinical Psychology*, 77, 26–37.
- Bradford, J.M.W. (1990). The antiandrogen and hormonal treatment of sex offenders. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 297–310). New York, NY: Plenum Press.
- Bradford, J.M.W. (1993). The pharmacological treatment of the adolescent sex offender. In H. E. Barbaree, W. L. Marshall, & S. M. Hudson (Eds.), *The juvenile sex offender* (pp. 278–288). New York, NY: Guilford Press.
- Briere, J., & Elliott, D. M. (2003). Prevalence and psychological sequelae of self-reported childhood physical and sexual abuse in a general population sample of men and women. *Child Abuse & Neglect: The International Journal*, 27, 1205–1222.
- Browne, A., & Finkelhor, D. (1986). Impact of child sexual abuse: A review of the research. *Psychological Bulletin*, 99, 16–77.
- Bumby, K. M. (1996). Assessing the cognitive distortions of child molesters and rapists: Development and validation of the MOLEST and RAPE scales. *Sexual Abuse: A Journal of Research and Treatment*, 8, 37–54.
- Burt, M. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology*, 38, 217–230.

- Calhoun, K. S., & Wilson, A. E. (2000). Rape and sexual aggression. In L. T. Szuchman & F. Muscarella (Eds.), *Psychological perspectives on human sexuality* (pp. 573–602). New York, NY: Wiley.
- Carnes, P. (1989). *Contrary to love: Helping the sexual addict*. Minneapolis, MN: CompCare.
- Carver, C. S. (2004). Self-regulation of action and affect. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 13–39). New York, NY: Guilford Press.
- Castellino, N. C., Bosco, F. M. B., Marshall, W. L., Marshall, L. E., & Veglia, F. V. (2011). Mind reading abilities in sexual offenders: An analysis of Theory of Mind processes. *Consciousness and Cognition*, 20, 1612–1624.
- Cicchetti, D. V., & Sparrow, S. S. (1981). Developing criteria for establishing interrater reliability of specific items: Applications of assessment of adaptive behavior. *American Journal of Mental Deficiency*, 86, 127–137.
- Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York, NY: Academic Press.
- Cortoni, F. A., & Marshall, W. L. (2001). Sex as a coping strategy and its relationship to juvenile sexual history and intimacy in sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 13, 27–43.
- Di Vasto, P. V., Kaufman, L. R., Jackson, R., Christy, J., Pearson, S., & Burgett, T. (1984). The prevalence of sexually stressful events among females in the general population. *Archives of Sexual Behavior*, 13, 59–67.
- Doren, D. M. (2002). *Evaluating sex offenders: A manual for civil commitments and beyond*. Thousand Oaks, CA: Sage.
- Dowden, C., & Andrews, D. A. (2003). The importance of staff practice in delivering effective correctional treatment: A meta-analysis. *International Journal of Offender Therapy and Comparative Criminology*, 48, 203–214.
- Dowden, C., & Andrews, D. A. (2004). The importance of staff practice in delivering effective correctional treatment: A meta-analytic review of core correctional practice. *International Journal of Offender Therapy and Comparative Criminology*, 48, 203–214.
- Drapeau, M. (2005). Research on the processes involved in treating sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 17, 117–125.
- Duwe, G., & Freske, P. J. (2012). Using logistic regression modelling to predict sexual recidivism: The Minnesota Sex Offender Screening Tool-3 (MnSOST-3). *Sexual Abuse: A Journal of Research and Treatment*, 24, 350–377.
- Endler, N. S., & Parker, J. D. A. (1990). Multidimensional assessment of coping: A critical evaluation. *Journal of Personality and Social Psychology*, 58, 844–854.
- Ford, H. J. (2009). Female sex offenders: Issues and considerations in working with this population. In A. R. Beech, L. A. Craig, & K. D. Browne (Eds.), *Assessment and treatment of sex offenders: A handbook* (pp. 473–488). Chichester, UK: Wiley-Blackwell.
- Freund, K. (1981). Assessment of pedophilia. In M. Cook & K. Howells (Eds.), *Adult sexual interest in children* (pp. 139–179). London, England: Academic Press.
- Freund, K., & Blanchard, R. (1989). Phallometric diagnosis of pedophilia. *Journal of Consulting and Clinical Psychology*, 57, 1–6.
- Freund, K., Chan, S., & Coulthard, R. (1979). Phallometric diagnoses with “nonadmitters.” *Behaviour Research and Therapy*, 17, 451–457.

- Freund, K., & Watson, R. J. (1991). Assessment of the sensitivity and specificity of a phallometric test: An update of phallometric diagnosis of pedophilia. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 3, 254–260.
- Gendreau, P., French, S. A., & Gionet, A. (2004). What works (what doesn't work): The principles of effective correctional treatment. *Journal of Community Corrections*, 13, 4–6.
- Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology*, 39, 281–291.
- Gross, J. J., & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 3–24). New York, NY: Guilford Press.
- Grubin, D. (1998). *Sex offending against children: Understanding the risk*. Police Research Series, Paper 99. London, England: Home Office.
- Hanson, R. K. (1997). *The development of a brief actuarial scale for sex offender recidivism*. Ottawa, Canada: Public Works and Government Services Canada.
- Hanson, R. K. (2006). Stability and change: Dynamic risk factors for sexual offenders. In W. L. Marshall, Y. M. Fernandez, L. E. Marshall, & G. A. Serran (Eds.), *Sexual offender treatment: Controversial issues* (pp. 17–31). Chichester, England: Wiley.
- Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior*, 36, 865–891.
- Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, 66, 348–362.
- Hanson, R. K., Gordon, A., Harris, A.J.R., Marques, J. K., Murphy, W., Quinsey, V. L., & Seto, M. C. (2002). First report of the Collaborative Outcome Data Project on the Effectiveness of Psychological Treatment for Sex Offenders. *Sexual Abuse: A Journal of Research and Treatment*, 14, 169–194.
- Hanson, R. K., & Morton-Bourgon, K. (2004). *Predictors of sexual recidivism: An updated meta-analysis*. Corrections User Report No. 2004–02. Ottawa, Canada: Public Safety and Emergency Preparedness Canada, Corrections Research.
- Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology*, 73, 1154–1163.
- Hanson, R. K., & Morton-Bourgon, K. E. (2007). *The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis*. Corrections User Report No. 2007–01. Ottawa, Canada: Public Safety and Emergency Preparedness Canada, Corrections Research.
- Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment*, 21, 1–21.
- Hanson, R. K., & Slater, S. (1988). Sexual victimization in the history of child sexual abusers: A review. *Annals of Sex Research*, 1, 485–499.
- Hanson, R. K., & Thornton, D. (1999). *Static-99: Improving actuarial risk assessments for sex offenders* (User report 1999–02). Ottawa, Canada: Department of the Solicitor General of Canada.
- Hart, S. D., Kropp, P. R., Laws, D. R., Klaver, J., Logan, C., & Watt, K. A. (2003). *Risk for Sexual Violence Protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence*. Burnaby, BC, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.

- Hart, S. D., Michie, C., & Cooke, D. J. (2007). Precision of actuarial risk assessment instruments: Evaluating the “margins of error” of group v. individual predictions of violence. *British Journal of Psychiatry*, *190*, 60–65.
- Howells, K. (1979). Some meanings of children for pedophiles. In M. Cook & G. Wilson (Eds.), *Love and attraction: An international conference* (pp. 59–526). Oxford, England: Pergamon Press.
- Hucker, S. J., & Bain, J. (1990). Androgenic hormones and sexual assault. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 93–102). New York, NY: Plenum Press.
- Hudson, S. M., Marshall, W. L., Wales, D., McDonald, E., Bakker, L. W., & McLean, A. (1993). Emotional recognition skills of sex offenders. *Annals of Sex Research*, *6*, 199–211.
- Iacono, W. G., & Patrick, C. J. (1999). Polygraph (“lie detector”) testing: The state of the art. In A. K. Hess & I. B. Weiner (Eds.), *The handbook of forensic psychology* (2nd ed., pp. 440–473). New York, NY: Wiley.
- Kingston, D. A., Firestone, P., Moulden, H. A., & Bradford, J. M. (2007). The utility of the diagnosis of pedophilia: A comparison of various classification procedures. *Archives of Sexual Behavior*, *36*, 423–436.
- Kriegman, D. (2006). The reduction of sexual offense recidivism following commitment and psychodynamic treatment. *Journal of Sexual Offender Civil Commitment: Science and the Law*, *1*, 90–98.
- Lalumière, M. L., & Quinsey, V. L. (1994). The discriminability of rapists from non-sex offenders using phallometric measures: A meta-analysis. *Criminal Justice and Behavior*, *21*, 150–175.
- Land, W. B. (1995). Psychopharmacological options for sex offenders. In B. K. Schwartz & H. R. Cellini (Eds.), *The sex offender: Corrections, treatment and legal practice* (pp. 18.1–18.7). Kingston, NJ: Civic Research Institute.
- Larsen, R. J., & Prizmic, Z. (2004). Affect regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 40–61). New York, NY: Guilford Press.
- Laws, D. R., & Marshall, W. L. (2003). A brief history of behavioral and cognitive behavioral approaches to sexual offenders: Part 1. Early developments. *Sexual Abuse: A Journal of Research and Treatment*, *15*, 75–92.
- Levenson, J. S. (2004). Reliability of sexually violent predator civil commitment criteria in Florida. *Law and Human Behavior*, *28*, 357–368.
- Licht, H. (1932). *Sexual life in ancient Greece*. London, England: Routledge & Kegan Paul.
- Linley, P. A., & Joseph, S. (Eds.). (2004). *Positive psychology in practice*. Hoboken, NJ: Wiley.
- Lösel, F., & Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology*, *1*, 117–146.
- Maletzky, B. M. (1991). *Treating the sexual offender*. Newbury Park, CA: Sage.
- Mann, R. E., Hanson, R. K., & Thornton, D. (2011). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse: A Journal of Research and Treatment*, *22*, 191–217.
- Mann, R. E., Webster, S. D., Wakeling, H. C., & Marshall, W. L. (2007). The measurement and influence of child sexual abusive supportive beliefs. *Psychology, Crime & Law*, *13*, 443–458.

- Marshall, L. E., & Marshall, W. L. (2001). Excessive sexual desire disorder among sexual offenders: The development of a research project. *Sexual Addiction & Compulsivity: The Journal of Treatment and Prevention*, 8, 301–307.
- Marshall, L. E., & Marshall, W. L. (2006). Sexual addiction in incarcerated sexual offenders. *Sexual Addiction & Compulsivity: The Journal of Treatment and Prevention*, 13, 377–390.
- Marshall, L. E., & Marshall, W. L. (2010). The factorial structure of the Sexual Addiction Screening Test (SAST) in sexual offenders and socio-economically matched community non-offenders. *Sexual Addiction and Compulsivity: The Journal of Treatment and Prevention*, 17, 210–218.
- Marshall, L. E., & Marshall, W. L. (2011). Empathy and antisocial behaviour. *Journal of Forensic Psychiatry and Psychology*, 22, 742–759.
- Marshall, L. E., & Marshall, W. L., Moulden, H. M., & Serran, G. A. (2008). The prevalence of sexual addiction in incarcerated sexual offenders and matched community nonoffenders. *Sexual Addiction and Compulsivity: The Journal of Treatment and Prevention*, 15, 271–283.
- Marshall, W. L. (1971). A combined treatment method for certain sexual deviations. *Behaviour Research and Therapy*, 9, 292–294.
- Marshall, W. L. (1989). Intimacy, loneliness and sexual offenders. *Behaviour Research and Therapy*, 27, 691–503.
- Marshall, W. L. (1991). *The Justifications Scale: A measure for assessing cognitive distortions in child molesters*. Unpublished scale, Queen's University, Kingston, Ontario, Canada.
- Marshall, W. L. (1996). Psychological evaluation in sexual offence cases. *Queen's Law Journal*, 21, 499–514.
- Marshall, W. L. (1997). Pedophilia: Psychopathology and theory. In D. R. Laws & W. O'Donohue (Eds.), *Handbook of sexual deviance: Theory and application* (pp. 152–179). New York, NY: Guilford Press.
- Marshall, W. L. (1998). Diagnosing and treating sexual offenders. In A. K. Hess & I. B. Weiner (Eds.), *The handbook of forensic psychology* (2nd ed., pp. 640–670). New York, NY: Wiley.
- Marshall, W. L. (2007). Diagnostic issues, multiple paraphilias, and comorbid disorders in sexual offenders: Their incidence and treatment. *Aggression and Violent Behavior*, 12, 16–35.
- Marshall, W. L. (2008). Czy pedofilia jest uleczalna? Wyniki badan pólnocnoamerykańskich (Are pedophiles treatable? Evidence from North American studies). *Seksuologia Polska*, 6, 1–6.
- Marshall, W. L., Anderson, D., & Champagne, F. (1997). The importance of self-esteem in sexual offenders. *Psychology, Crime and Law*, 3, 81–106.
- Marshall, W. L., Anderson, D., & Fernandez, Y. M. (1999). *Cognitive behavioural treatment of sexual offenders*. Chichester, England: Wiley.
- Marshall, W. L., & Burton, D. L. (2010). The importance of therapeutic processes in offender treatment. *Aggression and Violent Behavior*, 15, 141–149.
- Marshall, W. L., Earls, C. M., Segal, Z. V., & Darke, J. (1983). A behavioral program for the assessment and treatment of sexual aggressors. In K. Craig & R. McMahon (Eds.), *Advances in clinical behavior therapy* (pp. 148–174). New York, NY: Brunner/Mazel.
- Marshall, W. L., Eccles, A., & Barbaree, H. E. (1991). Treatment of exhibitionists: A focus on sexual deviance versus cognitive and relationship features. *Behaviour Research and Therapy*, 29, 129–135.

- Marshall, W. L., & Fernandez, Y. M. (2000a). Phallometric testing with sexual offenders: Limits to its value. *Clinical Psychology Review, 20*, 807–822.
- Marshall, W. L., & Fernandez, Y. M. (2000b). Phallometry in forensic practice. *Journal of Forensic Psychology Practice, 1*, 77–87.
- Marshall, W. L., & Fernandez, Y. M. (2003). *Phallometric testing with sexual offenders*. Brandon, VT: Safer Society Press.
- Marshall, W. L., Fernandez, Y. M., Serran, G. A., Mulloy, R., Thornton, D., Mann, R. E., & Anderson, D. (2003). Process variables in the treatment of sexual offenders: A review of the relevant literature. *Aggression and Violent Behavior, 8*, 205–234.
- Marshall, W. L., & Hucker, S. J. (2006). Issues in the diagnosis of sexual sadism. *Sexual Offender Treatment, 1*(2). Retrieved from <http://www.iatso.org/ejournal/>
- Marshall, W. L., & Kennedy, P. (2003). Sexual sadism in sexual offenders: An elusive diagnosis. *Aggression and Violent Behavior, 8*, 1–22.
- Marshall, W. L., Kennedy, P., & Yates, P. (2002). Issues concerning the reliability and validity of the diagnosis of sexual sadism applied in prison settings. *Sexual Abuse: A Journal of Research and Treatment, 14*, 301–311.
- Marshall, W. L., Kennedy, P., Yates, P., & Serran, G. (2002b). Diagnosing sexual sadism in sexual offenders: Reliability across diagnosticians. *International Journal of Offender Therapy and Comparative Criminology, 46*, 668–676.
- Marshall, W. L., & Laws, D. R. (2003). A brief history of behavioral and cognitive behavioral approaches to sexual offender treatment: Part 2. The modern era. *Sexual Abuse: A Journal of Research and Treatment, 15*, 93–120.
- Marshall, W. L., & Marshall, L. E. (2000). The origins of sexual offending. *Trauma, Violence, & Abuse, 1*, 250–263.
- Marshall, W. L., & Marshall, L. E. (2007). The utility of the Random Controlled Trial for evaluating sexual offender treatment: The gold standard or an inappropriate strategy? *Sexual Abuse: A Journal of Research and Treatment, 19*, 175–191.
- Marshall, W. L., & Marshall, L. E. (2010b). Attachment and intimacy in sexual offenders: An update. *Sexual and Marital Therapy, 25*, 86–90.
- Marshall, W. L., & Marshall, L. E. (2012). Treatment of sexual offenders: Effective elements and appropriate outcome evaluations. In E. Bowen & S. J. Brown (Eds.), *Perspectives on evaluating criminal justice and corrections* (pp. 71–94). Bingley, UK: Emerald Publishing.
- Marshall, W. L., Marshall, L. E., & Burton, D. L. (2013). Features of treatment delivery and group processes that maximize the effects of offender programs. In J. L. Wood & T. A. Gannon (Eds.), *Group crime: Psychological explanations* (pp. 159–176). London, England: Willan-Routledge.
- Marshall, W. L., Marshall, L. E., & Kingston, D. A. (2011). Are the cognitive distortions of child molesters in need of treatment? *Journal of Sexual Aggression, 17*, 118–129.
- Marshall, W. L., Marshall, L. E., & Ware, J. (2009). Cognitive distortions in sexual offenders: Should they all be treatment targets? *Sexual Abuse in Australia and New Zealand: An Interdisciplinary Journal, 2*, 70–78.
- Marshall, W. L., Marshall, L. E., Serran, G. A., & O'Brien, M. D. (2011). *Rehabilitating sexual offenders: A strength-based approach*. Washington, DC: American Psychological Association.
- Marshall, W. L., & O'Brien, M. D. (in press). Balancing clients' strengths and deficits in sexual offender treatment: The Rockwood Treatment Approach. In R. Tafrate & D. Mitchell (Eds.), *Forensic CBT: A practitioner's guide*. New York, NY: Wiley-Blackwell.

- Marshall, W. L., Payne, K., Barbaree, H. E., & Eccles, A. (1991). Exhibitionists: Sexual preferences for exposing. *Behaviour Research and Therapy*, 29, 37–40.
- Marshall, W. L., Serran, G. A., & Cortoni, F. A. (2000). Childhood attachments, sexual abuse, and their relationship to adult coping in child molesters. *Sexual Abuse: A Journal of Research and Treatment*, 12, 17–26.
- Marshall, W. L., Serran, G. A., Fernandez, Y. M., Mulloy, R., Mann, R. E., & Thornton, D. (2003). Therapist characteristics in the treatment of sexual offenders: Tentative data on their relationship with indices of behavior change. *Journal of Sexual Aggression*, 9, 25–30.
- Marshall, W. L., Serran, G. A., Moulden, H., Mulloy, R., Fernandez, Y. M., Mann, R. E., & Thornton, D. (2002). Therapist features in sexual offender treatment: Their reliable identification and influence on behavior change. *Clinical Psychology and Psychotherapy*, 9, 395–405.
- Maruna, S. (2001). *Making good: How ex-convicts reform and rebuild their lives*. Washington, DC: American Psychological Association.
- Maslow, A. H. (1968). *Toward a psychology of being* (2nd ed.). New York, NY: Van Nostrand Reinhold.
- Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds?: Coherence among emotion experience, behaviour, and physiology. *Emotion*, 5, 175–190.
- McGrath, R. J., Cumming, G. R., Burchard, B. L., Zeoli, S., & Ellerby, L. (2010). *Current practices and emerging trends in sexual abuser management*. Brandon, VT: Safer Society Press.
- Meehl, P.E. (1986). Causes and effects of my disturbing little book. *Journal of Personality Assessment*, 50, 370–375.
- Miller, R. S., & Lefcourt, H. M. (1982). The assessment of social intimacy. *Journal of Personality Assessment*, 46, 514–518.
- Miller, W. R., & Rollnick, S. (Eds.). (2002). *Motivational interviewing: Preparing people for change* (2nd ed.). New York, NY: Guilford Press.
- Monahan, J. (1981). *Predicting violent behavior: An assessment of clinical techniques*. Beverly Hills, CA: Sage.
- Moulden, H. M., Firestone, P., Kingston, D. A., & Bradford, J. M. W. (2009). Recidivism in pedophiles: An investigation using different methods of defining pedophilia. *Journal of Forensic Psychiatry and Psychology*, 20, 680–701.
- Murphy, K. R., & Davidshofer, C. O. (1998). *Psychological testing: Principles and applications* (4th ed.). Upper Saddle River, NJ: Prentice Hall.
- Murphy, W. D., & Barbaree, H. E. (1994). *Assessments of sex offenders by measures of erectile response: Psychometric properties and decision making*. Brandon, VT: Safer Society Press.
- Mussack, S. E., & Carich, M. S. (2001). Sexual abuser evaluation. In M. S. Carich & S. E. Mussack (Eds.), *Handbook for sexual abuser assessment and treatment* (pp. 11–29). Brandon, VT: Safer Society Press.
- National Research Council (2003). *The polygraph and lie detection*. Washington, DC: The National Academies Press.
- Nelson-Gray, R. D. (1991). DSM-IV: Empirical guidelines from psychometrics. *Journal of Abnormal Psychology*, 100, 308–315.
- Nitschke, J., Mokros, A., Osterheider, M., & Marshall, W. L. (in press). Sexual sadism: Diagnostics and treatment 125 years after the first description by Richard von Krafft-Ebing. *International Journal of Offender Therapy and Comparative Criminology*.

- O'Donohue, W. T., & Letourneau, E. (1992). The psychometric properties of the penile tumescence assessment of child molesters. *Journal of Psychopathology and Behavioral Assessment*, 14, 123–174.
- O'Donohue, W. T., Regev, L. G., & Hagstrom, A. (2000). Problems with the DSM-IV diagnosis of pedophilia. *Sexual Abuse: A Journal of Research and Treatment*, 12, 95–105.
- Packard, R. L., & Levenson, J. S. (2006). Revisiting the reliability of diagnostic decisions on sex offender civil commitment. *Sexual Offender Treatment*, 1, 3. Available online at <http://www.sexual-offender-treatment.org/>
- Paolucci, E., Genuis, M., & Violato, C. (2001). A meta-analysis of the published research on the effects of child sexual abuse. *Journal of Psychology: Interdisciplinary & Applied*, 135, 17–36.
- Perillo, A. D., Mercado, C. C., & Jeglic, E. L. (2011, November). Diagnostic reliability and recommendations for Sexually Violent Predator (SVP) commitment. Poster presented at the 30th annual conference of the Association for the Treatment of Sexual Abusers. Toronto, Canada.
- Peters, J. M., & Murphy, W. D. (1992). Profiling child sexual abusers: Legal considerations. *Criminal Justice and Behavior*, 19, 38–53.
- Peterson, C. (2006). *A primer in positive psychology*. New York, NY: Free Press.
- Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.
- Resick, P. (1993). The psychological impact of rape. *Journal of Interpersonal Violence*, 8, 223–255.
- Rich, P. (2006). *Attachment and sexual offending*. Chichester, England: Wiley.
- Rice, M. E., & Harris, G. T. (2003). The size and signs of treatment effects in sex offender therapy. In P. A. Prentky, E. S. Janus, & M. C. Seto (Eds.), *Annals of the New York Academy of Sciences* (pp. 428–440). New York, NY: New York Academy of Sciences.
- Rosen, R. C., & Beck, G. (1988). *Patterns of sexual arousal: Psychophysiological processes and clinical applications*. New York, NY: Guilford Press.
- Ross, R. R. (1995). The reasoning and rehabilitation program for high-risk probationers and prisoners. In R. R. Ross, D. H. Antonowicz, & G. K. Dhaliwal (Eds.), *Going straight: Effective delinquency prevention and offender rehabilitation* (pp. 195–222). Ottawa, Canada: Air Training and Publications.
- Russell, D., Peplau, L. A., & Cutrona, C. A. (1980). The Revised UCLA Loneliness Scale. *Journal of Personality and Social Psychology*, 39, 472–480.
- Schlenker, B. R., Pontari, B. A., & Christopher, A. N. (2001). Excuses and character: Personal and social implications of excuses. *Personality and Social Psychology Review*, 5, 15–32.
- Seligman, M.E.P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist*, 55, 5–14.
- Serran, G. A. (2002). The measure of empathy. In Y. M. Fernandez (Ed.), *In their shoes: Examining the issue of empathy and its place in the treatment of offenders* (pp. 16–35). Oklahoma City, OK: Wood 'N' Barnes.
- Serran, G. A., Firestone, P., Marshall, W. L., & Moulden, H. (2007). Changes in coping following treatment for child molesters. *Journal of Interpersonal Violence*, 22, 1199–1210.
- Seto, M. C. (2008). *Pedophilia and sexual offending against children: Theory, assessment, and intervention*. Washington, DC: American Psychological Association.

- Seto, M. C., Marques, J. K., Harris, G. T., Chaffin, M., Lalumière, M. L., Miner, M. H., . . . Quinsey, V. L. (2008). Good science and progress in sex offender treatment are intertwined. *Sexual Abuse: A Journal of Research and Treatment, 20*, 247–255.
- Simon, W. T., & Schouten, P.G.W. (1992). Problems in sexual preference testing in child sexual abuse cases: A legal and community perspective. *Journal of Interpersonal Violence, 7*, 503–516.
- Snyder, C. R., & Lopez, S. J. (Eds.). (2005). *Handbook of positive psychology*. New York, NY: Oxford University Press.
- Starzyk, K. B., & Marshall, W. L. (2003). Childhood family and personological risk factors for sexual offending. *Aggression and Violent Behavior, 8*, 93–105.
- Stinson, J. D., Sales, B. D., & Becker, J. V. (2008). *Sex offending: Causal theories to inform research, prevention, and treatment*. Washington, DC: American Psychological Association.
- Storey, J. E., Watt, K. A., Jackson, K. J., & Hart, S. D. (2012). Utilization and implications of the Static-99 in practice. *Sexual Abuse: A Journal of Research and Treatment*. Advance online publication. doi:10.1177/1079063211423943
- Taylor, G. R. (1954). *Sex in history*. New York, NY: Ballantine Books.
- Thornton, D., & Laws, D. R. (2009). *Cognitive approaches to the assessment of sexual interest in sexual offenders*. Chichester, England: Wiley.
- van Dijk, J.J.M., & Mayhew, P. (1992). *Criminal victimization in the industrialized world: Key findings of the 1989 and 1992 International Crime Surveys*. The Hague, the Netherlands: Ministry of Justice, Department of Crime Prevention.
- Vohs, K. D., & Baumeister, R. F. (2004). Understanding self-regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 1–9). New York, NY: Guilford Press.
- Ward, T. (2002). Good lives and the rehabilitation of offenders: Promises and problems. *Aggression and Violent Behavior, 7*, 513–528.
- Ward, T., Polaschek, D.L.L., & Beech, A. R. (2006). *Theories of sexual offending*. Chichester, England: Wiley.
- Webster, C. D., Harris, G. T., Rice, M. E., Cormier, C., & Quinsey, V. L. (1994). *The violence prediction scheme: Assessing dangerousness in high risk men*. Toronto, Canada: University of Toronto Press.
- Wilson, R. J., Abracen, J., Looman, J., Picheca, J. E., & Ferguson, M. (2011). Pedophilia: An evaluation of diagnostic and risk prediction methods. *Sexual Abuse: A Journal of Research and Treatment, 23*, 260–274.
- Wollert, R., Cramer, E., Waggoner, J., Skelton, A. & Vess, J. (2010). Recent research (N = 9,305) underscores the importance of using age-stratified actuarial tables in sex offender risk assessments. *Sexual Abuse: A Journal of Research and Treatment, 22*, 471–490.

Specialty Guidelines for Forensic Psychology

American Psychological Association

In the past 50 years forensic psychological practice has expanded dramatically. The American Psychological Association (APA) has a division devoted to matters of law and psychology (APA Division 41, the American Psychology–Law Society), a number of scientific journals devoted to interactions between psychology and the law exist (e.g., *Law and Human Behavior*; *Psychology, Public Policy, and Law*; *Behavioral Sciences & the Law*), and a number of key texts have been published and undergone multiple revisions (e.g., Grisso, 1986, 2003; Melton, Petrila, Poythress, & Slobogin, 1987, 1997, 2007; Rogers, 1988, 1997, 2008). In addition, training in forensic psychology is available in predoctoral, internship, and postdoctoral settings, and APA recognized forensic psychology as a specialty in 2001, with subsequent recertification in 2008.

Because the practice of forensic psychology differs in important ways from more traditional practice areas (Monahan, 1980) the “Specialty Guidelines for Forensic Psychologists” were developed and published in 1991 (Committee on Ethical Guidelines for Forensic Psychologists, 1991). Because of continued developments in the field in the ensuing 20 years, forensic practitioners’ ongoing need for guidance, and policy requirements of APA, the 1991 “Specialty Guidelines for Forensic Psychologists” were revised, with the intent of benefiting forensic practitioners and recipients of their services alike.

The goals of these Specialty Guidelines for Forensic Psychology (“the Guidelines”) are to improve the quality of forensic psychological services; enhance the practice and facilitate the systematic development of forensic psychology; encourage a high level of quality in professional practice; and encourage forensic practitioners to acknowledge and respect the rights of those they serve. These Guidelines are intended for use by psychologists when engaged in the practice of forensic psychology as described below and may also provide guidance on professional conduct to the legal system and other organizations and professions.

For the purposes of these Guidelines, *forensic psychology* refers to professional practice by any psychologist working within any subdiscipline of psychology (e.g., clinical, developmental, social, cognitive) when applying the scientific, technical, or specialized knowledge of psychology to the law to assist in addressing legal, contractual, and administrative matters. Application of the Guidelines does not depend on the practitioner’s typical areas of practice or expertise, but rather on the service provided in the case at hand. These Guidelines apply in all matters in which psychologists provide expertise to judicial, administrative, and

educational systems including, but not limited to, examining or treating persons in anticipation of or subsequent to legal, contractual, or administrative proceedings; offering expert opinion about psychological issues in the form of amicus briefs or testimony to judicial, legislative, or administrative bodies; acting in an adjudicative capacity; serving as a trial consultant or otherwise offering expertise to attorneys, the courts, or others; conducting research in connection with, or in the anticipation of, litigation; or involvement in educational activities of a forensic nature.

Psychological practice is not considered forensic solely because the conduct takes place in, or the product is presented in, a tribunal or other judicial, legislative, or administrative forum. For example, when a party (such as a civilly or criminally detained individual) or another individual (such as a child whose parents are involved in divorce proceedings) is ordered into treatment with a practitioner, that treatment is not necessarily the practice of forensic psychology. In addition, psychological testimony that is solely based on the provision of psychotherapy and does not include psycholegal opinions is not ordinarily considered forensic practice.

For the purposes of these Guidelines, *forensic practitioner* refers to a psychologist when engaged in the practice of forensic psychology as described above. Such professional conduct is considered forensic from the time the practitioner reasonably expects to, agrees to, or is legally mandated to provide expertise on an explicitly psycholegal issue.

The provision of forensic services may include a wide variety of psycholegal roles and functions. For example, as

This article was published online first October 1, 2012.

These Specialty Guidelines for Forensic Psychology were developed by the American Psychology–Law Society (Division 41 of the American Psychological Association [APA]) and the American Academy of Forensic Psychology. They were adopted by the APA Council of Representatives on August 3, 2011.

The previous version of the Guidelines (“Specialty Guidelines for Forensic Psychologists”; Committee on Ethical Guidelines for Forensic Psychologists, 1991) was approved by the American Psychology–Law Society (Division 41 of APA) and the American Academy of Forensic Psychology in 1991. The current revision, now called the “Specialty Guidelines for Forensic Psychology” (referred to as “the Guidelines” throughout this document), replaces the 1991 “Specialty Guidelines for Forensic Psychologists.”

These guidelines are scheduled to expire August 3, 2021. After this date, users are encouraged to contact the American Psychological Association Practice Directorate to confirm that this document remains in effect.

Correspondence concerning these guidelines should be addressed to the Practice Directorate, American Psychological Association, 750 First Street, NE, Washington, DC 20002-4242.

researchers, forensic practitioners may participate in the collection and dissemination of data that are relevant to various legal issues. As advisors, forensic practitioners may provide an attorney with an informed understanding of the role that psychology can play in the case at hand. As consultants, forensic practitioners may explain the practical implications of relevant research, examination findings, and the opinions of other psycholegal experts. As examiners, forensic practitioners may assess an individual's functioning and report findings and opinions to the attorney, a legal tribunal, an employer, an insurer, or others (APA, 2010b, 2011a). As treatment providers, forensic practitioners may provide therapeutic services tailored to the issues and context of a legal proceeding. As mediators or negotiators, forensic practitioners may serve in a third-party neutral role and assist parties in resolving disputes. As arbiters, special masters, or case managers with decision-making authority, forensic practitioners may serve parties, attorneys, and the courts (APA, 2011b).

These Guidelines are informed by APA's "Ethical Principles of Psychologists and Code of Conduct" (hereinafter referred to as the EPPCC; APA, 2010a). The term *guidelines* refers to statements that suggest or recommend specific professional behavior, endeavors, or conduct for psychologists. Guidelines differ from standards in that standards are mandatory and may be accompanied by an enforcement mechanism. Guidelines are aspirational in intent. They are intended to facilitate the continued systematic development of the profession and facilitate a high level of practice by psychologists. Guidelines are not intended to be mandatory or exhaustive and may not be applicable to every professional situation. They are not definitive, and they are not intended to take precedence over the judgment of psychologists.

As such, the Guidelines are advisory in areas in which the forensic practitioner has discretion to exercise professional judgment that is not prohibited or mandated by the EPPCC or applicable law, rules, or regulations. The Guidelines neither add obligations to nor eliminate obligations from the EPPCC but provide additional guidance for psychologists. The modifiers used in the Guidelines (e.g., *reasonably*, *appropriate*, *potentially*) are included in recognition of the need for professional judgment on the part of forensic practitioners; ensure applicability across the broad range of activities conducted by forensic practitioners; and reduce the likelihood of enacting an inflexible set of guidelines that might be inapplicable as forensic practice evolves. The use of these modifiers, and the recognition of the role of professional discretion and judgment, also reflects that forensic practitioners are likely to encounter facts and circumstances not anticipated by the Guidelines and they may have to act upon uncertain or incomplete evidence. The Guidelines may provide general or conceptual guidance in such circumstances. The Guidelines do not, however, exhaust the legal, professional, moral, and ethical considerations that inform forensic practitioners, for no complex activity can be completely defined by legal rules, codes of conduct, and aspirational guidelines.

The Guidelines are not intended to serve as a basis for disciplinary action or civil or criminal liability. The standard of care is established by a competent authority, not by the Guidelines. No ethical, licensure, or other administrative action or remedy, nor any other cause of action, should be taken *solely* on the basis of a forensic practitioner acting in a manner consistent or inconsistent with these Guidelines.

In cases in which a competent authority references the Guidelines when formulating standards, the authority should consider that the Guidelines attempt to identify a high level of quality in forensic practice. Competent practice is defined as the conduct of a reasonably prudent forensic practitioner engaged in similar activities in similar circumstances. Professional conduct evolves and may be viewed along a continuum of adequacy, and "minimally competent" and "best possible" are usually different points along that continuum.

The Guidelines are designed to be national in scope and are intended to be consistent with state and federal law. In cases in which a conflict between legal and professional obligations occurs, forensic practitioners make known their commitment to the EPPCC and the Guidelines and take steps to achieve an appropriate resolution consistent with the EPPCC and the Guidelines.

The format of the Guidelines is different from most other practice guidelines developed under the auspices of APA. This reflects the history of the Guidelines as well as the fact that the Guidelines are considerably broader in scope than any other APA-developed guidelines. Indeed, these are the only APA-approved guidelines that address a complete specialty practice area. Despite this difference in format, the Guidelines function as all other APA guideline documents.

This document replaces the 1991 "Specialty Guidelines for Forensic Psychologists," which were approved by the American Psychology–Law Society (Division 41 of APA) and the American Board of Forensic Psychology. The current revision has also been approved by the Council of Representatives of APA. Appendix A includes a discussion of the revision process, enactment, and current status of these Guidelines. Appendix B includes definitions and terminology as used for the purposes of these Guidelines.

1. Responsibilities

Guideline 1.01: Integrity

Forensic practitioners strive for accuracy, honesty, and truthfulness in the science, teaching, and practice of forensic psychology and they strive to resist partisan pressures to provide services in any ways that might tend to be misleading or inaccurate.

Guideline 1.02: Impartiality and Fairness

When offering expert opinion to be relied upon by a decision maker, providing forensic therapeutic services, or teaching or conducting research, forensic practitioners strive for accuracy, impartiality, fairness, and independence (EPPCC Standard 2.01). Forensic practitioners rec-

ognize the adversarial nature of the legal system and strive to treat all participants and weigh all data, opinions, and rival hypotheses impartially.

When conducting forensic examinations, forensic practitioners strive to be unbiased and impartial, and avoid partisan presentation of unrepresentative, incomplete, or inaccurate evidence that might mislead finders of fact. This guideline does not preclude forceful presentation of the data and reasoning upon which a conclusion or professional product is based.

When providing educational services, forensic practitioners seek to represent alternative perspectives, including data, studies, or evidence on both sides of the question, in an accurate, fair and professional manner, and strive to weigh and present all views, facts, or opinions impartially.

When conducting research, forensic practitioners seek to represent results in a fair and impartial manner. Forensic practitioners strive to utilize research designs and scientific methods that adequately and fairly test the questions at hand, and they attempt to resist partisan pressures to develop designs or report results in ways that might be misleading or unfairly bias the results of a test, study, or evaluation.

Guideline 1.03: Avoiding Conflicts of Interest

Forensic practitioners refrain from taking on a professional role when personal, scientific, professional, legal, financial, or other interests or relationships could reasonably be expected to impair their impartiality, competence, or effectiveness, or expose others with whom a professional relationship exists to harm (EPPCC Standard 3.06).

Forensic practitioners are encouraged to identify, make known, and address real or apparent conflicts of interest in an attempt to maintain the public confidence and trust, discharge professional obligations, and maintain responsibility, impartiality, and accountability (EPPCC Standard 3.06). Whenever possible, such conflicts are revealed to all parties as soon as they become known to the psychologist. Forensic practitioners consider whether a prudent and competent forensic practitioner engaged in similar circumstances would determine that the ability to make a proper decision is likely to become impaired under the immediate circumstances.

When a conflict of interest is determined to be manageable, continuing services are provided and documented in a way to manage the conflict, maintain accountability, and preserve the trust of relevant others (also see Guideline 4.02 below).

2. Competence

Guideline 2.01: Scope of Competence

When determining one's competence to provide services in a particular matter, forensic practitioners may consider a variety of factors including the relative complexity and specialized nature of the service, relevant training and experience, the preparation and study they are able to devote to the matter, and the opportunity for consultation with a professional of established competence in the sub-

ject matter in question. Even with regard to subjects in which they are expert, forensic practitioners may choose to consult with colleagues.

Guideline 2.02: Gaining and Maintaining Competence

Competence can be acquired through various combinations of education, training, supervised experience, consultation, study, and professional experience. Forensic practitioners planning to provide services, teach, or conduct research involving populations, areas, techniques, or technologies that are new to them are encouraged to undertake relevant education, training, supervised experience, consultation, or study.

Forensic practitioners make ongoing efforts to develop and maintain their competencies (EPPCC Standard 2.03). To maintain the requisite knowledge and skill, forensic practitioners keep abreast of developments in the fields of psychology and the law.

Guideline 2.03: Representing Competencies

Consistent with the EPPCC, forensic practitioners adequately and accurately inform all recipients of their services (e.g., attorneys, tribunals) about relevant aspects of the nature and extent of their experience, training, credentials, and qualifications, and how they were obtained (EPPCC Standard 5.01).

Guideline 2.04: Knowledge of the Legal System and the Legal Rights of Individuals

Forensic practitioners recognize the importance of obtaining a fundamental and reasonable level of knowledge and understanding of the legal and professional standards, laws, rules, and precedents that govern their participation in legal proceedings and that guide the impact of their services on service recipients (EPPCC Standard 2.01).

Forensic practitioners aspire to manage their professional conduct in a manner that does not threaten or impair the rights of affected individuals. They may consult with, and refer others to, legal counsel on matters of law. Although they do not provide formal legal advice or opinions, forensic practitioners may provide information about the legal process to others based on their knowledge and experience. They strive to distinguish this from legal opinions, however, and encourage consultation with attorneys as appropriate.

Guideline 2.05: Knowledge of the Scientific Foundation for Opinions and Testimony

Forensic practitioners seek to provide opinions and testimony that are sufficiently based upon adequate scientific foundation and reliable and valid principles and methods that have been applied appropriately to the facts of the case.

When providing opinions and testimony that are based on novel or emerging principles and methods, forensic practitioners seek to make known the status and limitations of these principles and methods.

Guideline 2.06: Knowledge of the Scientific Foundation for Teaching and Research

Forensic practitioners engage in teaching and research activities in which they have adequate knowledge, experience, and education (EPPCC Standard 2.01), and they acknowledge relevant limitations and caveats inherent in procedures and conclusions (EPPCC Standard 5.01).

Guideline 2.07: Considering the Impact of Personal Beliefs and Experience

Forensic practitioners recognize that their own cultures, attitudes, values, beliefs, opinions, or biases may affect their ability to practice in a competent and impartial manner. When such factors may diminish their ability to practice in a competent and impartial manner, forensic practitioners may take steps to correct or limit such effects, decline participation in the matter, or limit their participation in a manner that is consistent with professional obligations.

Guideline 2.08: Appreciation of Individual and Group Differences

When scientific or professional knowledge in the discipline of psychology establishes that an understanding of factors associated with age, gender, gender identity, race, ethnicity, culture, national origin, religion, sexual orientation, disability, language, socioeconomic status, or other relevant individual and cultural differences affects implementation or use of their services or research, forensic practitioners consider the boundaries of their expertise, make an appropriate referral if indicated, or gain the necessary training, experience, consultation, or supervision (EPPCC Standard 2.01; APA, 2003, 2004, 2011c, 2011d, 2011e).

Forensic practitioners strive to understand how factors associated with age, gender, gender identity, race, ethnicity, culture, national origin, religion, sexual orientation, disability, language, socioeconomic status, or other relevant individual and cultural differences may affect and be related to the basis for people's contact and involvement with the legal system.

Forensic practitioners do not engage in unfair discrimination based on such factors or on any basis proscribed by law (EPPCC Standard 3.01). They strive to take steps to correct or limit the effects of such factors on their work, decline participation in the matter, or limit their participation in a manner that is consistent with professional obligations.

Guideline 2.09: Appropriate Use of Services and Products

Forensic practitioners are encouraged to make reasonable efforts to guard against misuse of their services and exercise professional discretion in addressing such misuses.

3. Diligence

Guideline 3.01: Provision of Services

Forensic practitioners are encouraged to seek explicit agreements that define the scope of, time-frame of, and

compensation for their services. In the event that a client breaches the contract or acts in a way that would require the practitioner to violate ethical, legal, or professional obligations, the forensic practitioner may terminate the relationship.

Forensic practitioners strive to act with reasonable diligence and promptness in providing agreed-upon and reasonably anticipated services. Forensic practitioners are not bound, however, to provide services not reasonably anticipated when retained, nor to provide every possible aspect or variation of service. Instead, forensic practitioners may exercise professional discretion in determining the extent and means by which services are provided and agreements are fulfilled.

Guideline 3.02: Responsiveness

Forensic practitioners seek to manage their workloads so that services can be provided thoroughly, competently, and promptly. They recognize that acting with reasonable promptness, however, does not require the forensic practitioner to acquiesce to service demands not reasonably anticipated at the time the service was requested, nor does it require the forensic practitioner to provide services if the client has not acted in a manner consistent with existing agreements, including payment of fees.

Guideline 3.03: Communication

Forensic practitioners strive to keep their clients reasonably informed about the status of their services, comply with their clients' reasonable requests for information, and consult with their clients about any substantial limitation on their conduct or performance that may arise when they reasonably believe that their clients expect a service that is not consistent with their professional obligations. Forensic practitioners attempt to keep their clients reasonably informed regarding new facts, opinions, or other potential evidence that may be relevant and applicable.

Guideline 3.04: Termination of Services

The forensic practitioner seeks to carry through to conclusion all matters undertaken for a client unless the forensic practitioner–client relationship is terminated. When a forensic practitioner's employment is limited to a specific matter, the relationship may terminate when the matter has been resolved, anticipated services have been completed, or the agreement has been violated.

4. Relationships

Whether a forensic practitioner–client relationship exists depends on the circumstances and is determined by a number of factors, which may include the information exchanged between the potential client and the forensic practitioner prior to, or at the initiation of, any contact or service, the nature of the interaction, and the purpose of the interaction.

In their work, forensic practitioners recognize that relationships are established with those who retain their services (e.g., retaining parties, employers, insurers, the

court) and those with whom they interact (e.g., examinees, collateral contacts, research participants, students). Forensic practitioners recognize that associated obligations and duties vary as a function of the nature of the relationship.

Guideline 4.01: Responsibilities to Retaining Parties

Most responsibilities to the retaining party attach only after the retaining party has requested and the forensic practitioner has agreed to render professional services and an agreement regarding compensation has been reached. Forensic practitioners are aware that there are some responsibilities, such as privacy, confidentiality, and privilege, that may attach when the forensic practitioner agrees to consider whether a forensic practitioner–retaining party relationship shall be established. Forensic practitioners, prior to entering into a contract, may direct the potential retaining party not to reveal any confidential or privileged information as a way of protecting the retaining party’s interest in case a conflict exists as a result of preexisting relationships.

At the initiation of any request for service, forensic practitioners seek to clarify the nature of the relationship and the services to be provided including the role of the forensic practitioner (e.g., trial consultant, forensic examiner, treatment provider, expert witness, research consultant); which person or entity is the client; the probable uses of the services provided or information obtained; and any limitations to privacy, confidentiality, or privilege.

Guideline 4.02: Multiple Relationships

A multiple relationship occurs when a forensic practitioner is in a professional role with a person and, at the same time or at a subsequent time, is in a different role with the same person; is involved in a personal, fiscal, or other relationship with an adverse party; at the same time is in a relationship with a person closely associated with or related to the person with whom the forensic practitioner has the professional relationship; or offers or agrees to enter into another relationship in the future with the person or a person closely associated with or related to the person (EPPCC Standard 3.05).

Forensic practitioners strive to recognize the potential conflicts of interest and threats to objectivity inherent in multiple relationships. Forensic practitioners are encouraged to recognize that some personal and professional relationships may interfere with their ability to practice in a competent and impartial manner, and they seek to minimize any detrimental effects by avoiding involvement in such matters whenever feasible or limiting their assistance in a manner that is consistent with professional obligations.

Guideline 4.02.01: Therapeutic–Forensic Role Conflicts

Providing forensic and therapeutic psychological services to the same individual or closely related individuals involves multiple relationships that may impair objectivity and/or cause exploitation or other harm. Therefore, when requested or ordered to provide either concurrent or se-

quential forensic and therapeutic services, forensic practitioners are encouraged to disclose the potential risk and make reasonable efforts to refer the request to another qualified provider. If referral is not possible, the forensic practitioner is encouraged to consider the risks and benefits to all parties and to the legal system or entity likely to be impacted, the possibility of separating each service widely in time, seeking judicial review and direction, and consulting with knowledgeable colleagues. When providing both forensic and therapeutic services, forensic practitioners seek to minimize the potential negative effects of this circumstance (EPPCC Standard 3.05).

Guideline 4.02.02: Expert Testimony by Practitioners Providing Therapeutic Services

Providing expert testimony about a patient who is a participant in a legal matter does not necessarily involve the practice of forensic psychology even when that testimony is relevant to a psycholegal issue before the decision maker. For example, providing testimony on matters such as a patient’s reported history or other statements, mental status, diagnosis, progress, prognosis, and treatment would not ordinarily be considered forensic practice even when the testimony is related to a psycholegal issue before the decision maker. In contrast, rendering opinions and providing testimony about a person on psycholegal issues (e.g., criminal responsibility, legal causation, proximate cause, trial competence, testamentary capacity, the relative merits of parenting arrangements) would ordinarily be considered the practice of forensic psychology.

Consistent with their ethical obligations to base their opinions on information and techniques sufficient to substantiate their findings (EPPCC Standards 2.04, 9.01), forensic practitioners are encouraged to provide testimony only on those issues for which they have adequate foundation and only when a reasonable forensic practitioner engaged in similar circumstances would determine that the ability to make a proper decision is unlikely to be impaired. As with testimony regarding forensic examinees, the forensic practitioner strives to identify any substantive limitations that may affect the reliability and validity of the facts or opinions offered, and communicates these to the decision maker.

Guideline 4.02.03: Provision of Forensic Therapeutic Services

Although some therapeutic services can be considered forensic in nature, the fact that therapeutic services are ordered by the court does not necessarily make them forensic.

In determining whether a therapeutic service should be considered the practice of forensic psychology, psychologists are encouraged to consider the potential impact of the legal context on treatment, the potential for treatment to impact the psycholegal issues involved in the case, and whether another reasonable psychologist in a similar position would consider the service to be forensic and these Guidelines to be applicable.

Therapeutic services can have significant effects on current or future legal proceedings. Forensic practitioners

are encouraged to consider these effects and minimize any unintended or negative effects on such proceedings or therapy when they provide therapeutic services in forensic contexts.

Guideline 4.03: Provision of Emergency Mental Health Services to Forensic Examinees

When providing forensic examination services an emergency may arise that requires the practitioner to provide short-term therapeutic services to the examinee in order to prevent imminent harm to the examinee or others. In such cases the forensic practitioner is encouraged to limit disclosure of information and inform the retaining attorney, legal representative, or the court in an appropriate manner. Upon providing emergency treatment to examinees, forensic practitioners consider whether they can continue in a forensic role with that individual so that potential for harm to the recipient of services is avoided (EPPCC Standard 3.04).

5. Fees

Guideline 5.01: Determining Fees

When determining fees forensic practitioners may consider salient factors such as their experience providing the service, the time and labor required, the novelty and difficulty of the questions involved, the skill required to perform the service, the fee customarily charged for similar forensic services, the likelihood that the acceptance of the particular employment will preclude other employment, the time limitations imposed by the client or circumstances, the nature and length of the professional relationship with the client, the client's ability to pay for the service, and any legal requirements.

Guideline 5.02: Fee Arrangements

Forensic practitioners are encouraged to make clear to the client the likely cost of services whenever it is feasible and make appropriate provisions in those cases in which the costs of services is greater than anticipated or the client's ability to pay for services changes in some way.

Forensic practitioners seek to avoid undue influence that might result from financial compensation or other gains. Because of the threat to impartiality presented by the acceptance of contingent fees and associated legal prohibitions, forensic practitioners strive to avoid providing professional services on the basis of contingent fees. Letters of protection, financial guarantees, and other security for payment of fees in the future are not considered contingent fees unless payment is dependent on the outcome of the matter.

Guideline 5.03: Pro Bono Services

Forensic psychologists recognize that some persons may have limited access to legal services as a function of financial disadvantage and strive to contribute a portion of their professional time for little or no compensation or personal advantage (EPPCC Principle E).

6. Informed Consent, Notification, and Assent

Because substantial rights, liberties, and properties are often at risk in forensic matters, and because the methods and procedures of forensic practitioners are complex and may not be accurately anticipated by the recipients of forensic services, forensic practitioners strive to inform service recipients about the nature and parameters of the services to be provided (EPPCC Standards 3.04, 3.10).

Guideline 6.01: Timing and Substance

Forensic practitioners strive to inform clients, examinees, and others who are the recipients of forensic services as soon as is feasible about the nature and extent of reasonably anticipated forensic services.

In determining what information to impart, forensic practitioners are encouraged to consider a variety of factors including the person's experience or training in psychological and legal matters of the type involved and whether the person is represented by counsel. When questions or uncertainties remain after they have made the effort to explain the necessary information, forensic practitioners may recommend that the person seek legal advice.

Guideline 6.02: Communication With Those Seeking to Retain a Forensic Practitioner

As part of the initial process of being retained, or as soon thereafter as previously unknown information becomes available, forensic practitioners strive to disclose to the retaining party information that would reasonably be anticipated to affect a decision to retain or continue the services of the forensic practitioner.

This disclosure may include, but is not limited to, the fee structure for anticipated services; prior and current personal or professional activities, obligations, and relationships that would reasonably lead to the fact or the appearance of a conflict of interest; the forensic practitioner's knowledge, skill, experience, and education relevant to the forensic services being considered, including any significant limitations; and the scientific bases and limitations of the methods and procedures which are expected to be employed.

Guideline 6.03: Communication With Forensic Examinees

Forensic practitioners inform examinees about the nature and purpose of the examination (EPPCC Standard 9.03; American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA, & NCME], in press). Such information may include the purpose, nature, and anticipated use of the examination; who will have access to the information; associated limitations on privacy, confidentiality, and privilege including who is authorized to release or access the information contained in the forensic practitioner's records; the voluntary or involuntary nature of participation, including potential consequences of par-

ticipation or nonparticipation, if known; and, if the cost of the service is the responsibility of the examinee, the anticipated cost.

Guideline 6.03.01: Persons Not Ordered or Mandated to Undergo Examination

If the examinee is not ordered by the court to participate in a forensic examination, the forensic practitioner seeks his or her informed consent (EPPCC Standards 3.10, 9.03). If the examinee declines to proceed after being notified of the nature and purpose of the forensic examination, the forensic practitioner may consider postponing the examination, advising the examinee to contact his or her attorney, and notifying the retaining party about the examinee's unwillingness to proceed.

Guideline 6.03.02: Persons Ordered or Mandated to Undergo Examination or Treatment

If the examinee is ordered by the court to participate, the forensic practitioner can conduct the examination over the objection, and without the consent, of the examinee (EPPCC Standards 3.10, 9.03). If the examinee declines to proceed after being notified of the nature and purpose of the forensic examination, the forensic practitioner may consider a variety of options, including postponing the examination, advising the examinee to contact his or her attorney, and notifying the retaining party about the examinee's unwillingness to proceed.

When an individual is ordered to undergo treatment but the goals of treatment are determined by a legal authority rather than the individual receiving services, the forensic practitioner informs the service recipient of the nature and purpose of treatment and any limitations on confidentiality and privilege (EPPCC Standards 3.10, 10.01).

Guideline 6.03.03: Persons Lacking Capacity to Provide Informed Consent

Forensic practitioners appreciate that the very conditions that precipitate psychological examination of individuals involved in legal proceedings can impair their functioning in a variety of important ways, including their ability to understand and consent to the evaluation process.

For examinees adjudicated or presumed by law to lack the capacity to provide informed consent for the anticipated forensic service, the forensic practitioner nevertheless provides an appropriate explanation, seeks the examinee's assent, and obtains appropriate permission from a legally authorized person, as permitted or required by law (EPPCC Standards 3.10, 9.03).

For examinees whom the forensic practitioner has concluded lack capacity to provide informed consent to a proposed, non-court-ordered service but who have not been adjudicated as lacking such capacity, the forensic practitioner strives to take reasonable steps to protect their rights and welfare (EPPCC Standard 3.10). In such cases, the forensic practitioner may consider suspending the pro-

posed service or notifying the examinee's attorney or the retaining party.

Guideline 6.03.04: Evaluation of Persons Not Represented by Counsel

Because of the significant rights that may be at issue in a legal proceeding, forensic practitioners carefully consider the appropriateness of conducting a forensic evaluation of an individual who is not represented by counsel. Forensic practitioners may consider conducting such evaluations or delaying the evaluation so as to provide the examinee with the opportunity to consult with counsel.

Guideline 6.04: Communication With Collateral Sources of Information

Forensic practitioners disclose to potential collateral sources information that might reasonably be expected to inform their decisions about participating that may include, but may not be limited to, who has retained the forensic practitioner; the nature, purpose, and intended use of the examination or other procedure; the nature of and any limits on privacy, confidentiality, and privilege; and whether their participation is voluntary (EPPCC Standard 3.10).

Guideline 6.05: Communication in Research Contexts

When engaging in research or scholarly activities conducted as a service to a client in a legal proceeding, forensic practitioners attempt to clarify any anticipated use of the research or scholarly product, disclose their role in the resulting research or scholarly products, and obtain whatever consent or agreement is required.

In advance of any scientific study, forensic practitioners seek to negotiate with the client the circumstances under and manner in which the results may be made known to others. Forensic practitioners strive to balance the potentially competing rights and interests of the retaining party with the inappropriateness of suppressing data, for example, by agreeing to report the data without identifying the jurisdiction in which the study took place. Forensic practitioners represent the results of research in an accurate manner (EPPCC Standard 5.01).

7. Conflicts in Practice

In forensic psychology practice, conflicting responsibilities and demands may be encountered. When conflicts occur, forensic practitioners seek to make the conflict known to the relevant parties or agencies and consider the rights and interests of the relevant parties or agencies in their attempts to resolve the conflict.

Guideline 7.01: Conflicts With Legal Authority

When their responsibilities conflict with law, regulations, or other governing legal authority, forensic practitioners make known their commitment to the EPPCC and take steps to resolve the conflict. In situations in which the

EPPCC or the Guidelines are in conflict with the law, attempts to resolve the conflict are made in accordance with the EPPCC (EPPCC Standard 1.02).

When the conflict cannot be resolved by such means, forensic practitioners may adhere to the requirements of the law, regulations, or other governing legal authority, but only to the extent required and not in any way that violates a person's human rights (EPPCC Standard 1.03).

Forensic practitioners are encouraged to consider the appropriateness of complying with court orders when such compliance creates potential conflicts with professional standards of practice.

Guideline 7.02: Conflicts With Organizational Demands

When the demands of an organization with which they are affiliated or for whom they are working conflict with their professional responsibilities and obligations, forensic practitioners strive to clarify the nature of the conflict and, to the extent feasible, resolve the conflict in a way consistent with professional obligations and responsibilities (EPPCC Standard 1.03).

Guideline 7.03: Resolving Ethical Issues With Fellow Professionals

When an apparent or potential ethical violation has caused, or is likely to cause, substantial harm, forensic practitioners are encouraged to take action appropriate to the situation and consider a number of factors, including the nature and the immediacy of the potential harm; applicable privacy, confidentiality, and privilege; how the rights of the relevant parties may be affected by a particular course of action; and any other legal or ethical obligations (EPPCC Standard 1.04). Steps to resolve perceived ethical conflicts may include, but are not limited to, obtaining the consultation of knowledgeable colleagues, obtaining the advice of independent counsel, and conferring directly with the client.

When forensic practitioners believe there may have been an ethical violation by another professional, an attempt is made to resolve the issue by bringing it to the attention of that individual, if that attempt does not violate any rights or privileges that may be involved, and if an informal resolution appears appropriate (EPPCC Standard 1.04). If this does not result in a satisfactory resolution, the forensic practitioner may have to take further action appropriate to the situation, including making a report to third parties of the perceived ethical violation (EPPCC Standard 1.05). In most instances, in order to minimize unforeseen risks to the party's rights in the legal matter, forensic practitioners consider consulting with the client before attempting to resolve a perceived ethical violation with another professional.

8. Privacy, Confidentiality, and Privilege

Forensic practitioners recognize their ethical obligations to maintain the confidentiality of information relating to a client or retaining party, except insofar as disclosure is

consented to by the client or retaining party, or required or permitted by law (EPPCC Standard 4.01).

Guideline 8.01: Release of Information

Forensic practitioners are encouraged to recognize the importance of complying with properly noticed and served subpoenas or court orders directing release of information, or other legally proper consent from duly authorized persons, unless there is a legally valid reason to offer an objection. When in doubt about an appropriate response or course of action, forensic practitioners may seek assistance from the retaining client, retain and seek legal advice from their own attorney, or formally notify the drafter of the subpoena or order of their uncertainty.

Guideline 8.02: Access to Information

If requested, forensic practitioners seek to provide the retaining party access to, and a meaningful explanation of, all information that is in their records for the matter at hand, consistent with the relevant law, applicable codes of ethics and professional standards, and institutional rules and regulations. Forensic examinees typically are not provided access to the forensic practitioner's records without the consent of the retaining party. Access to records by anyone other than the retaining party is governed by legal process, usually subpoena or court order, or by explicit consent of the retaining party. Forensic practitioners may charge a reasonable fee for the costs associated with the storage, reproduction, review, and provision of records.

Guideline 8.03: Acquiring Collateral and Third-Party Information

Forensic practitioners strive to access information or records from collateral sources with the consent of the relevant attorney or the relevant party, or when otherwise authorized by law or court order.

Guideline 8.04: Use of Case Materials in Teaching, Continuing Education, and Other Scholarly Activities

Forensic practitioners using case materials for purposes of teaching, training, or research strive to present such information in a fair, balanced, and respectful manner. They attempt to protect the privacy of persons by disguising the confidential, personally identifiable information of all persons and entities who would reasonably claim a privacy interest; using only those aspects of the case available in the public domain; or obtaining consent from the relevant clients, parties, participants, and organizations to use the materials for such purposes (EPPCC Standard 4.07; also see Guidelines 11.06 and 11.07 of these Guidelines).

9. Methods and Procedures

Guideline 9.01: Use of Appropriate Methods

Forensic practitioners strive to utilize appropriate methods and procedures in their work. When performing examinations, treatment, consultation, educational activities, or scholarly investigations, forensic practitioners seek to

maintain integrity by examining the issue or problem at hand from all reasonable perspectives and seek information that will differentially test plausible rival hypotheses.

Guideline 9.02: Use of Multiple Sources of Information

Forensic practitioners ordinarily avoid relying solely on one source of data and corroborate important data whenever feasible (AERA, APA, & NCME, in press). When relying upon data that have not been corroborated, forensic practitioners seek to make known the uncorroborated status of the data, any associated strengths and limitations, and the reasons for relying upon the data.

Guideline 9.03: Opinions Regarding Persons Not Examined

Forensic practitioners recognize their obligations to only provide written or oral evidence about the psychological characteristics of particular individuals when they have sufficient information or data to form an adequate foundation for those opinions or to substantiate their findings (EPPCC Standard 9.01). Forensic practitioners seek to make reasonable efforts to obtain such information or data, and they document their efforts to obtain it. When it is not possible or feasible to examine individuals about whom they are offering an opinion, forensic practitioners strive to make clear the impact of such limitations on the reliability and validity of their professional products, opinions, or testimony.

When conducting a record review or providing consultation or supervision that does not warrant an individual examination, forensic practitioners seek to identify the sources of information on which they are basing their opinions and recommendations, including any substantial limitations to their opinions and recommendations.

10. Assessment

Guideline 10.01: Focus on Legally Relevant Factors

Forensic examiners seek to assist the trier of fact to understand evidence or determine a fact in issue, and they provide information that is most relevant to the psycholegal issue. In reports and testimony, forensic practitioners typically provide information about examinees' functional abilities, capacities, knowledge, and beliefs, and address their opinions and recommendations to the identified psycholegal issues (American Bar Association & American Psychological Association, 2008; Grisso, 1986, 2003; Heilbrun, Marczyk, DeMatteo, & Mack-Allen, 2007).

Forensic practitioners are encouraged to consider the problems that may arise by using a clinical diagnosis in some forensic contexts and to consider and qualify their opinions and testimony appropriately.

Guideline 10.02: Selection and Use of Assessment Procedures

Forensic practitioners use assessment procedures in the manner and for the purposes that are appropriate in light of

the research on or evidence of their usefulness and proper application (EPPCC Standard 9.02; AERA, APA, & NCME, in press). This includes assessment techniques, interviews, tests, instruments, and other procedures and their administration, adaptation, scoring, and interpretation, including computerized scoring and interpretation systems.

Forensic practitioners use assessment instruments whose validity and reliability have been established for use with members of the population assessed. When such validity and reliability have not been established, forensic practitioners consider and describe the strengths and limitations of their findings. Forensic practitioners use assessment methods that are appropriate to an examinee's language preference and competence, unless the use of an alternative language is relevant to the assessment issues (EPPCC Standard 9.02).

Assessment in forensic contexts differs from assessment in therapeutic contexts in important ways that forensic practitioners strive to take into account when conducting forensic examinations. Forensic practitioners seek to consider the strengths and limitations of employing traditional assessment procedures in forensic examinations (AERA, APA, & NCME, in press). Given the stakes involved in forensic contexts, forensic practitioners strive to ensure the integrity and security of test materials and results (AERA, APA, & NCME, in press).

When the validity of an assessment technique has not been established in the forensic context or setting in which it is being used, the forensic practitioner seeks to describe the strengths and limitations of any test results and explain the extrapolation of these data to the forensic context. Because of the many differences between forensic and therapeutic contexts, forensic practitioners consider and seek to make known that some examination results may warrant substantially different interpretation when administered in forensic contexts (AERA, APA, & NCME, in press).

Forensic practitioners consider and seek to make known that forensic examination results can be affected by factors unique to, or differentially present in, forensic contexts, including response style, voluntariness of participation, and situational stress associated with involvement in forensic or legal matters (AERA, APA, & NCME, in press).

Guideline 10.03: Appreciation of Individual Differences

When interpreting assessment results, forensic practitioners consider the purpose of the assessment as well as the various test factors, test-taking abilities, and other characteristics of the person being assessed, such as situational, personal, linguistic, and cultural differences that might affect their judgments or reduce the accuracy of their interpretations (EPPCC Standard 9.06). Forensic practitioners strive to identify any significant strengths and limitations of their procedures and interpretations.

Forensic practitioners are encouraged to consider how the assessment process may be impacted by any disability an examinee is experiencing, make accommodations as

possible, and consider such when interpreting and communicating the results of the assessment (APA, 2011d).

Guideline 10.04: Consideration of Assessment Settings

In order to maximize the validity of assessment results, forensic practitioners strive to conduct evaluations in settings that provide adequate comfort, safety, and privacy.

Guideline 10.05: Provision of Assessment Feedback

Forensic practitioners take reasonable steps to explain assessment results to the examinee or a designated representative in language they can understand (EPPCC Standard 9.10). In those circumstances in which communication about assessment results is precluded, the forensic practitioner explains this to the examinee in advance (EPPCC Standard 9.10).

Forensic practitioners seek to provide information about professional work in a manner consistent with professional and legal standards for the disclosure of test data or results, interpretation of data, and the factual bases for conclusions.

Guideline 10.06: Documentation and Compilation of Data Considered

Forensic practitioners are encouraged to recognize the importance of documenting all data they consider with enough detail and quality to allow for reasonable judicial scrutiny and adequate discovery by all parties. This documentation includes, but is not limited to, letters and consultations; notes, recordings, and transcriptions; assessment and test data, scoring reports and interpretations; and all other records in any form or medium that were created or exchanged in connection with a matter.

When contemplating third-party observation or audio/video-recording of examinations, forensic practitioners strive to consider any law that may control such matters, the need for transparency and documentation, and the potential impact of observation or recording on the validity of the examination and test security (Committee on Psychological Tests and Assessment, American Psychological Association, 2007).

Guideline 10.07: Provision of Documentation

Pursuant to proper subpoenas or court orders or other legally proper consent from authorized persons, forensic practitioners seek to make available all documentation described in Guideline 10.05, all financial records related to the matter, and any other records including reports (and draft reports if they have been provided to a party, attorney, or other entity for review) that might reasonably be related to the opinions to be expressed.

Guideline 10.08: Record Keeping

Forensic practitioners establish and maintain a system of record keeping and professional communication (EPPCC Standard 6.01; APA, 2007) and attend to relevant laws and rules. When indicated by the extent of the rights, liberties,

and properties that may be at risk, the complexity of the case, the amount and legal significance of unique evidence in the care and control of the forensic practitioner, and the likelihood of future appeal, forensic practitioners strive to inform the retaining party of the limits of record-keeping times. If requested to do so, forensic practitioners consider maintaining such records until notified that all appeals in the matter have been exhausted or sending a copy of any unique components/aspects of the record in their care and control to the retaining party before destruction of the record.

11. Professional and Other Public Communications

Guideline 11.01: Accuracy, Fairness, and Avoidance of Deception

Forensic practitioners make reasonable efforts to ensure that the products of their services, as well as their own public statements and professional reports and testimony, are communicated in ways that promote understanding and avoid deception (EPPCC Standard 5.01).

When in their role as expert to the court or other tribunals, the role of forensic practitioners is to facilitate understanding of the evidence or dispute. Consistent with legal and ethical requirements, forensic practitioners do not distort or withhold relevant evidence or opinion in reports or testimony. When responding to discovery requests and providing sworn testimony, forensic practitioners strive to have readily available for inspection all data which they considered, regardless of whether the data supports their opinion, subject to and consistent with court order, relevant rules of evidence, test security issues, and professional standards (AERA, APA, & NCME, in press; Committee on Legal Issues, American Psychological Association, 2006; Bank & Packer, 2007; Golding, 1990).

When providing reports and other sworn statements or testimony in any form, forensic practitioners strive to present their conclusions, evidence, opinions, or other professional products in a fair manner. Forensic practitioners do not, by either commission or omission, participate in misrepresentation of their evidence, nor do they participate in partisan attempts to avoid, deny, or subvert the presentation of evidence contrary to their own position or opinion (EPPCC Standard 5.01). This does not preclude forensic practitioners from forcefully presenting the data and reasoning upon which a conclusion or professional product is based.

Guideline 11.02: Differentiating Observations, Inferences, and Conclusions

In their communications, forensic practitioners strive to distinguish observations, inferences, and conclusions. Forensic practitioners are encouraged to explain the relationship between their expert opinions and the legal issues and facts of the case at hand.

Guideline 11.03: Disclosing Sources of Information and Bases of Opinions

Forensic practitioners are encouraged to disclose all sources of information obtained in the course of their professional services and to identify the source of each piece of information that was considered and relied upon in formulating a particular conclusion, opinion, or other professional product.

Guideline 11.04: Comprehensive and Accurate Presentation of Opinions in Reports and Testimony

Consistent with relevant law and rules of evidence, when providing professional reports and other sworn statements or testimony, forensic practitioners strive to offer a complete statement of all relevant opinions that they formed within the scope of their work on the case, the basis and reasoning underlying the opinions, the salient data or other information that was considered in forming the opinions, and an indication of any additional evidence that may be used in support of the opinions to be offered. The specific substance of forensic reports is determined by the type of psycholegal issue at hand as well as relevant laws or rules in the jurisdiction in which the work is completed.

Forensic practitioners are encouraged to limit discussion of background information that does not bear directly upon the legal purpose of the examination or consultation. Forensic practitioners avoid offering information that is irrelevant and that does not provide a substantial basis of support for their opinions, except when required by law (EPPCC Standard 4.04).

Guideline 11.05: Commenting upon Other Professionals and Participants in Legal Proceedings

When evaluating or commenting upon the work or qualifications of other professionals involved in legal proceedings, forensic practitioners seek to represent their disagreements in a professional and respectful tone and base them on a fair examination of the data, theories, standards, and opinions of the other expert or party.

When describing or commenting upon clients, examinees, or other participants in legal proceedings, forensic practitioners strive to do so in a fair and impartial manner.

Forensic practitioners strive to report the representations, opinions, and statements of clients, examinees, or other participants in a fair and impartial manner.

Guideline 11.06: Out-of-Court Statements

Ordinarily, forensic practitioners seek to avoid making detailed public (out-of-court) statements about legal proceedings in which they have been involved. However, sometimes public statements may serve important goals, such as educating the public about the role of forensic practitioners in the legal system, the appropriate practice of forensic psychology, and psychological and legal issues that are relevant to the matter at hand. When making public statements, forensic practitioners refrain from releasing

private, confidential, or privileged information and attempt to protect persons from harm, misuse, or misrepresentation as a result of their statements (EPPCC Standard 4.05).

Guideline 11.07: Commenting Upon Legal Proceedings

Forensic practitioners strive to address particular legal proceedings in publications or communications only to the extent that the information relied upon is part of a public record or when consent for that use has been properly obtained from any party holding any relevant privilege (also see Guideline 8.04).

When offering public statements about specific cases in which they have not been involved, forensic practitioners offer opinions for which there is sufficient information or data and make clear the limitations of their statements and opinions resulting from having had no direct knowledge of or involvement with the case (EPPCC Standard 9.01).

REFERENCES

- American Bar Association & American Psychological Association. (2008). *Assessment of older adults with diminished capacity: A handbook for psychologists*. Washington, DC: Authors.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (in press). *Standards for educational and psychological testing* (3rd ed.). Washington, DC: Authors.
- American Psychological Association. (2002). Criteria for practice guideline development and evaluation. *American Psychologist*, 57, 1048–1051. doi:10.1037/0003-066X.57.12.1048
- American Psychological Association. (2003). Guidelines on multicultural education, training, research, practice, and organizational change for psychologists. *American Psychologist*, 58, 377–402. doi:10.1037/0003-066X.58.5.377
- American Psychological Association. (2004). Guidelines for psychological practice with older adults. *American Psychologist*, 59, 4, 236–260. doi:10.1037/0003-066X.59.4.236
- American Psychological Association. (2007). Record keeping guidelines. *American Psychologist*, 62, 993–1004. doi:10.1037/0003-066X.62.9.993
- American Psychological Association. (2010a). *Ethical principles of psychologists and code of conduct* (2002, Amended June 1, 2010). Retrieved from <http://www.apa.org/ethics/code/index.aspx>
- American Psychological Association. (2010b). Guidelines for child custody evaluations in family law proceedings. *American Psychologist*, 65, 863–867. doi:10.1037/a0021250
- American Psychological Association. (2011a). *Guidelines for psychological evaluations in child protection matters*. Washington, DC: Author. (*American Psychologist*, 2013, 68, 20–31. doi:10.1037/a0029891)
- American Psychological Association. (2011b). *Guidelines for the practice of parenting coordination*. Washington, DC: Author. (*American Psychologist*, 2012, 67, 63–71. doi:10.1037/a0024646)
- American Psychological Association. (2011c). *Guidelines for the evaluation of dementia and age related cognitive change*. Washington, DC: Author. (*American Psychologist*, 2012, 67, 1–9. doi:10.1037/a0024643)
- American Psychological Association. (2011d). *Guidelines for assessment of and intervention with persons with disabilities*. Washington, DC: Author. (*American Psychologist*, 2012, 67, 43–62. doi:10.1037/a0025892)
- American Psychological Association. (2011e). *Guidelines for psychological practice with lesbian, gay, and bisexual clients*. Washington, DC: Author. (*American Psychologist*, 2012, 67, 10–42. doi:10.1037/a0024659)
- Bank, S., & Packer, R. (2007). Expert witness testimony: Law, ethics, and

- practice. In A. M. Goldstein (Ed.), *Forensic psychology: Emerging topics and expanding roles* (pp. 421–445). Hoboken, NJ: Wiley.
- Committee on Ethical Guidelines for Forensic Psychologists. (1991). Specialty guidelines for forensic psychologists. *Law and Human Behavior*, 15, 655–665.
- Committee on Legal Issues, American Psychological Association. (2006). Strategies for private practitioners coping with subpoenas or compelled testimony for client records or test data. *Professional Psychology: Research and Practice*, 37, 215–222. doi:10.1037/0735-7028.37.2.215
- Committee on Psychological Tests and Assessment, American Psychological Association. (2007). *Statement on third party observers in psychological testing and assessment: A framework for decision making*. Washington, DC: Author. Retrieved from <http://www.apa.org/science/programs/testing/third-party-observers.pdf>
- Golding, S. L. (1990). Mental health professionals and the courts: The ethics of expertise. *International Journal of Law and Psychiatry*, 13, 281–307. doi:10.1016/0160-2527(90)90023-V
- Grisso, T. (1986). *Evaluating competencies: Forensic assessments and instruments*. New York, NY: Plenum Press.
- Grisso, T. (2003). *Evaluating competencies: Forensic assessments and instruments* (2nd ed.). New York, NY: Kluwer/Plenum.
- Heilbrun, K., Marczyk, G., DeMatteo, D., & Mack-Allen, J. (2007). A principles-based approach to forensic mental health assessment: Utility and update. In A. M. Goldstein (Ed.), *Forensic psychology: Emerging topics and expanding roles* (pp. 45–72). Hoboken, NJ: Wiley.
- Melton, G., Pettila, J., Poythress, N., & Slobogin, C. (1987). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers*. New York, NY: Guilford Press.
- Melton, G., Pettila, J., Poythress, N., & Slobogin, C. (1997). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (2nd ed.). New York, NY: Guilford Press.
- Melton, G., Pettila, J., Poythress, N., & Slobogin, C. (2007). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (3rd ed.). New York, NY: Guilford Press.
- Monahan, J. (Ed.). (1980). *Who is the client? The ethics of psychological intervention in the criminal justice system*. Washington, DC: American Psychological Association. doi:10.1037/10051-000
- Rogers, R. (Ed.). (1988). *Clinical assessment of malingering and deception*. New York, NY: Guilford Press.
- Rogers, R. (Ed.). (1997). *Clinical assessment of malingering and deception* (2nd ed.). New York, NY: Guilford Press.
- Rogers, R. (Ed.). (2008). *Clinical assessment of malingering and deception* (3rd ed.). New York, NY: Guilford Press.

Appendix A

Revision Process of the Guidelines

This revision of the Guidelines was coordinated by the Committee for the Revision of the Specialty Guidelines for Forensic Psychology (“the Revisions Committee”), which was established by the American Academy of Forensic Psychology and the American Psychology–Law Society (Division 41 of the American Psychological Association [APA]) in 2002 and which operated through 2011. This committee consisted of two representatives from each organization (Solomon Fulero, PhD, JD; Stephen Golding, PhD, ABPP; Lisa Piechowski, PhD, ABPP; Christina Studebaker, PhD), a chairperson (Randy Otto, PhD, ABPP), and a liaison from Division 42 (Psychologists in Independent Practice) of APA (Jeffrey Younggren, PhD, ABPP).

This document was revised in accordance with APA Rule 30.08 and the APA policy document “Criteria for Practice Guideline Development and Evaluation” (APA, 2002). The Revisions Committee posted announcements regarding the revision process to relevant electronic discussion lists and professional publications (i.e., the Psy-Law-L e-mail listserv of the American Psychology–Law Society, the American Academy of Forensic Psychology listserv, the American Psychology–Law Society Newslet-

ter). In addition, an electronic discussion list devoted solely to issues concerning revision of the Guidelines was operated between December 2002 and July 2007, followed by establishment of an e-mail address in February 2008 (sgfp@yahoo.com). Individuals were invited to provide input and commentary on the existing Guidelines and proposed revisions via these means. In addition, two public meetings were held throughout the revision process at biennial meetings of the American Psychology–Law Society.

Upon development of a draft that the Revisions Committee deemed suitable, the revised Guidelines were submitted for review to the Executive Committee of the American Psychology–Law Society (Division 41 of APA) and the American Board of Forensic Psychology. Once the revised Guidelines were approved by these two organizations, they were submitted to APA for review, commentary, and acceptance, consistent with APA’s “Criteria for Practice Guideline Development and Evaluation” (APA, 2002) and APA Rule 30-8. They were subsequently revised by the Revisions Committee and were adopted by the APA Council of Representatives on August 3, 2011.

(Appendices continue)

Appendix B

Definitions and Terminology

For the purposes of these Guidelines:

Appropriate, when used in relation to conduct by a forensic practitioner, means that, according to the prevailing professional judgment of competent forensic practitioners, the conduct is apt and pertinent and is considered befitting, suitable, and proper for a particular person, place, condition, or function. **Inappropriate** means that, according to the prevailing professional judgment of competent forensic practitioners, the conduct is not suitable, desirable, or properly timed for a particular person, occasion, or purpose; and may also denote improper conduct, improprieties, or conduct that is discrepant for the circumstances.

Agreement refers to the objective and mutual understanding between the forensic practitioner and the person or persons seeking the professional service and/or agreeing to participate in the service. See also Assent, Consent, and Informed Consent.

Assent refers to the agreement, approval, or permission, especially regarding verbal or nonverbal conduct, that is reasonably intended and interpreted as expressing willingness, even in the absence of unmistakable consent. Forensic practitioners attempt to secure assent when consent and informed consent cannot be obtained or when, because of mental state, the examinee may not be able to consent.

Consent refers to agreement, approval, or permission as to some act or purpose.

Client refers to the attorney, law firm, court, agency, entity, party, or other person who has retained, and who has a contractual relationship with, the forensic practitioner to provide services.

Conflict of Interest refers to a situation or circumstance in which the forensic practitioner's objectivity, impartiality, or judgment may be jeopardized due to a relationship, financial, or any other interest that would reasonably be expected to substantially affect a forensic practitioner's professional judgment, impartiality, or decision making.

Decision Maker refers to the person or entity with the authority to make a judicial decision, agency determination, arbitration award, or other contractual determination after consideration of the facts and the law.

Examinee refers to a person who is the subject of a forensic examination for the purpose of informing a decision maker or attorney about the psychological functioning of that examinee.

Forensic Examiner refers to a psychologist who examines the psychological condition of a person whose psychological condition is in controversy or at issue.

Forensic Practice refers to the application of the scientific, technical, or specialized knowledge of psychol-

ogy to the law and the use of that knowledge to assist in resolving legal, contractual, and administrative disputes.

Forensic Practitioner refers to a psychologist when engaged in forensic practice.

Forensic Psychology refers to all forensic practice by any psychologist working within any subdiscipline of psychology (e.g., clinical, developmental, social, cognitive).

Informed Consent denotes the knowledgeable, voluntary, and competent agreement by a person to a proposed course of conduct after the forensic practitioner has communicated adequate information and explanation about the material risks and benefits of, and reasonably available alternatives to, the proposed course of conduct.

Legal Representative refers to a person who has the legal authority to act on behalf of another.

Party refers to a person or entity named in litigation, or who is involved in, or is witness to, an activity or relationship that may be reasonably anticipated to result in litigation.

Reasonable or Reasonably, when used in relation to conduct by a forensic practitioner, denotes the conduct of a prudent and competent forensic practitioner who is engaged in similar activities in similar circumstances.

Record or Written Record refers to all notes, records, documents, memorializations, and recordings of considerations and communications, be they in any form or on any media, tangible, electronic, handwritten, or mechanical, that are contained in, or are specifically related to, the forensic matter in question or the forensic service provided.

Retaining Party refers to the attorney, law firm, court, agency, entity, party, or other person who has retained, and who has a contractual relationship with, the forensic practitioner to provide services.

Tribunal denotes a court or an arbitrator in an arbitration proceeding, or a legislative body, administrative agency, or other body acting in an adjudicative capacity. A legislative body, administrative agency, or other body acts in an adjudicative capacity when a neutral official, after the presentation of legal argument or evidence by a party or parties, renders a judgment directly affecting a party's interests in a particular matter.

Trier of Fact refers to a court or an arbitrator in an arbitration proceeding, or a legislative body, administrative agency, or other body acting in an adjudicative capacity. A legislative body, administrative agency, or other body acts in an adjudicative capacity when a neutral official, after the presentation of legal argument or evidence by a party or parties, renders a judgment directly affecting a party's interests in a particular matter.

Author Index

- Aamodt, M. G., 594
Abbott, E., 763
Abdollahi-Arena, M. K., 649
Abe, N., 643, 644
Abel, G. G., 844, 845, 849, 853
Abeles, N., 102
Abidin, R., 157
Abracen, J., 841
Abraham, K. S., 478
Abram, K. M., 811
Abramson, J., 482
Achenbach, T. M., 158
Ackerman, M. C., 137, 157, 163
Ackerman, M. J., 137, 156, 157,
163, 294, 726
Acklin, M. W., 290, 294
Adamo, K. R., 482
Adams, G. C., 625
Aderhold, B., 112
Adler, S. J., 482
Afifi, T., 239
Agnew, R., 764, 766
Ahern, E. C., 583
Ahrons, C., 142, 148
Aiken, L. S., 823
Aitken, P., 798
Akers, R. L., 766
Alban, A. D., 227
Albertson, S., 479
Albrecht, S. F., 209
Alexander, K. W., 565, 570,
571, 576
Alexander, W., 392
Alfini, J. J., 471
Alia-Klein, N., 238
Alink, L., 237, 238, 243
Alkon, A., 570
Allen, C. J., 636
Allen, F. A., 763
Allen, J. J., 636
Allen, J. J. B., 634–636, 638,
640–642
Allen, S. W., 457
Alley, D., 563, 590
Allhusen, V. D., 582
Alloway, W. R., 626
Almvik, R., 392
Alterman, A. I., 800
Althouse, R., 760, 768, 781
Althouse, R. L. R., 796
Amacher, E., 561
Amato, S., 626
Amigo, E., 226
Anastasi, J. S., 525
Anderson, C. M., 570
Anderson, D., 849, 854, 855
Anderson, S. D., 289
Anderson, S. K., 88
Andreasen, N., 331
Andres-Pueyo, A., 406
Andrews, D. A., 48, 119, 392,
395, 401, 402, 414, 427,
759, 760, 766, 769–774,
777, 782, 796, 797, 801,
802, 804, 805, 810, 815,
817, 848, 849, 853, 854,
857
Angell, A., 632
Angell, M., 333
Annas, G. J., 231
Antonacci, D. J., 814
Antonowicz, D. H., 802
Aos, S., 777
App, D. C., 670
Appelbaum, P., 49, 283, 392
Appelbaum, P. S., 115, 128,
232, 233, 334, 417, 419,
420
Apple, B. N., 338
Applebaum, P. S., 231, 693
Applegate, B. K., 766
Appleyard, K., 246, 254, 256
Arbach-Lucioni, K., 406
Arbour, L., 781
Arbuckle, B., 244
Archer, R. P., 187
Areh, I., 526
Arenella, P., 364
Arlen, J. H., 478
Armesto, J. C., 233
Armitage, T. E., 625
Armstrong, B., 773, 805
Armstrong, D., 820
Armstrong, G. S., 769
Arndt, S., 331, 795
Arnsberger, P., 239
Arons, H., 668, 694
Arredondo, P., 587
Arrigo, B. A., 322
Artwohl, A., 463
Ashford, J. B., 819
Atabaki, A., 584
Athaide-Victor, E., 502
Augimeri, L. K., 396, 406
Augusti, E. M., 563
Ault, R. L., 669
Austin, W., 149, 151, 154, 159
Ax, R. F., 21
Ax, R. K., 760, 768, 811–814
Axelrad, A. D., 694
Ayers, M. S., 572
Ayoub, C., 250, 254, 261–263,
265, 269
Aytes, K. E., 805
Azikri, D., 663
Babchishin, K. M., 28
Babcock, J. C., 800
Babiker, G., 241, 265
Babitsky, S., 740, 744, 755
Bachrach, L. L., 806
Bader, S., 406
Bae, Hwa-ok, 238
Baer, J., 241
Bagby, R. M., 297
Bailey, F. L., 694
Bailis, D., 341–343
Bain, J., 852
Bakanosky, S., 573

- Baker, M. T., 462
 Baker, S. J., 584, 590
 Baker, T., 471
 Bakermans-Kranenburg, M., 241
 Baker-Ward, L., 567, 569, 581
 Bakker, L. W., 851
 Baksheev, G., 43
 Bala, N., 150, 561, 569, 585
 Balaratnasingam, S., 715
 Bales, W. D., 770, 771
 Ball, P. J., 592
 Ball, S., 778
 Ballantine, H. W., 320
 Baltruweit, W. J., 689
 Bandura, A., 27, 773
 Banes, J., 808
 Banister, P. A., 767
 Banks, S., 115, 128
 Banyard, V., 243
 Baradoran, L., 244
 Barbarash, R. A., 233
 Barbaree, H. E., 843–846
 Barber, T. X., 682, 690
 Barcus, E., 156
 Barden, R. C., 697
 Barger, S. D., 628, 629
 Bar-Hillel, M., 650
 Barland, G. H., 9, 628, 629
 Barlow, D. H., 853
 Barnes, H. E., 19
 Barnes, M., 691
 Barnhardt, J., 643
 Barnhouse, R., 331
 Barnier, A. J., 682, 688, 689
 Barnum, D., 251, 252
 Barnum, R., 250, 269, 303
 Baron, R. A., 27
 Barratt, E. S., 353
 Barrett, B., 251, 252, 820
 Barrett, C. L., 225, 227, 228
 Barrett, G. V., 711
 Barron, P., 808
 Bartel, P., 396, 402
 Bartholomew, R. E., 715
 Bartol, A., 111
 Bartol, A. M., 23
 Bartol, C., 42, 111
 Bartol, C. R., 22, 23
 Barton, C. D., 232
 Bartos, L., 747
 Batastini, A. B., 810
 Batchelor, E. S., 297
 Bathurst, K., 157
 Batista Gonçalves, I., 597
 Batterman-Faunce, J. M., 566, 569, 571, 576, 596
 Batts, A. W., 450, 456
 Baucom, D. H., 823
 Bauer, R., 807
 Bauer, W. D., 243
 Bauermann, T., 646
 Baugerud, G. A., 570, 571
 Baum, K. M., 457
 Baumeister, R. F., 850
 Baumrind, N., 570
 Beach, S., 238, 239
 Beahrs, J. O., 685, 689, 694
 Bearse, J., 96
 Beasley, W. R., 323
 Beatty, J., 518
 Beccaria, C., 761
 Beck, C., 46
 Beck, C. J. A., 117, 131
 Beck, G., 844
 Beck, J. C., 385, 417, 426
 Becker, J. V., 800, 840, 849, 853
 Beddow, P. A., 203
 Bederian-Gardner, D., 573
 Beech, A., 805
 Beech, A. R., 819, 840, 854
 Beecher, H. K., 228
 Beer, J., 580
 Beggs, S. M., 817, 818
 Beidleman, W. B., 682
 Beier, K., 256
 Bekerian, D. A., 572
 Belbin, E., 552
 Belenko, S., 42
 Belfrage, H., 396, 409, 411, 413, 428
 Belli, M. M., 694
 Belsky, J., 238–247, 254
 Belue, K., 226
 Benedict, K., 479
 Beneteau, J. L., 402
 Benjamin, G., 426
 Benjamin, G. A. H., 112
 Benjamin, L. T., 10
 Bennell, C., 43
 Bennett, B. B., 711
 Bennett, L. A., 766
 Ben-Porath, Y. S., 187, 188, 448, 449
 Ben-Shakhar, G., 627, 628, 631, 637, 650
 Beran, B., 663
 Berdie, R., 19
 Beresford, J., 588
 Berg, J. W., 693
 Berge, G., 283, 286–288, 290, 291, 293, 336
 Bergman, B. A., 800
 Bergman, E., 572
 Bergstrom, M. K., 215
 Berkowitz, L., 27
 Berkowitz, N., 226
 Berlin, L., 246, 254, 256
 Berliner, L., 267
 Bermant, G., 479
 Bernheim, H., 667, 691
 Bernstein, D., 586
 Bernstein, H., 471
 Bernstein, M. J., 525
 Berntsen, D., 571
 Berry, D. T. R., 716
 Berry, S., 802
 Bersoff, D., 426
 Bersoff, D. N., 5, 28, 198, 199, 207
 Bertelli, M., 197
 Bertman, L. J., 289
 Bertman-Pate, L. J., 334
 Best, S. R., 454
 Beutler, L. E., 721
 Bewley, M. T., 801, 802, 806
 Bexton, W. H., 780
 Bilker, W. B., 644
 Binder, R. L., 388, 420, 426
 Binet, A., 6, 8
 Biringen, Z., 148
 Birmingham, L., 803
 Bishopp, D., 245
 Bithoney, W., 240
 Bizzi, E., 645
 Bjork, R., 215
 Bjorklund, D. F., 574
 Black, H. C., 317, 357, 358
 Blackburn, R., 28
 Blades, M., 588
 Blair, P. R., 394, 395, 401, 847
 Blais, M. A., 726
 Blakely, C. H., 777
 Blanchard, E. B., 844, 853
 Blanchard, R., 844
 Blandon-Gitlin, I., 536, 575
 Blank, M. B., 813
 Blau, T., 446
 Blau, T. H., 733, 734, 746, 751
 Bleechmore, J., 16
 Blinder, M., 365, 366
 Bloch, R. M., 814
 Block, E. B., 668
 Block, S. D., 570, 593

- Bloom, J. D., 322
 Blue, J., 243
 Boaz, T. L., 642
 Boccaccini, M. T., 187, 287,
 288, 290, 291, 336, 394,
 395, 401, 741, 753, 847
 Bocian, K. M., 208
 Boddy, J., 297
 Bodemar, O. A., 19
 Boer, D. P., 396, 406, 407,
 847, 848
 Bohl, N., 452
 Bohl-Penrod, N., 453
 Bohni, M. K., 571
 Böker, W., 330
 Bolt, D. M., 203
 Bolt, S. E., 203
 Bolton, F. E., 6
 Bolton, N., 767
 Bonanno, G. A., 584
 Bonczar, T., 764
 Bond, G. R., 807
 Bond, I., 853
 Bonnema, M. C., 665
 Bonnie, R., 50, 359
 Bonnie, R. J., 281, 283, 285, 286,
 297–299, 301, 302, 323,
 326
 Bonta, J., 48, 119, 392, 395, 401,
 414, 759, 760, 766–771,
 773, 777, 780, 796, 797,
 801, 802, 804, 805, 810,
 815, 817, 848, 853, 854,
 857
 Bordens, K. S., 475, 479, 493,
 495
 Borduin, C., 856
 Boren, B., 645
 Borgida, E., 42, 592
 Boring, E. G., 19
 Bornstein, B., 476
 Bornstein, B. H., 10, 470, 473,
 525, 528, 536, 564, 565,
 586, 591
 Bornstein, R. F., 103
 Borthwick-Duffy, S. L., 210
 Borum, R., 43, 50, 254, 256,
 287, 315, 322, 328, 341,
 396, 402, 411, 417–419,
 426, 448, 449, 451, 810
 Bosco, F. M. B., 850
 Bosh, G., 634, 638–641
 Bothwell, R. K., 690
 Bottoms, B. L., 49, 562, 572,
 575, 585, 589–591
 Boucher, R. J., 386–388, 408,
 411
 Boulas, J., 112
 Bouman, Y., 406
 Bourgeois, M., 493–495, 497,
 501
 Bourgeois, M. J., 479
 Bourgon, G., 760, 773, 805, 853
 Bow, J. N., 138
 Bowers, J. M., 572
 Bowers, K. S., 682, 690, 691
 Bowlby, J., 241, 565
 Bowman, J., 474, 476
 Boyce, W. T., 570
 Brackenridge, C., 245
 Bradford, J. M., 841
 Bradford, J. M. W., 841, 852
 Bradley, V. J., 210
 Braff, J. J., 330, 336
 Brainerd, C. J., 572–575
 Braithwaite, E., 406
 Braithwaite, R. L., 340
 Brakel, S. J., 294
 Brame, R., 821
 Brandenburg, C. A., 297
 Brandon, P., 238
 Braver, S., 148, 151
 Bremner, J. D., 181, 571
 Brent, B., 241
 Brett, A. S., 231
 Brewer, N., 527, 530, 541, 542
 Bricklin, B., 156
 Bricklin, P., 711
 Briere, J., 47, 839
 Briggs, C. S., 765, 766
 Briggs, S. R., 297
 Brigham, J., 36
 Brigham, J. C., 4, 5, 13, 15, 16,
 525, 527, 532
 Bright-Paul, A., 582
 Brimacombe, C. A. E., 538
 Brink, J., 396, 406, 407
 Britt, S. H., 12
 Brobst, K. E., 449
 Brocato, J., 820
 Brockman, J., 36
 Brockway, J. H., 475
 Brockway, Z. R., 763
 Broder, J. M., 470
 Brodie-Innes, J. W., 664
 Brodsky, A., 426
 Brodsky, S. L., 21, 22, 44, 45,
 104, 187, 712, 740, 741,
 744, 746, 750–753, 755
 Brody, G., 238, 239
 Brody, J. L., 229, 230
 Brodzinsky, D., 250
 Bromet, E., 181
 Bromley, D. L., 338
 Brondino, M. J., 339, 340, 810
 Brooks, S., 27
 Brooks-Gunn, J., 245
 Brower, J., 447
 Brown, D., 680, 682, 688, 689,
 691, 692, 694
 Brown, K., 208
 Brown, M., 13
 Brown, N., 502
 Brown, T. L., 810
 Browne, A., 839
 Browne, K., 240, 805
 Brubacher, S. P., 583
 Bruck, M., 49, 566, 572, 575,
 581, 583
 Brumariu, L. E., 567
 Brunello, N., 181
 Brunia, C. H. M., 636
 Brussel, J. A., 42
 Bryan, A. I., 19
 Bryan, W. J., 668, 695, 696
 Bryant, F. B., 475
 Bryant, R. A., 453
 Rubner, S., 800
 Buchanan, A., 330, 331
 Buck, J. A., 592, 593, 595
 Buckhout, R., 532
 Buckner, F., 385, 426
 Buckner, R. L., 564
 Budd, K., 250, 251, 259, 262,
 263
 Budson, A. E., 564
 Buehler, R. E., 766
 Buffington-Vollum, J. K., 187
 Bugental, D., 243
 Bukatman, B. A., 294
 Bukstel, L. H., 766
 Bulkley, J., 589
 Bumby, K. M., 318, 849
 Buntinx, W. H. E., 210
 Burchard, B. L., 849
 Burdon, W. M., 800, 805
 Burgess, A. W., 43
 Burgess, R. L., 245
 Burgett, T., 839
 Burgwyn-Bailes, E., 567, 581
 Burl, T. G., 764
 Burl, J., 28, 111, 120, 122, 127
 Burnett, D. R., 334
 Burns, H. J., 536
 Burns, T. E., 675

- Burt, M., 850
 Burton, D. C., 813
 Burton, D. L., 852
 Burttt, H. E., 13
 Busch, S. I., 644
 Bush, J., 774
 Bush, S. S., 739
 Bushway, S. D., 821
 Bussey, K., 41
 Bussière, M. T., 845, 848
 Butcher, J. N., 89, 157, 187, 188
 Butler, B. T., 294
 Butler, J. A., 212
 Button, J., 682
 Butzin, C. A., 800
 Byng, R., 798
- Cady, H. M., 12
 Cahill, B. S., 530
 Cahill, L., 569
 Cairns, H., 13
 Calhoun, K. S., 839
 Caliso, J., 242
 Callaghan, M. G. O., 592
 Callahan, L. A., 336–338, 355, 367
 Calverley, D. S., 690
 Camilleri, J. A., 28
 Camp, N., 849
 Camp, S., 768
 Camp, S. D., 766, 800
 Campagna, V., 157
 Camparo, L. B., 597
 Campbell, J., 798
 Campbell, M. A., 401, 407, 760, 770, 771
 Campbell, R., 586
 Candel, I., 575
 Cannell, J., 697
 Cannell, J. J., 689
 Cano, E., 484, 486
 Canter, J., 43
 Cantos, A., 260
 Cantwell, B., 631, 632
 Cao, Y., 799
 Caon, A., 542
 Capaldi, D., 240, 241
 Caploe, J., 688
 Carey, K. B., 819
 Carich, M. S., 849
 Carlier, I. V. E., 453
 Carlson, C. A., 517, 540
 Carnes, P., 851
 Carol, R. N., 528
 Caron, A., 778
- Carpenter, W., 331
 Carrick, N., 570
 Carrigan, D. O., 762, 763
 Carson, D., 532
 Carson, J. L., 476
 Carulla, L. S., 197
 Carver, C. S., 851
 Casey, J. O., 800
 Cashmore, J., 151
 Casper, J. D., 479
 Cassar, D., 116
 Cassidy, J., 241
 Castellano, T. C., 765, 766
 Castelli, P., 584, 585, 594, 595
 Castellino, N. C., 850
 Cather, C., 475, 476
 Cattell, J. M., 6
 Cauffman, E., 303, 305
 Cavanaugh, J. L., 328
 Ceci, S., 49
 Ceci, S. J., 525, 562, 563, 572, 574, 576, 577, 579–581, 583
 Cecil, J., 479, 738
 Cederborg, A.-C., 564
 Cellard, A., 761
 Cerrito, B. M., 570
 Cerundolo, P., 385, 426
 Chae, Y., 563, 565, 567, 573, 580, 581
 Chaffin, M., 855
 Chaiken, S., 765
 Chambelin, N. C., 24
 Chambers, J. C., 801
 Champagne, F., 849
 Champine, P., 219, 222–224, 227
 Chan, A. S., 339, 342
 Chan, C. C., 643
 Chan, S., 844
 Chapman, G. B., 473
 Chapman, J., 407
 Chard, S., 245, 247
 Charette, Y., 406
 Charman, S. D., 532
 Chatterjee, A., 226
 Chávez, V., 406
 Chedid, E., 641
 Chen, J.-R., 764
 Chen, Z., 241
 Cheng, E. K., 648, 650
 Chermack, T. J., 415
 Childress, A. R., 643, 644
 Chiroro, P. M., 525
 Choate, P. A., 335, 338
- Choe, I., 453
 Chopra, S., 470, 472
 Choudhry, N. K., 232
 Christensen, L., 463
 Christianson, S.-Å., 564, 565, 568, 569
 Christie, J., 529
 Christopher, A. N., 850
 Christy, J., 839
 Chuang, E., 251, 252
 Cicchetti, D., 237–239, 241–243, 246, 255, 570
 Cicchetti, D. V., 840
 Cirincione, C., 333, 334, 336, 337, 341, 355, 367
 Clark, C. R., 353, 356
 Clark, H. L., 590
 Clark, J., 251, 259, 813
 Clark, J. A., 716
 Clark, J. W., 233
 Clark, L., 243, 254
 Clark, S. E., 537, 538, 540, 546, 553
 Clarke-Stewart, K. A., 582, 591–593
 Clarparedo, E., 6
 Clayton, L. D., 591
 Cleary, P. D., 233
 Cleckley, H., 27, 49
 Clegg, C., 818
 Cleland, C. M., 802
 Clemens, B., 673
 Clements, C. B., 21, 760, 768
 Clemmer, D., 760, 766
 Clermont, K. M., 481
 Clingempeel, W. G., 810
 Coates, C. A., 47
 Cobbina, J., 799
 Coccozza, J. J., 286
 Cochran, J. C., 770, 771
 Cochran, R. E., 288, 289, 291
 Cocozza, J., 389
 Cody, H. A., 232
 Coffaro, F., 342
 Cohen, E., 366
 Cohen, F., 795
 Cohen, J., 780, 823, 840
 Cohen, L., 257
 Cohen, M., 254, 777
 Cohen, M. I., 338
 Cohen, P., 823
 Cohn, N., 283, 286–288, 290, 291, 293
 Cohn, N. B., 336
 Coid, J., 402, 760, 820

- Collie, R. M., 818
 Collins, P. A., 454
 Collins, P. I., 452
 Collins, S., 283
 Collishaw, S., 242
 Collman, P., 157
 Colman, R., 259
 Colvin, M., 764, 766, 769
 Comer, S. L., 689
 Condie, D., 725
 Condie, L., 238, 245, 247–254,
 256, 258–269
 Condie, L. C., 725
 Condon, R., 666
 Conger, R., 240, 241
 Conley, M., 818
 Connell, M., 41, 49, 219,
 222–224, 227, 251, 259,
 597
 Connell, M. A., 739
 Connery, D. S., 682
 Connolly, D. A., 592, 593
 Connor, E., 157
 Connor, M., 342
 Connor, P., 814
 Connors, A. D., 819, 821
 Connors, G., 253
 Conrad, A. F., 471
 Conroy, M. A., 726
 Considine, M. G., 578
 Contreras, R., 112
 Cook, D., 813
 Cook, D. J., 812, 814
 Cook, E. F., 233
 Cook, N., 407
 Cooke, D. J., 393, 847
 Coons, P. M., 681
 Coontz, P. D., 399, 420
 Cooper, D. K., 302, 303, 306
 Cooper, S. A., 197
 Cooper, V. G., 287, 288, 291,
 303
 Copas, J., 402
 Corber, S., 588
 Cordón, I. M., 565, 571
 Cordon, I. M., 561, 570, 593
 Corey, D. M., 43, 447–451
 Cormier, C., 116, 254, 386, 392,
 394, 395, 397, 405, 408,
 426, 846
 Cormier, C. A., 392, 846, 847
 Cornelis, I., 818
 Corrigan, P. W., 807
 Corse, S., 239, 246
 Corsini, R., 19
 Cortoni, F., 808
 Cortoni, F. A., 851
 Corwyn, R., 241
 Coulter, D. L., 210
 Coulthard, R., 844
 Coulton, C., 245–247
 Couwels, J., 462
 Coviello, D., 800
 Cowden, V. L., 303
 Cox, C., 238, 239, 260
 Cox, D. N., 395, 771
 Cox, D. R., 447
 Craig, E. M., 210
 Craig, R. J., 43, 157
 Craig, W., 588
 Cramer, E., 847
 Crasilneck, H. B., 660, 661, 680,
 682, 692–694
 Crawford, M., 820
 Crayton, J., 568, 570, 581
 Creelman, C. D., 518
 Crenshaw, W., 251, 252
 Crescini, C., 588
 Crittenden, P. M., 246
 Crockenberg, S., 578
 Crocker, A. G., 406
 Crombez, G., 623, 625
 Cronbach, L. J., 215
 Cronch, L. E., 43
 Crooks, C. V., 150
 Cross, H. J., 584
 Cross, T. P., 597
 Crossman, A. M., 585
 Crotty, H. D., 316
 Crouch, J. L., 810
 Crowley, M. J., 592
 Cruise, K. R., 406
 Crutcher, N., 764
 Csikszentmihalyi, M., 852
 Ct, L. A., 662
 Cuadra, L., 406
 Cuc, A., 564
 Cullen, F. T., 759, 764–766,
 773, 777, 781, 782, 796,
 797, 804, 848, 853
 Cullen, R. M., 43
 Culver, C. M., 229, 230
 Culver, M., 584
 Cumbleton, A., 818
 Cumming, G., 760, 780, 800
 Cumming, G. R., 849
 Curwen, T., 396, 406, 811
 Custer, H., 594
 Cutajar, M., 47
 Cutler, B., 340, 341, 541
 Cutler, B. L., 480, 538
 Cutrona, C. A., 850
 Cuttler, M. J., 447
 Cyr, C., 241
 Daftary-Kapur, T., 342
 Dahl, L. M., 40
 Dahlstrom, W. G., 157, 187,
 188
 Dailey, S. B., 540
 Dale, M., 148
 Dalton, C., 149
 Daniel, A. E., 812
 Daniels, C. W., 648
 Daniels, S., 470, 471
 Danielson, K. D., 634–636, 638
 Dann, B. M., 479, 483, 486, 488,
 496
 Dannenfelser, P. L., 597
 Dansereau, D. F., 800
 Dantzker, M. L., 449
 Danzig, E. R., 12
 Dark, F., 820
 Darke, J., 853
 Darley, D., 341, 343
 Darley, J., 341–343
 Darley, J. G., 19
 Darley, J. M., 473, 474, 477
 Darling, S., 527
 Daston, L., 518
 Dattilio, F. M., 747
 Davidshofer, C. O., 844
 Davidson, W., 759
 Davidson, W. S., 777
 Davies, G., 531, 532, 536, 549
 Davies, G. M., 480
 Davis, C., 643
 Davis, D., 697
 Davis, D. L., 289
 Davis, J. P., 545
 Davis, K., 812
 Davis, M. R., 48
 Davis, R., 157
 Davis, R. D., 449
 Davis, R. E., 545
 Davis, S., 282
 Davis, S. L., 568, 570, 572, 575,
 581, 590
 Davison, K. S., 597
 Dawes, R. M., 394
 Day, A., 800, 801
 Deane, F. P., 798
 DeAngelis, T., 456
 de Beaumont, G., 761, 762
 De Bellis, M., 241, 242

- De Clercq, A., 623
 de Clercq, B., 818
 Dee, J., 480
 Deffenbacher, K. A., 525, 528,
 536, 564, 565, 586
 de Grazia, E., 294
 de Groh, M., 689
 Deisinger, G., 457
 DeJong, C., 764
 deJong, M., 527
 deLaosa, M., 297
 DeLeon, P., 283
 Delgado, J., 120
 DeLisa, A., 340, 480
 DeLisi, M., 766
 Delprino, R. P., 454, 462
 DeMatteo, D., 28, 111, 120,
 122, 127, 288, 810
 Dempsey, J., 588
 Denkhaus, H., 809
 Denney, R. L., 739
 Denno, D. W., 353
 Denson, T. F., 528
 Dent, H. R., 562
 DePaulo, B. M., 594
 Derogatis, L. R., 400
 de Ruiters, C., 258, 406, 407
 Desmarais, S. L., 50, 396, 406,
 407
 de Tocqueville, A., 761, 762
 Deutsch, R., 47
 Devine, D. J., 472, 591
 Devine, P. G., 517, 533, 537
 Devlin, Patrick, 538
 de Vogel, V., 406, 407
 de Vries Robbé, M., 406
 DeWitt, J., 392
 Deyoub, P. L., 666
 Dhaliwal, G. K., 799
 Dhanens, T. P., 690
 Diamond, B., 361
 Diamond, B. L., 317, 318, 670,
 682, 689, 691
 Diamond, P. M., 809
 Diamond, S., 479, 485, 489,
 490, 499, 500
 Diamond, S. S., 339, 479, 485,
 490, 491, 591
 Diaz, P., 148
 Dibble, A. E., 334
 Dickens, C., 762
 Dickenson, J. J., 593
 Dickey, W., 294
 Dickinson, J. J., 536, 538
 Dickover, R. M., 766
 DiClemente, C. C., 819
 Dietrich, H. L., 541
 Dietz, E. F., 21
 Di Fazio, R., 818
 Diggett, C., 668
 Dillow, S. A., 210, 212
 Dilorenzo, T. M., 764
 Dinges, D. F., 677, 680, 685
 Di Placido, C., 809, 819
 Dirks-Linhorst, P. A., 338
 Dishion, T. J., 778
 Di Vasto, P. V., 839
 Dix, D., 762
 Dix, G. E., 356, 359, 360, 363,
 386
 Dixon, B. G., 805, 808
 Dixon, L., 240–242, 253, 737
 Djanogly, S. E., 186
 Dlugacz, H. A., 219, 222–224,
 227
 D'Mello, M., 576, 579
 Dobbs, S., 240, 241, 254
 Dodd, E., 585
 Dodd, J., 593
 Dodge, K., 246, 254, 256
 Doi Fick, L., 143
 Dolan, M. C., 407
 Doll, H., 403, 404
 Dollar, K. M., 590
 Donchin, E., 627, 631–640
 Donnell, D. C., 233
 Donnelly, D. A., 590
 Doob, A. N., 524, 532, 533
 Dorado, J. S., 585
 Doreleijers, T., 258
 Doreleijers, Th. A. H., 407
 Doren, D. M., 842
 Dorken, S. K., 796, 806, 807
 Dougherty, C., 338
 Douglas, J. E., 43
 Douglas, K. S., 50, 171,
 179–181, 187, 385, 388,
 394–397, 403, 406–409,
 411, 413, 416, 426, 760,
 771, 817
 Douglas-Brown, R., 574
 Dowden, C., 848, 849
 Dowling, F., 451
 Drake, B., 242, 247, 255
 Drake, E., 777
 Drake, R. E., 807, 819
 Drapeau, M., 854
 Draper, P., 245
 Dressler, J., 665
 Dreyfuss, D. A., 663
 Drogin, E. Y., 225, 227, 228
 Drozd, L., 149
 D'Souza, D. C., 528
 Duan, N., 806
 Duara, R., 226
 Dubin, L., 692
 Dubowitz, H., 238–240, 258,
 260
 Dudgeon, P., 580, 581
 Dukes, C., 574
 Dulcan, M. K., 811
 Dumas, J., 241
 DuMaurier, G., 663
 Duncan, G., 245
 Dunford, B. B., 591
 Dunn, C., 304
 Dunn, L. A., 636
 Dunn, M. A., 738
 Dunning, D., 525, 542
 Durham, R., 475, 476
 Duwe, G., 847
 Dvoskin, J., 797, 798
 Dvoskin, J. A., 113, 176, 177,
 179, 180, 723, 760, 795
 Dwight, T. W., 763
 Dyer, F. J., 157
 Dysart, J., 540
 Dysart, J. E., 529
 Dywan, J., 682, 690, 691

 Eagley, A., 765
 Eakin, D. K., 572
 Earls, C. M., 853
 Eaves, D., 296, 297, 334, 335,
 337, 394–396, 403, 411
 Eaves, S. D., 771
 Ebbinghaus, H., 529
 Eberhardt, N. L., 646
 Eccelstone, C. E. J., 778
 Eccles, A., 843, 845
 Eccleston, L., 801
 Echterling, L. G., 662
 Eckes, S. E., 208
 Edelstein, B., 233
 Edelstein, R. S., 565, 567–571,
 573, 583, 594
 Edens, J. F., 156, 297–299
 Edgerton, R. B., 301
 Efron, B., 633
 Egeland, B., 240, 242
 Ehlers, A., 453
 Eimer, B. N., 662
 Eisen, M. L., 568, 570, 581, 584
 Eisenberg, M., 281, 299

- Eisenberg, T., 454, 471, 472, 476, 481
 Eisner, S. L., 493, 495
 Ekman, P., 594
 Elaad, E., 627, 628
 Ellenberger, H. F., 667
 Ellerby, L., 849
 Elliott, D. M., 839
 Elliott, I., 43
 Elliott, S. N., 203
 Ellis, H. D., 527
 Ellis, L., 485, 499, 500
 Ellison, K. W., 532
 Ellman, I., 151
 Ellsworth, P. C., 479
 Ellwanger, J. W., 636
 Elsass, P., 407
 Elwork, A., 289, 471
 Emery, R., 147
 Emery, R. E., 46, 139
 Emmanuel, E., 231
 Emmanuel, L. L., 231
 Emmerling, D. A., 662
 Empey, D. D., 799
 Emshoff, J. G., 777
 Endler, N. S., 851
 Enebrink, P., 407
 England, P., 593
 English, L., 186
 English, P., 471
 English, P. W., 115
 Ennis, B. J., 419
 Eno Loudon, J., 341
 Enquist, A., 76
 Epperson, D. L., 392
 Epstein, M. A., 590
 Erard, R. E., 157
 Erb, A. L., 96
 Erchull, M. J., 823
 Erdahl, A., 761
 Erdelyi, M. H., 531, 690
 Erickson, M. H., 662
 Erickson, S. K., 338
 Erlinder, C. P., 359
 Ermann, M. D., 478
 Eronen, M., 807
 Ertem, I. O., 240, 241, 254
 Esplin, P., 154
 Esplin, P. W., 569
 Essock, S. M., 819
 Estupinan-Kane, L., 289
 Etchells, E., 232
 Evans, D., 853
 Evans, D. C., 524
 Evans, D. L., 473, 476, 480
 Evans, F. B., 157
 Evans, J., 341
 Evans, J. R., 527, 528
 Evens, M., 818
 Everington, C., 301, 302, 304
 Everington, C. T., 302–304
 Everson, M., 256, 268, 269
 Ewin, D., 663
 Ewing, C. P., 61, 328, 336
 Ewles, C. D., 802
 Eysenck, H. J., 26
 Fabiano, E. A., 802
 Fabiano, R., 802
 Fabricius, B., 151
 Fabricius, W., 148
 Fagan, J., 821
 Fagan, J. A., 454
 Fagan, T. J., 760, 768, 781, 796, 811–814
 Fagan, T. K., 197
 Faigman, D., 385, 420
 Faigman, D. L., 648, 650, 678
 Falkin, G. P., 799, 805
 Faller, K., 256, 268, 269
 Fang, F., 642
 Farkas, G., 283
 Farley, E. J., 479
 Farrar, M. J., 569
 Farrigan, T., 812
 Farrington, D. P., 777, 803
 Farrow, T. F. D., 643
 Farwell, L. A., 627, 631–640, 642
 Favole, R. J., 282
 Fawcett, J. M., 529
 Fazel, S., 403–405, 427
 Feehan, J. W., 480
 Feemster, S. L., 456
 Felitti, V. J., 570
 Felix, E., 250
 Felthous, A. R., 353, 385, 426
 Feng, C. M., 643
 Ferguson, M., 326, 841
 Ferguson, R. J., 620
 Fernandez, Y. M., 840, 844–846, 854, 855
 Feshbach, N., 241, 243
 Feucht-Haviar, T., 281
 Fico, J. M., 448
 Fiedler, K., 625, 650
 Field, C. E., 778
 Figueredo, A. J., 799
 Filone, S., 111, 120, 127
 Fincham, F., 341, 343, 344
 Findley, J. D., 117
 Finkel, N. J., 335, 337, 339, 342, 343
 Finkelhor, D., 839
 Finkelstein, S., 662
 Finn, P., 451, 453–455, 459, 764
 Firestone, M., 385, 426
 Firestone, P., 841, 851
 First, M., 331
 Fischhoff, B., 475
 Fischler, G. L., 450
 Fischler, I. S., 642
 Fishbein, M., 765
 Fisher, B. S., 766
 Fisher, D., 805
 Fisher, M. A., 106, 445
 Fisher, R. P., 530, 646
 Fisher, W., 806
 Fisher, W. H., 286, 806
 Fiske, S. T., 340
 Fite, R., 681, 685, 689, 692
 Fitzgerald, J. F., 286
 Fivush, R., 567
 Flanagan, C. L., 451
 Flanagan, T., 767
 Flango, V. E., 493, 494, 497, 498
 Flannagan, D. A., 569
 Flannery, D., 769
 Flaum, M., 331, 795
 Flens, J., 157
 Fletcher, K., 283
 Flora, D. B., 772, 796, 801, 804, 806
 Flowe, H. D., 588
 Foa, E. B., 181
 Foa, E. G., 571
 Fogel, M. H., 338
 Foley, M. A., 563
 Folkman, S., 767
 Follette, W. C., 697, 823
 Folstein, M. F., 232
 Folstein, S. E., 232
 Ford, H. J., 839
 Ford, M. C., 479
 Fordham, A. S., 854
 Forrest, T. J., 574
 Forrester, D., 646
 ForsterLee, L., 479, 493–498, 501–503, 590
 ForsterLee, R., 590
 Forth, A., 396, 402
 Forth, A. E., 402, 818
 Fortune, Z., 820
 Foster, E., 111, 120, 127
 Foster, S., 767

- Fowers, D., 392
 Fox, K. C., 814
 Foy, J. L., 294
 Fradin, S., 580
 Frank, J., 470
 Fraser, J., 530
 Fraser, S., 527
 Fraynt, B. R., 619
 Freedman, N. L., 778
 Freeman, A., 804
 Freeman, N., 22
 Freese, R., 807
 Fremouw, W., 818
 French, S., 401, 759, 760, 766,
 770–772, 781
 French, S. A., 849
 Frenda, S., 257
 Freske, P. J., 847
 Freund, K., 844
 Frierson, R. L., 341
 Friman, P. C., 778
 Frischholz, E., 660, 661, 680,
 682, 692–694
 Frischholz, E. J., 663, 668, 679,
 688, 691
 Fromuth, M. E., 590
 Frowd, C. D., 532
 Frueh, B. C., 774, 814
 Fryman, H. M., 590
 Fujii, T., 643
 Fulero, S., 302, 322, 341, 426,
 540
 Fulero, S. M., 538
 Fulton, J. P., 322
 Furedy, J. J., 643
 Furniss, J. M., 766
- Gabora, N. J., 592
 Gaborit, M., 474
 Gabrieli, J. D. E., 564
 Gaes, G., 768
 Gaes, G. G., 766
 Galanter, M., 476
 Galiotta, M., 342
 Galynker, I., 257
 Gamer, M., 644, 646
 Ganis, G., 619, 643
 Gao, J. H., 643
 Garbarino, J., 247
 Garber, S., 471
 Garbolino, J., 140, 141
 Garcia, J. A., 156
 Garcia, T., 528
 Garcia-Mansilla, A., 427
 Gardner, D. S., 686
- Gardner, K., 289
 Gardner, M., 372
 Gardner, R., 149
 Gardner, W., 398–400
 Garety, P., 330–332
 Garg, R., 476
 Garofalo, R. F., 386–388, 408,
 411
 Garrett, B. L., 517
 Garrett, C. J., 759
 Garrido, S., 457
 Garrity, T. F., 798
 Garry, M., 536
 Garven, S., 578
 Garver, R. B., 660, 661, 680,
 682, 692–694
 Gasser, D., 338
 Gaudin, J., 255, 258, 264
 Ge, L., 587
 Geddie, L., 580
 Geer, T. M., 800
 Gelber, C., 462
 Gelfand, D., 243
 Gelles, M. G., 457, 458
 Gelles, R., 238
 Gendreau, P., 43, 48, 401, 759,
 760, 764–774, 777–782,
 796, 797, 804, 805, 815,
 817, 848, 849, 853
 Genet, B., 451
 Gensheimer, L., 759
 Gentz, D., 453
 Genuis, M., 839
 George, C., 148, 268
 George, M. S., 644, 645
 Gerard, A., 158
 Gerend, M. A., 823
 Gershon, R., 455
 Gersons, B. P. R., 453
 Gert, B., 229, 230
 Ghetti, S., 565, 570, 571, 574,
 575
 Giannelli, P. C., 678, 684
 Giardini, G. I., 19
 Gibas, A. L., 428
 Gibbons, F., 238, 239
 Gibbons, J. J., 768
 Gibbs, A. C., 454
 Gibbs, B., 244
 Gibling, F., 527
 Gibson, M., 767
 Gigerenzer, G., 518
 Gignac, G. E., 581
 Gill, B., 737
 Gillard, N. D., 188
- Gilstrap, L. L., 577
 Ginton, A., 628
 Gionet, A., 849
 Girard, L., 817
 Gittelman, B., 226
 Glaser, D., 21, 770
 Glasmeier, A. K., 812
 Glasser, M., 419
 Glatt, R. L., 689
 Glaze, L. E., 795, 797
 Glisky, E., 569
 Glueck, E. T., 762
 Glueck, S., 762
 Goddard, H. H., 3
 Godinet, M., 239
 Goerdts, J., 476
 Goggin, C., 48, 759, 760, 764,
 765, 767–770, 777, 781,
 815
 Goldberg, T., 663
 Goldfinger, K., 726
 Golding, G., 341, 343, 344
 Golding, J. M., 589–591
 Golding, S., 254, 256, 337,
 341–344
 Golding, S. L., 92, 282, 283,
 286–291, 293–296, 316,
 318, 328, 329, 333–337,
 341, 343, 747
 Goldkamp, K. S., 474
 Goldman, J., 259
 Goldman, R., 392
 Goldstein, A., 773
 Goldstein, A. M., 58–61, 111,
 712, 718
 Goldstein, A. S., 321
 Goldstein, L., 570
 Goldstein, R., 238
 Goldstein, R. H., 688
 Goldstein, R. L., 290
 Goldwyn, R., 242
 Gollan, J. K., 112
 Gollop, M., 151
 Gomar-Sones, J., 406
 Goncalves, R. A., 407
 Goode, P. S., 226
 Goodman, G. S., 49, 561–573,
 575, 576, 578, 579, 581,
 583–585, 587, 589, 590,
 593–596
 Goodman, J., 473, 480
 Goodman, S., 243, 254
 Goodman-Brown, T. B., 583
 Goodsell, C. A., 517, 540
 Goodstein, L., 767

- Goodstein, L. D., 107
 Goodwin, R. C., 286
 Gordon, A., 802, 808, 809, 817, 855, 856
 Gordon, A. E., 392
 Gordon, B. N., 567, 581
 Gordon, D. A., 799
 Gordon, D. S., 583
 Gordon, H. M., 585
 Gordon, J. R., 773
 Gore, D. K., 849
 Gore, K. S., 392
 Gorny, I., 102
 Gosling, S. D., 823
 Gossop, M., 805
 Gottdiener, W. H., 286, 287, 290, 306, 307
 Gottesman, J., 24, 25
 Gottfried, A. E., 157
 Gottfried, A. W., 157
 Gottlieb, M., 799
 Gottschalk, J. A., 726
 Gottschalk, R., 759
 Gould, J., 137, 148
 Gould, J. W., 143
 Gould-Saltman, D. J., 143
 Gowensmith, W. N., 288, 290, 291, 336
 Grace, R. C., 817, 818
 Graesser, A. C., 592
 Graham, J. R., 157, 187, 188
 Graham, M., 517
 Graham, S., 303, 305
 Grandjean, N., 299, 300
 Granhag, P. A., 564, 594, 595
 Grann, M., 394, 403–405, 427
 Grant, B. A., 767
 Grassian, S., 778, 779
 Gravitz, M. A., 659–661, 667, 680, 682, 692–694
 Gray, A. L., 819, 821
 Gray, E., 592
 Gray, J., 592
 Gray, S., 317
 Gray, S. R., 800
 Greathouse, S. M., 538
 Greeley, H. T., 645
 Greely, H. T., 645
 Green, E., 476
 Green, M. J., 231
 Greenberg, J., 14
 Greenberg, L. R., 143
 Greenberg, L. S., 820
 Greenberg, S., 92, 119
 Greenberg, S. A., 95, 171, 173, 176, 177, 179, 182, 187, 712, 723
 Greene, E., 470, 472–476, 478, 480, 527
 Greene, L. W., 453, 462
 Greene, R. L., 101, 721
 Greenfield, D. P., 726
 Greenhoot, A. F., 569
 Greer, A., 286
 Greif, G. L., 245, 254, 260
 Grenesko, E. L., 644, 645
 Gresham, A. W., 592
 Gresham, F. M., 208
 Gries, L., 260
 Griffin, P. A., 41, 338
 Griffith, H. R., 226
 Griffith, J. D., 593
 Grimbeck, E. J., 41
 Grisso, J. T., 4, 5, 13, 15, 16
 Grisso, T., 39, 49, 58–61, 111, 130, 179, 220–222, 231–233, 281, 283, 286, 287, 294, 302–306, 392, 712, 716, 718
 Gronlund, S. D., 517, 540
 Groscup, J., 343
 Groscup, J. L., 221, 342, 408, 428, 737
 Gross, H., 26
 Gross, J. J., 851
 Gross, N. R., 777, 797
 Gross, S. R., 517
 Grossman, S., 157
 Groth-Marnat, G., 721
 Grove, W., 128
 Grove, W. M., 390, 391
 Grover, T., 46
 Grubin, D., 330, 331, 803, 846
 Grundmann, D., 256
 Gruzelier, J., 662
 Gu, D., 808, 809, 819
 Guarnaccia, C. A., 298
 Gudjonsson, G. H., 675, 682
 Guevremont, D. C., 762
 Guilbault, R. L., 475
 Guilford, J. P., 762
 Guilmette, T., 289
 Gumpert, C. H., 407
 Gunnin, D. D., 338
 Gunstad, J., 102
 Gur, R. C., 643, 644
 Gurevich, M., 643
 Gurrera, R. J., 233
 Guterman, A., 226
 Gutheil, T., 426
 Gutheil, T. G., 104, 334, 417, 419, 689, 694, 697
 Guthrie, C., 481
 Guy, L. S., 391, 403–405, 407
 Guzzo, L., 767
 Gwaltney, L., 805
 Gwynn, M. I., 689
 Haas, A., 42
 Haberman, M. A., 662
 Haberstick, B., 244
 Haecerich, M., 590
 Hafemeister, T. L., 113
 Häfner, H., 330
 Hagstrom, A., 842
 Hahlweg, K., 823
 Hale, M., 8, 10
 Hall, D., 337
 Hall, E. A., 799
 Hall, H. V., 387, 419
 Hall, S. R., 113, 593, 595
 Ham, L. S., 528
 Hamberger, L. K., 799
 Hambridge, J. A., 800
 Hamilton-Giachritsis, C., 240–242, 253
 Hamilton-Giachritsis, C. E., 854
 Hammelmann, H. A., 537
 Hammond, D. C., 660, 661, 669, 680, 682, 689, 691–694
 Hampton, R., 240
 Hancock, P. J., 532
 Handel, R. W., 187
 Handel, S., 342, 343
 Handwerk, M. L., 778
 Haney, C., 36, 37, 342, 768, 778
 Hanley, D., 769
 Hannaford, P., 499, 500
 Hannaford-Agor, P. L., 481, 482
 Hanoteau, F., 589
 Hans, V., 499, 500
 Hans, V. P., 470, 473, 478–481, 499
 Hansen, D. J., 43
 Hanson, R. K., 392–395, 401, 404, 405, 760, 802, 817, 820, 845–853, 855, 856
 Hanson, T., 256, 257
 Harding, E. C., 663
 Hardt, J., 240, 241, 264
 Hare, R., 49, 401

- Hare, R. D., 27, 288, 291, 402, 765, 771
 Harel, A., 471
 Haring, M., 188, 189
 Harker, D. L., 538
 Harlow, C. W., 803
 Haroun, A. M., 291
 Harpold, J. A., 456
 Harrell, L., 226
 Harrell, L. E., 226, 232
 Harris, A. J. R., 802, 820, 855, 856
 Harris, E., 711
 Harris, G., 116, 254
 Harris, G. T., 386, 392, 394, 395, 397, 405, 408, 426, 846, 847, 855
 Harris, L., 338
 Harris, L. S., 563, 571
 Harris, M., 328
 Harris, R., 664, 667
 Harrow, M., 331
 Hart, A. J., 471, 473, 476, 480, 481
 Hart, S. D., 288, 291, 393–396, 403, 406–409, 411, 413, 414, 416, 428, 771, 847, 848
 Hartman, C. R., 43
 Hartwig, M., 594, 595
 Harvey, V. S., 726
 Harwood, T., 721
 Hashtroudi, S., 572, 574
 Hassin, R., 524
 Hassiotis, A., 808
 Hastie, R., 44, 340, 471, 473, 475, 481, 493, 497, 498
 Hatfield, E., 715
 Hautman, L., 818
 Haw, R. M., 536
 Hawk, G. L., 813
 Hawkins, S. A., 475
 Hayaski, R., 805
 Hayne, H., 567
 Hays, V., 453
 Healy, D., 818
 Hebert, K., 227
 Hebert, K. S., 42
 Hebert, T., 227
 Heiblum, N., 856
 Heilbrun, K., 4, 27, 28, 57–61, 185, 187–189, 283, 288, 338, 712, 733, 734, 739
 Heilbrun, K. S., 407, 408
 Heiman, M., 268
 Heise, M., 471, 472, 481
 Helmus, L., 28, 760, 820, 853
 Hemmati, T., 818, 821
 Hemsley, M., 330, 332
 Henaghan, R. M., 151
 Henderson, C. E., 287, 290
 Hendry, M., 403
 Henggeler, S. W., 810
 Henning, K., 818
 Henning, K. R., 774
 Hepburn, J. R., 799
 Hepps, D., 563
 Herbel, B. L., 288, 289, 291
 Herbert, K. S., 479, 481
 Herbert, M., 241, 265
 Herford, A. E., 643
 Herlitz, A., 526
 Hermann, D., 330
 Hernandez, R. S., 636
 Heron, W., 780
 Herrenkohl, E., 240
 Herrenkohl, R., 240, 243
 Herrenkohl, T., 243
 Herrick, S., 813
 Herring, D., 240, 245
 Herrnstein, R. J., 803
 Hersch, J., 481
 Hershkowitz, I., 154, 266, 267, 584, 594
 Herve, H. F. M., 528
 Heskin, K. J., 767
 Hess, A. K., 105
 Hess, D., 798
 Hess, T. M., 569
 Hetherington, E., 147
 Heuer, L., 481, 485–491, 493–498
 Hewitt, J., 244, 289
 Heyne, L. K., 96
 Hibbard, W. S., 668
 Hibler, N. S., 448, 458, 669
 Hicken, B., 233
 Hickman, L. J., 802
 Hicks, S. J., 43
 Hicks, T., 294
 Hilgard, E. R., 666, 682
 Hilgard, J. R., 682
 Hiller, M. L., 42, 798–800
 Hilliar, K. F., 528
 Hillyard, S. A., 642
 Hilton, N. Z., 392, 405
 Hinsz, V. B., 473
 Hira, S., 639
 Hirschi, T., 766
 Hirschman, J. E., 563
 Hirst, W., 564
 Hithwaite, J., 227
 Ho, R., 590
 Hobbs, S. D., 573
 Hodge, D., 575
 Hodges, W. F., 148
 Hodgins, S., 807
 Hodgson, S., 760, 853
 Hoge, R., 48
 Hoge, R. D., 392, 402, 759, 773, 777, 796, 797, 804, 848, 853
 Hoge, S., 50
 Hoge, S. K., 281, 297, 298
 Hoke, S. E., 800
 Holdford, R., 818
 Holdsworth, M., 250, 262, 263
 Holland, C. L., 845, 849
 Holliday, R. E., 588
 Hollin, C. R., 817
 Hollinger, L. A., 473
 Hollon, S. D., 764
 Holloway, J., 43
 Holohean, E. J., 797, 798
 Holsinger, A. M., 817
 Holt, A. R., 590
 Holt, R. R., 390
 Holton, J., 237
 Holton, S. M. B., 811, 812
 Homant, R. J., 452
 Honig, A. L., 455
 Honts, C. R., 618, 622, 624, 626, 628, 629
 Hooijschuur, A., 407
 Hooks, G., 812
 Hooley, J. M., 781
 Hopfer, C., 244
 Horowitz, D., 266, 267, 594
 Horowitz, I., 475, 493–498, 501–503
 Horowitz, I. A., 13, 479, 590
 Horowitz, S. W., 628
 Horton, M. L., 301
 Horvath, A. O., 820
 Horvath, F., 618
 Hosp, J. L., 208, 209
 Hough, L. M., 449
 Houlding, C., 407
 Howard, J., 761
 Howard, P. D., 394
 Howe, M. L., 562, 563, 565, 569, 570
 Howells, K., 800, 801, 850
 Howerton, A., 798
 Howes, P., 239, 241, 243, 246

- Howgego, I. M., 820
 Hubbard, K. L., 288, 291
 Huckler, S. J., 843, 852
 Hudson, J. I., 697
 Hudson, S. M., 851
 Huebner, B. M., 799
 Huening, J., 806
 Huffman, J., 845
 Hug, J. J., 767
 Hugenberg, K., 525
 Hughes, M., 181
 Humphreys, L., 229
 Humphries, J. E., 588
 Hunt, E., 810
 Hunt, M., 42
 Hunter, J. A., 799
 Hunter, J. E., 780
 Huot, S. J., 392
 Hupp, R. T., 452
 Hurling, T., 812
 Hurwitz, D., 226
 Huss, M., 35
 Huss, M. T., 112, 221, 737
 Hutchins, R. M., 5, 12
 Hyman, S. E., 645
- Iacono, W. G., 615, 618,
 622–625, 634–636, 638,
 649–651, 845
- Ibabe, I., 568
 Illes, J., 645
 Immarigeon, R., 779
 Imwinkelried, E., 584, 589, 590
 Indahl, K. E., 473
 Ingram, K. K., 232
 Innes, C. A., 781
 Inwald, R., 448
 Inwald, R. E., 449
 Inwinkelried, E. J., 678
 Ireson, R., 805
 Iselin, A., 810
 Itoh, M., 643
 Ivanov, S., 632
 Ives, G., 760–762
- Jackson, J. P., 15, 16
 Jackson, K. J., 848
 Jackson, M. A., 282
 Jackson, R., 839
 Jackson, R. L., 300
 Jackson, S., 578
 Jackson, S. L., 597
 Jacobs, C., 336
 Jacobs, S. C., 452
 Jacobson, L., 537
- Jacobson, N. S., 823
 Jacobvitz, D., 240, 242
 Jacoby, J. E., 389
 Jaffe, P. G., 150
 Jaffee, S., 240, 241
 Jaman, D. R., 766
 James, D. J., 797
 James, K. A., 541
 Janca, A., 715
 Janofsky, J. S., 232, 336
 Janus, E. S., 423
 Jarrold, C., 582
 Jaspers, K., 332
 Jaudes, P., 244
 Jauncey, S., 800
 Jay Rhee, S., 244
 Jefferson, S. D., 777
 Jeffrey, R., 341
 Jeffries, J. C., 478
 Jeglic, E. L., 840
 Jellicoe, D., 799
 Jemelka, R., 770
 Jenkins, F., 532, 536, 549
 Jensen, G., 297
 Jenson, F. A. S., 289
 Jesness, C., 21
 Jewkes, R., 47
 Joab, A., 592
 Joffe, S., 233
 Johansson-Love, J., 818
 John, O. P., 823
 Johns, M., 474, 476
 Johnson, C., 342
 Johnson, D. M., 714
 Johnson, J. L., 575
 Johnson, K. A., 644, 645
 Johnson, M., 632, 778
 Johnson, M. K., 563, 564, 572,
 574
 Johnson, M. M., 632
 Johnson, M. P., 149, 150
 Johnson, M. T., 738
 Johnson, R., 643
 Johnson, W. G., 297
 Johnston, J., 259
 Johnston, J. R., 46, 149, 150
 Johnston, N., 762
 Jones, B. L., 302
 Jones, D. P., 570
 Jones, D. P. H., 563, 583, 593
 Jones, E. E., 650
 Jones, L. M., 597
 Jones, M. N., 395, 803
 Jones, O. D., 645
 Jonson, C. L., 765, 766, 769
- Jonson-Reid, M., 242, 247, 255
 Jordan, C. E., 47
 Josephs, R. A., 528
 Jungman, N., 628
 Jurden, F. H., 589
- Kachigian, C., 385, 426
 Kaemmer, B., 157, 187, 188
 Kahneman, D., 471, 473, 477,
 478
 Kalven, H., 481
 Kamena, M. D., 453
 Kamin, K. A., 475
 Kane, A., 176, 177, 179, 180
 Kane, A. W., 723
 Kanwisher, N., 645
 Kaplan, H., 241
 Kaplan, M. F., 45
 Kaplan, N., 241
 Karel, M. J., 233
 Karibian, D., 574
 Karlin, R. A., 671, 678, 679, 692
 Karlovac, M., 474, 477
 Karmen, A., 47
 Karson, M., 726
 Kashgarian, M., 286, 295
 Kastanakis, J. A., 561
 Katsavdakias, K., 257
 Katz, C., 584
 Katz, M. R., 232
 Katzenbach, N. B., 768
 Kaufman, L. R., 839
 Kaul, J. D., 392
 Kavale, K. A., 214
 Kaye, D., 420
 Kaye, D. H., 479, 678
 Kazantzis, N., 804
 Kazdin, A. E., 799, 822
 Kazemian, L., 821
 Keedy, E. R., 320
 Keeling, J. A., 819
 Keeney, J. M., 589, 595
 Keeney, K. S., 561
 Keeton, P., 474, 475
 Keilitz, I., 322
 Keith-Spiegel, P. C., 88
 Keller, E. I., 17
 Kelley-Poulos, S., 163
 Kelly, J., 147
 Kelly, J. B., 147, 149, 150
 Kelly, J. D., 148
 Kelly, R. K., 143
 Kelso, J. C., 482
 Kelves, D. J., 18
 Kemp, R. I., 528

- Kendler, K., 331
 Kennedy, C., 454
 Kennedy, D. B., 452
 Kennedy, P., 842, 843
 Kensingler, E. A., 565
 Kent, L., 503
 Kenyon, S., 590
 Kerns, K. A., 567
 Kesic, D., 43
 Kessler, R. C., 181
 Kettler, R., 203
 Kettler, R. J., 203
 Keyes, D., 760, 769, 772
 Kicklighter, R. H., 198
 Kiernan, B., 574
 Kihlstrom, J. F., 571, 659, 687
 Killias, M., 765
 Kilmann, P. R., 766
 Kilpatrick, D. G., 47
 Kim, D., 259
 Kim, J., 237, 238, 242, 243, 255, 570
 Kim, K., 240, 241, 258
 Kimball, P., 768
 King, E. M., 471
 King, R. S., 812
 Kingston, D. A., 841, 850
 Kingstone, A., 528
 Kinscherff, R., 250, 254, 261–263, 265, 269
 Kinsella, C., 813
 Kircher, J., 622, 624, 626, 629
 Kircher, J. C., 628
 Kirkland, K., 47, 139, 164
 Kirkland, K. L., 139, 164
 Kirsch, I., 681, 685, 688–690, 692
 Kirschman, E., 457
 Kirschner, S. M., 386
 Kirshenbaum, H. M., 524, 532, 533
 Kitaeff, J., 446
 Kitchener, K. S., 88
 Klaver, J., 396, 411, 416, 847
 Klebe, K. J., 778–780
 Kleinhauz, M., 663
 Klemfuss, J. Z., 582
 Klettke, B., 592
 Kluckhohn, C., 413
 Kluger, R., 15
 Knap, S. J., 175
 Knapp, R. G., 814
 Knapp, S., 385, 426, 711
 Knapp, S. J., 88, 95
 Knight, A., 226, 227
 Knight, E., 238, 239, 260
 Knight, K., 799, 800
 Knox, C., 778
 Knox, P. L., 800
 Koch, W. J., 171, 179–181, 187–189
 Koegl, C. J., 396, 406
 Kohl, P., 242, 247, 255
 Kola, L., 819
 Kolva, E. A., 221, 224
 Kondrat, D., 338
 Koocher, G., 265
 Koocher, G. P., 88
 Korbin, J., 245–247
 Kose, S., 645
 Kosslyn, S. M., 643
 Kosson, D. S., 402, 623
 Koster, E. H., 623
 Kotch, J., 238, 239, 260
 Kovera, M. B., 42, 44, 470, 472, 479, 481, 538, 592, 594–596, 738
 Kowaz, A., 334, 335, 337
 Kozak, M. J., 571
 Kozel, F. A., 644, 645
 Kozinski, A., 470
 Kozol, H. L., 386–388, 408, 411
 Krafka, C., 738
 Kraus, F. R., 367
 Krauss, D., 111, 116, 117, 122, 128, 130, 800
 Krauss, D. A., 28
 Kravitz, J., 517, 525
 Kremnitzer, M., 650
 Kressel, D. F., 45
 Kressel, N. J., 45
 Kriegman, D., 856
 Kroger, W., 666
 Kroner, D. G., 395, 796, 801, 803, 804, 806, 807, 815, 817–819, 821
 Kronstrand, R., 807
 Kropp, P. R., 395, 396, 403, 406, 407, 411, 416, 428, 847
 Krüger, L., 518
 Krüger, T., 571
 Krugman, R. D., 563
 Kruh, I., 303
 Kruhl, J., 448
 Krupp, B. H., 289
 Krzywanski, S., 226
 Kuczynski, L., 241, 246, 254
 Kuehnle, K., 41, 46, 49
 Kugler, K. E., 287, 290, 295
 Kugler, T., 471
 Kuhn, J., 566, 569, 571, 576
 Kuhns, B. W., 668
 Kuhns, J. B., 772
 Kulkofsky, S., 562, 582
 Kuna, D. P., 11
 Kupers, T. A., 778, 779
 Kurus, S., 338
 Kurz, A., 203
 Kutas, M., 642
 Kutzer, D., 417
 Kwartner, P. P., 592
 Kyle, L. L., 338
 Lab, S. P., 777
 Laben, J. K., 286, 295
 LaBine, G., 475
 LaBine, S. J., 475
 Labkovsky, E., 641
 LaDuke, C., 810
 LaFave, W., 316, 358
 LaFountain, N., 481
 Laken, S. J., 644, 645
 Lally, S., 187
 Lalumière, M. L., 844, 855
 Lamb, H. R., 806
 Lamb, M., 244, 246, 266, 267
 Lamb, M. E., 147, 148, 154, 564, 569, 572, 583, 584, 594
 Lamberti, J. S., 338
 Lammers, S., 406
 Lampropoulos, G. K., 804
 Land, W. B., 852
 Landenberger, N., 802, 803
 Landes, W. M., 471
 Landsman, S., 481
 Landström, S., 594–596
 Landy, F. J., 10
 Lang, C., 392
 Langan, N., 768
 Langleben, D. D., 643, 644
 Langley, R. L., 715
 Laravuso, J. J., 636
 LaRooy, D. J., 564, 572
 Larsen, D., 812
 Larsen, R. J., 851
 Larsen, S., 492
 Latendresse, M., 808
 Latessa, E. J., 766, 767, 774, 781, 817
 Latter, R., 649
 Laub, J., 821
 Laub, J. H., 821
 Laureillo, J., 229, 230
 Laurell, O., 564

- Laurence, J.-R., 659, 663,
 677–679, 691
 Law, M., 401
 Law, M. A., 768, 770
 Lawler, M. J., 561, 593
 Lawrence, K., 586
 Lawrence, L., 406
 Laws, D. R., 394, 396, 411, 416,
 839, 845, 847
 Lazarus, R. S., 767
 Lazzaro, T., 157
 Leak, G. K., 804
 Leal, R. M., 197
 Leal, S., 646
 Leander, L., 564
 LeBlanc, S. M., 372
 Lecente, S. W., 800
 Lee, J., 561, 569
 Lee, J. S., 587
 Lee, K., 41, 584, 585, 587, 588
 Lee, T. M., 643
 Lees-Haley, P., 186
 Lees-Haley, P. R., 186
 Lefcourt, H. M., 850
 Legaspi, A., 777
 Leggett, W., 778–780
 Lei, M., 238, 239
 Leichtman, M. D., 576, 577, 579
 Leipziger, M., 767
 Leippe, M. R., 533, 589
 Leis, T., 809
 Losos, D., 287, 294
 Lemert, E. M., 766
 Lempert, R. O., 479
 Lench, H. C., 569, 578, 582
 Lessem, J., 244
 Letourneau, E., 844
 Leukefeld, C. G., 798
 Levene, K. S., 396, 406
 Levenson, J. S., 840
 Levenson, R. W., 851
 Leventhal, J. M., 240, 241, 254
 Levi, B. H., 231
 Levine, J. A., 646
 Levine, L. J., 568
 Levine, R. J., 229
 Levy, R. J., 42
 Lewis, J., 243
 Lewis, M., 585
 Lewis, T., 139
 Lexcen, F., 303, 305
 Lexcen, F. J., 811, 813
 Li, F., 239
 Liber, E., 809
 Liberman, A. M., 454
 Libkuman, T. M., 593
 Licht, H., 839
 Liddle, C., 242
 Lidz, C., 229, 230
 Lidz, C. W., 388, 398–400, 417,
 420, 426, 693
 Lieberman, J., 471
 Lind, E., 479
 Lindner, H., 19
 Lindsay, D. S., 536, 541, 572,
 574, 577, 579
 Lindsay, M., 527
 Lindsay, R., 527
 Lindsay, R. C. L., 529, 538, 540,
 563, 585, 587–589, 595
 Lindsay, W. R., 808
 Lines, K. J., 392
 Linz, D., 485
 Lion, J. R., 417, 418
 Lipmann, O., 5, 13
 Lippert, T., 597
 Lipsey, M., 802, 803
 Lipsey, M. W., 773, 777, 805
 Lipsitt, P. D., 287, 294
 Lipton, D. S., 799, 802, 805
 Listwan, S. J., 764, 769, 772
 Little, T., 48, 759, 770, 815
 Littleton, K. S., 567
 Litwack, T. R., 330, 386, 388,
 389, 391, 400, 408, 419,
 424, 426, 428
 Liu, H. L., 643
 Liu, R., 241
 Liu, S., 587
 Liu, Y., 642
 Livermore, J. M., 321
 Livingston, J. A., 800
 Lloyd, K. P., 288, 289, 291,
 333
 Lobao, L., 812
 Lodewijks, H., 258
 Lodewijks, H. P. B., 407
 Loeb, P. A., 226
 Loeber, R., 258
 Lofquist, W. S., 480
 Loftus, E., 41, 49
 Loftus, E. F., 5, 473, 480, 536,
 563, 572, 575, 588, 675,
 682, 689, 697
 Logan, C., 396, 409, 411, 414,
 416, 847
 Loh, W. D., 5, 13, 14, 27
 Lohr, J. M., 799
 London, K., 583, 585, 592, 593
 London, P., 690
 London, R. W., 692
 Long, A., 182, 187
 Long, A. C., 723
 Long, S. A., 809
 Longabaugh, R., 253
 Looman, J., 841
 Lopez, M. F., 208
 Lösel, F., 48, 855, 856
 Losen, D. J., 209
 Loughead, J. W., 644
 Louisell, D. W., 5, 14
 Lovell, D., 770
 Lovins, L. B., 767, 774
 Lowe, D. J., 475
 Lowenkamp, C. T., 766, 774,
 781, 817
 Lowenstein, D. A., 226
 Lower, J. S., 16
 Lowes, R., 812
 Loza, L., 818
 Loza, W., 817, 818
 Loza-Fanous, A., 818
 Lu, F., 332
 Lu, X., 645
 Luban, D., 476
 Luckasson, R., 303, 304
 Ludlum, J., 228
 Ludolph, P., 148
 Luebbers, S., 43
 Lui, M., 642
 Lui, M. A., 641
 Lujan, J., 777
 Lukoff, D., 332
 Lund, J., 808
 Lund, M. E., 156
 Lundy, R. M., 690
 Luong, D., 20, 22
 Lush, D. E., 12
 Luten, T. L., 594
 Luus, C. A. E., 541
 Lyketsos, C., 229, 230
 Lykken, D. T., 616, 619,
 625–627, 649
 Lymburner, J. A., 335
 Lynham, S. A., 415
 Lynn, S. J., 531, 660, 681, 685,
 689, 690, 692, 697
 Lyon, D., 416
 Lyon, T. D., 567, 569, 583, 585,
 592
 Lyons, P., 36, 39, 171, 172, 180,
 187, 735
 Lyons, P. M., 281, 288, 294,
 355–357
 Långström, N., 394, 406, 407

- Maccoby, E., 148
 MacCoun, R., 472, 478
 MacCoun, R. J., 473
 MacKenzie, D. L., 767, 769, 778, 802
 Mackey-Bilaver, L., 244
 MacKinnon-Lewis, C., 244
 MacLaren, V. V., 627
 MacLin, M. K., 524
 MacLin, O. H., 517, 524, 540
 MacMillan, D. L., 208
 MacMillan, H., 238, 239
 Macmillan, N. A., 518
 Maddi, S. R., 463
 Magaletta, P. R., 21, 760, 768, 781, 812–814
 Maggard, D. L., 462
 Magnussen, S., 587
 Magruder, J., 238, 256
 Mahankali, S., 643
 Main, M., 241, 242
 Makarios, M., 767, 774
 Malcolm, A. S., 462
 Maldjian, J. A., 643
 Malesky, A., 123
 Maletzky, B. M., 843
 Malkin, C., 246
 Malle, B. F., 353
 Mallik, H. S., 232
 Malloy, L. C., 566, 572, 575, 576, 579, 582–584, 593
 Malone, B. E., 594
 Malpass, R. S., 517, 524, 525, 533, 534, 537, 538, 540, 578
 Manchak, S., 806
 Manchak, S. M., 817
 Mandell, W., 818
 Mandracchia, J. T., 806
 Mandy, W., 586
 Mangan, D. J., 625
 Mangraviti, J. J., 740, 744, 755
 Manguno-Mire, G. M., 334
 Manly, J., 242, 243, 255
 Mann, R. E., 847–850, 854
 Mann, S., 646
 Manson, S. M., 322
 Mapes, K. S., 645
 Marbe, K., 8, 9
 Marcon, J. L., 517
 Marcus, D. K., 394, 395, 401, 847
 Marczyk, G., 28, 111, 122
 Marczyk, G. R., 288
 Margolin, G., 823
 Margulies, S., 186
 Marks, J., 666
 Marlatt, G. A., 773
 Marmar, C. R., 454
 Marques, J. K., 802, 855, 856
 Marsden, J., 805
 Marshall, L. E., 845, 849–852, 854–856
 Marshall, N., 238, 256
 Marshall, P., 402
 Marshall, W. L., 839–846, 849–856
 Marsil, D. F., 590, 595
 Marson, D. C., 226, 227, 232
 Marston, W. M., 11, 12
 Mart, E. G., 227, 744, 746
 Martin, J., 471
 Martin, J. A., 289
 Martin, M. A., 300
 Martin, M. H., 797
 Martin, M. L., 396, 406
 Martin, M. S., 796, 806, 807
 Martin, N., 452
 Martin, R. C., 226
 Martindale, D., 137
 Martinez, C., 241
 Martinson, R., 778
 Martinsson, R., 818
 Maruna, S., 850
 Marx, M. L., 361
 Marxsen, D., 528
 Maslow, A. H., 855
 Mason, D., 294, 803
 Mason, P., 766
 Masters, J. C., 764
 Mather, M., 565
 Mathias, R. E., 797
 Matson, J. L., 764
 Matson, J. V., 755
 Matsuda, T., 639
 Mattson, R. E., 294
 Mauer, M., 812
 Maughan, B., 242
 Maurice, D. M., 597
 Mauss, I. B., 851
 Mayberg, H. S., 181
 Mayer, G. R., 214
 Mayer, J., 759
 Mayhew, P., 839
 Maylor, E. A., 528
 Mazzeri, L., 631, 632
 Mazzoni, G., 531, 688
 McAuliff, B., 738
 McAuliff, B. D., 42, 479, 481, 594–596, 738
 McCabe, J., 128
 McCann, J., 157
 McCann, T., 691
 McCarter, L., 851
 McCarthy, R. T., 232
 McCarty, D. G., 13
 McCarty, D. W., 798
 McCary, J. L., 13
 McCauley, M. R., 589
 McClelland, G. M., 811
 McCloskey, L., 569
 McClung, J., 538
 McConkey, K. M., 671, 682, 687–689, 692
 McCord, J., 778
 McCoy, K., 818
 McCullers, E., 814
 McCulloch, A., 800
 McDaniel, M., 215
 McDonald, B. R., 777, 797, 804, 814
 McDonald, E., 851
 McDonald, S., 643
 McDonnell, L. M., 203
 McDonough, E. M., 474
 McElhaney, J. W., 750
 McElroy, H. K., 450
 McFadden, S., 128
 McFarlane, F., 580, 581
 McGarry, A. L., 282, 287, 294
 McGorty, E. K., 528, 564, 565
 McGovern, A., 590
 McGrath, R. J., 800, 849
 McGreevy, M., 336
 McGreevy, M. A., 336, 337, 355, 367
 McGuire, J., 48, 765, 796
 McHugh, P. R., 232
 McIntosh, J., 148
 McInturff, B., 226
 McKee, G. R., 303
 McKelvie, S. J., 527
 McKevitt, B. C., 203
 McLaughlin, M. A., 484, 488, 489, 497
 McLaughlin, M. J., 203
 McLean, A., 851
 McLean, M., 43
 McMinn, M. R., 96
 McMurrin, M., 414, 799, 800
 McNally, R. J., 453
 McNamara, E., 561, 569
 McNiel, D. E., 388, 420, 426
 McQuiston-Surrett, D., 540

- McWilliams, K., 563, 569, 571, 573
- Mead, G., 406
- Meade, M. L., 572
- Mears, D. P., 770, 771
- Meehl, P., 128
- Meehl, P. E., 321, 390, 391, 426, 846
- Megargee, E. I., 21, 27
- Meijer, E., 625, 631, 650
- Meijer, E. H., 625, 637
- Meisel, A., 229, 230
- Meisenhelder, T., 795
- Meissner, C. A., 517, 524, 525, 527, 530, 532, 536, 540
- Meixner, J. B., 641
- Meldrum, L., 820
- Melinder, A., 565, 570, 571, 576, 579, 591
- Melnick, G., 799
- Melnyk, L., 566, 575, 581
- Meloy, J. R., 409, 452
- Melton, G., 112, 171, 172, 180, 187, 247
- Melton, G. B., 36, 39, 137, 142, 221, 224, 225, 229, 281, 288, 294, 315, 324, 326–329, 336, 337, 355–357, 735
- Melville, J. D., 323
- Memon, A., 530
- Mendoza, M., 563, 569
- Mennen, F., 258
- Menzies, R. J., 282, 294, 419
- Mercado, C. C., 840
- Merckelbach, H. L., 625
- Merenbach, D. G., 755
- Mericle, A. A., 811
- Merkelbach, H., 625, 650
- Merrill, M. A., 24
- Merten, T., 102
- Mertens, R., 638, 640–642
- Meskeil, M. W., 761
- Mesout, J., 526
- Messer, J., 242
- Messick, S., 291
- Messina, N. P., 800, 805
- Metzler, T. J., 454
- Metzner, J. L., 779, 780
- Meyer, R. G., 179
- Michie, C., 393, 847
- Middleberg, L., 209
- Middleton, H., 809
- Milan, M. A., 772
- Miletich, D., 738
- Milgram, D., 797
- Milgram, S., 537
- Miller, B. S., 589
- Miller, C. K., 233
- Miller, D. G., 536
- Miller, H. A., 188
- Miller, H. L., 16
- Miller, L., 450
- Miller, M., 777
- Miller, M. O., 120
- Miller, R., 333, 338
- Miller, R. S., 850
- Miller, T., 254
- Miller, T. W., 813
- Miller, W., 253
- Milling, L. S., 688
- Millon, C., 157
- Millon, T., 157, 733
- Mills, J. F., 395, 796, 801, 803, 804, 806, 807, 815, 817–819, 821
- Milner, J., 242
- Miner, M. H., 855
- Minkoff, K., 819
- Miraglia, R., 337
- Mishara, B. L., 452
- Mitchell, C. M., 777
- Mitchell, C. N., 282
- Mitchell, E. B., 584, 593
- Mitchell, G., 530
- Mitchell-Herzfeld, S., 259
- Mnookin, R., 148
- Moan, S. F., 564
- Moeller, S., 238
- Mohandie, K., 452
- Mojtabai, R., 257, 331
- Mokros, A., 843
- Moll, A., 667
- Monaghan, J., 668
- Monahan, J., 5, 8, 40, 41, 50, 91, 115, 128, 281, 297–299, 385, 386, 388, 392, 398, 408, 417, 419, 424, 425, 846
- Monnier, J., 814
- Monson, C. M., 338
- Moore, C. C., 10
- Moore, D. L., 800
- Moore, J. T., 297
- Moore, K., 251, 252
- Moos, R. H., 768
- Moran, G., 340, 341, 480
- Moran, P., 820
- Moran, P. B., 589
- Mordell, S., 402
- Morey, L. C., 157, 188
- Morgan, P. S., 645
- Morgan, R. D., 772, 777, 781, 795–798, 800–804, 806, 807, 812–815, 821
- Morganroth, F., 291
- Mori, E., 643
- Morison, P., 203
- Morris, G., 567
- Morris, G. H., 291
- Morris, S. B., 711
- Morrissey, J., 336, 337
- Morrissey, J. P., 355, 367, 798
- Morry, M. M., 476
- Morse, S. J., 59, 355, 358, 359, 361–364, 366, 368, 371, 645
- Morton-Bourgon, K., 847
- Morton-Bourgon, K. E., 401, 404, 405, 760, 847, 848, 850, 851
- Mosher, C., 812
- Mosmann, A. L., 634, 635
- Mossman, D., 394
- Motiuk, L., 767
- Motiuk, L. L., 770
- Motivans, M., 800
- Mott, N., 485, 490
- Mott, N. L., 473
- Moulden, H., 851, 854
- Moulden, H. A., 841
- Moulden, H. M., 841, 851
- Moussalli, S., 245
- Movius, H. L., 636
- Mowrer, O. H., 762
- Moye, J., 233
- Mu, Q., 644, 645
- Mudd, D., 820
- Mueser, K. T., 807
- Mülberger, A., 5, 7, 9, 13
- Mullen, P., 332
- Mullen, P. E., 47
- Müller-Isberner, R., 807
- Mullett, N., 297
- Mulligan, W. C., 668
- Mulloy, R., 854
- Mulvaney, F., 800
- Mulvey, E., 283, 392, 810
- Mulvey, E. P., 115, 128, 388, 398–400, 417, 420, 426
- Mumley, D., 61
- Mumley, D. L., 306
- Münsterberg, H., 10, 51, 517, 686, 734
- Munsterman, G. T., 481, 482

- Munsterman, T., 499, 500
 Murai, J., 646
 Murphey, M., 584
 Murphy, B., 479, 485, 489–491, 499, 500
 Murphy, K. R., 844
 Murphy, W., 802, 855, 856
 Murphy, W. D., 844, 846
 Murray, C., 803
 Murray, D., 806
 Murray, H. A., 413
 Murray, P., 805
 Murrie, D. C., 287, 288, 290, 291, 336
 Murrin, M., 42
 Musick, J. E., 300, 301
 Mussack, S. E., 849
 Mutter, C. B., 660, 661, 675, 680, 682, 692–694
 Myers, B., 649
 Myers, J. E. B., 561, 584, 585, 589–591, 594, 595
 Myers, L. B., 800
- Nader, R., 188, 189
 Nadkarni, L., 726
 Nadon, R., 682
 Nagin, D. S., 764, 765, 821
 Naik-Polan, A., 250
 Naimark, D., 291, 323
 Nairne, J. S., 565
 Najdowski, C. J., 49
 Nakayama, M., 628
 Napoli, D. S., 17
 Narby, D., 341
 Narr, R., 563, 569
 Nasman, V. T., 631, 632
 Neale, J. M., 330
 Nedd, A. N. B., 798
 Needed, B., 238, 256
 Neil, K. M., 221
 Neisser, U., 564, 566
 Nelson, C. B., 181
 Nelson, E., 812–814
 Nelson, F. R., 297
 Nelson, S. E., 353
 Nelson, W., 176, 177, 179, 180
 Nelson-Gray, R. D., 840, 844
 Nemeroff, C. B., 181
 Neo, L. H., 818
 Neppl, T., 240, 241
 Nessa, D. B., 286, 295
 Neuschatz, J., 681, 685, 689, 692
 Neuschatz, J. S., 517, 538
- Neutze, J., 256
 Neves, A. C., 407
 Nevins-Saunders, E., 372
 Newberger, E., 240
 Newbern, D., 800
 Newcomb, A., 663
 Newhill, C. E., 420
 Newman, B. S., 597
 Newman, R., 283
 Newton, R., 238, 239, 260
 Nicholas, E., 564
 Nicholls, T., 407, 427
 Nicholls, T. L., 171, 187, 396, 406
 Nicholson, R., 283, 297, 298, 331
 Nicholson, R. A., 287, 290, 295, 297, 299
 Nicole, H., 564
 Nicoletti, J., 457
 Niederhoffer, A., 453
 Niederhoffer, E., 453
 Niendam, T., 637, 638
 Nietzel, M., 25
 Nightingale, N. N., 589, 590
 Nisbett, R., 596
 Nitschke, J., 843
 Nitz, K., 243
 Noll, J. G., 584
 Noor, F., 524
 Norcross, J. C., 819
 Norgate, S. H., 567
 North, C. S., 181
 Norwood, S., 283
 Nose, I., 646
 Nossiter, A., 764
 Notario-Smull, H., 301
 Nottingham, E. J., 294
 Novaco, R. W., 760
 Nuñez, N., 585, 592, 595
 Nuffield, J., 401
 Nunes, K., 808
 Nunez, D. T., 532
 Nunez, G., 479
 Nussbaum, D., 297, 812–814
 Nysse-Carris, K. L., 585, 590
- Oates, L. C., 76
 O'Brien, M. D., 845, 852–855
 O'Connor, S. D., 469
 O'Donnell, I., 818
 O'Donohue, W. T., 139, 394, 650, 842, 844
 Oeberst, A., 530, 531
 Ogan, G. D., 800
- Ogle, C. M., 561, 564, 567, 570, 571, 583, 593
 Oglesby, T. W., 24, 25
 Ogloff, J., 337, 339, 342
 Ogloff, J. R., 328
 Ogloff, J. R. P., 5, 28, 36, 43, 47, 48, 51, 326, 388, 395, 406, 407
 O'Keefe, M. L., 778–780
 Oksol, E. M., 650
 Olczak, P. V., 45
 Olesen, N., 149
 Oliver, P., 762
 Olsen, S. S., 805
 Olson, E. A., 532
 Olson, J. R., 818
 Oltmanns, T., 330, 331
 Olver, M., 402
 Olver, M. E., 28, 800, 819
 O'Neill, M., 179–181, 584
 O'Neill, M. L., 171, 187
 Ones, D. S., 449
 Opton, E., 659, 666
 O'Quin, K., 454
 Oran, D., 570
 Oran, H. S., 570
 Orbach, Y., 154, 266, 267, 564, 569, 594
 Orcutt, H. K., 596
 Orne, E. C., 677, 680, 685, 692
 Orne, M. T., 623, 671, 675, 677–680, 685, 687, 691, 692, 780
 Ornstein, P. A., 567, 581
 Orr, W. B., 232
 Orsagh, T., 764
 Orsak, C., 584
 Ostapuik, E., 240
 Osterheider, M., 843
 Ostrom, B., 476, 481
 Ostrom, T. M., 533
 Ostrov, E., 451
 Otgaar, H., 565, 569, 575
 Otto, R., 182, 187, 254, 256
 Otto, R. K., 4, 28, 36, 39, 50, 139, 156, 171, 172, 180, 187, 281, 288, 294, 297–301, 355–357, 723, 733–735
 Ovenstad, K., 571
 Owen, C., 820
 Owen, D. G., 474
 Owen, L., 240, 241
 Owens, C., 798
 Owens, R. M., 809

- Ownby, R., 251
 Oxtoby, O., 532
- Packard, R. L., 840
 Packard, T., 747
 Packer, I. K., 315, 321–324,
 329, 333, 718
 Padgett, T. M., 644
 Page, K. S., 452
 Palfrey, J. S., 212
 Palma-Oliveira, J. M., 407
 Palmer, C. A., 323
 Palmer, E. J., 817
 Palmer, J. C., 536
 Palmer, T., 769, 778
 Palomares, R. S., 812
 Pandeirada, J. N. S., 565
 Paolucci, E., 839
 Papatola, K., 240, 242
 Papierno, P. B., 562
 Pardini, D., 258
 Park, N., 113
 Parker, A. L. S., 334
 Parker, H. A., 781
 Parker, J. D. A., 851
 Parker, J. F., 517, 526, 540,
 589
 Parker, L. S., 693
 Parkinson, C., 589
 Parkinson, P., 151
 Part, II., 682, 687
 Pasewark, R., 335, 336, 341
 Pashler, H., 215
 Paternoster, R., 821
 Paterson, D. G., 25
 Patnaik, P., 619
 Patrick, A. R., 813, 814
 Patrick, C. J., 615, 618, 622, 623,
 625, 845
 Patrick, J., 209
 Patry, M. W., 21
 Patterson, G., 240, 241
 Patterson, G. R., 766
 Patterson, M. E., 297
 Pavlidis, I., 646
 Payne, J., 471
 Payne, J. W., 471, 475
 Payne, K., 843, 845
 Paz-Alonso, P. M., 564, 568,
 583
 Peace, K. A., 529
 Peacock, E. J., 846
 Pearlstone, Z., 530
 Pears, K., 240, 241
 Pearson, F. S., 802
- Pearson, J., 259, 260
 Pearson, R., 586
 Pearson, S., 839
 Pedersen, L., 407
 Pekarik, G., 799
 Pelissier, B., 800
 Pendleton, D., 597
 Pendleton, L., 289
 Penner, E., 297
 Pennington, N., 44, 340, 473
 Penrod, S., 45, 470, 472, 485,
 541
 Penrod, S. D., 10, 42, 221, 481,
 485–491, 493–498, 525,
 528, 536, 538, 564, 565,
 586, 737
 Pensky, E., 242
 Peplau, L. A., 850
 Perillo, A. D., 840
 Perlin, M., 341–343
 Perlin, M. L., 219, 222–224,
 227, 284, 285
 Perlini, A. H., 689
 Perretta, S., 542
 Perry, A. R., 590
 Perry, C., 659, 663, 666,
 677–679, 682, 687, 691,
 692
 Perry, J. L., 479
 Perry, N. W., 642
 Peszke, M. A., 286
 Peters, J. M., 846
 Peters, R. H., 42
 Peterson, C., 113, 563,
 566–568, 576, 582, 585,
 852
 Peterson, D., 114
 Peterson, J. K., 806
 Peterson, M. A., 471, 476
 Petrella, R. C., 302
 Petrila, J., 36, 39, 41, 112, 137,
 142, 171, 172, 180, 187,
 221, 224, 225, 229, 281,
 288, 294, 315, 324,
 326–329, 336, 337,
 355–357, 735
 Petrosino, A., 777
 Peyrot, M. F., 812–814
 Pezdek, K., 536, 572, 573, 575,
 581, 587
 Pfohl, S. J., 419
 Phelps, E. A., 564, 565, 645
 Philibert, R., 238, 239
 Phillips, J., 372
 Phillipson, C., 761
- Picarello, K., 189
 Picheca, J. E., 841
 Pickel, K. L., 529
 Pickering, A., 527
 Pickles, A., 242
 Pickrel, S. G., 810
 Pickrell, J. E., 536, 575
 Pidgeon, A., 243
 Piehl, A. M., 765
 Pierce, P. S., 226
 Pilpel, M. E., 12
 Pinta, E., 426
 Pipe, M.-E., 564, 565, 569,
 585
 Piquero, A., 777
 Pirelli, G., 286, 287, 290, 306,
 307
 Pitre, U., 800
 Pitt, S. E., 176, 177, 179, 180
 Pizzi, W. T., 288
 Platt, A. M., 317, 318
 Platt-Houston, C., 245, 247
 Plomin, R., 243
 Poindexter, L., 250
 Polaschek, D. L. L., 805, 808,
 818, 840
 Polcari, A., 570
 Polizzi, D. M., 802
 Polvi, N. H., 419
 Pomarol-Clotet, E., 406
 Pomerantz, A. M., 726
 Pomerol, J.-C., 415
 Pontari, B. A., 850
 Poole, D. A., 574, 577, 579
 Pope, H. G., 697
 Porporino, F., 767–769
 Porporino, F. J., 767
 Port, L. K., 593
 Porter, B., 43
 Porter, S., 528
 Porter, T., 518
 Posavac, E. J., 475
 Posey, A. J., 40, 43, 45
 Posner, R. A., 469, 471, 482
 Potter, J., 823
 Poulin, F., 778
 Poulson, R. L., 339, 340
 Powell, J. E., 188
 Powell, M., 43
 Powell, M. B., 580, 581, 592
 Powell, T. A., 774
 Poythress, N., 50, 112, 171,
 172, 180, 187, 281, 297,
 298, 327, 735

- Poythress, N. G., 36, 39, 137,
142, 221, 224, 225, 229,
281, 288, 290, 294, 299,
315, 324, 326, 328, 329,
336, 337, 355–357
- Pozzulo, J. D., 563, 587, 588
- Prado, L., 593
- Prag, B., 532
- Prendergast, M. L., 799, 800,
805
- Prentky, R. A., 423
- Present-Koller, J., 810
- Preston, D. L., 808
- Price, H. L., 593
- Prins, H., 411, 417, 418
- Pritchard, C. C., 474
- Prizmic, Z., 851
- Prizmich, L. P., 584, 589, 590
- Prochaska, J. O., 819
- Proctor, S., 123
- Proporino, F., 802
- Proporino, F. J., 802
- Prosser, W. L., 474, 475
- Pryce, J., 591
- Purver, J. M., 484
- Putman-Hornstein, E., 238, 256
- Putnam, F. W., 584
- Putnam, W. H., 687
- Putney, G. J., 763
- Putney, S., 763
- Qian, J., 632
- Qin, J., 568, 570, 581, 584, 585,
594, 595
- Quas, J. A., 561, 565, 566,
568–572, 575, 576, 578,
579, 582, 583, 585,
589–595
- Quay, H. C., 27
- Quayle, D., 470
- Quigley, C. A., 689
- Quinlivan, D. S., 535, 538, 545,
588
- Quinn, P. C., 587
- Quinnell, F. A., 138
- Quinsey, F., 254
- Quinsey, V., 116, 392
- Quinsey, V. L., 386, 392, 394,
395, 397, 405, 408, 426,
802, 804, 844, 846, 847,
855, 856
- Quinton, D., 242
- Rabasca, L., 446, 461
- Rabbitt, P. M., 528
- Rachlinski, J. J., 475, 481
- Raczynski, J. L., 541
- Radaelli, S., 525
- Raginsky, B. B., 685
- Ragland, J. D., 643
- Rai, A., 61
- Raichle, M. E., 645
- Raine, A., 258
- Rakos, R. F., 481
- Ralston, C. A., 392
- Ramsey, S. H., 143
- Randolph, J. J., 294
- Raney, G., 642
- Ranganathan, M., 528
- Rankin, J. H., 25
- Rappaport, J. R., 336
- Rappaport, R. G., 797
- Rapson, R. L., 715
- Rasch, W., 767
- Raskin, D., 617, 622, 624, 626,
629
- Raskin, D. C., 626, 628
- Rasmussen, K., 392, 407
- Rathner, J., 849
- Rattenbury, F., 331
- Ray, 330, 331
- Ray, I., 319
- Raynor, G., 585
- Read, J. D., 536, 592, 593
- Read, R., 541
- Reardon, M. L., 288, 289, 291
- Reckers, P. M. J., 475
- Redding, R. E., 810, 811
- Reddy, M., 426
- Reder, L. M., 572
- Redlich, A. D., 41, 303, 565,
570, 571, 584, 585, 589,
590, 594, 595
- Redner, R., 777
- Reed, A., 330, 331
- Reed, G. M., 197
- Reed, R. S., 562
- Rees, M., 567
- Reese, J. T., 26, 443
- Regan, P. C., 584, 590, 592
- Regev, L. G., 842
- Rehnman, J., 526
- Reinsmith-Meyer, C. L., 798
- Reiser, M., 25, 26, 443, 668,
672
- Reisner, R., 61
- Reiss, S. A., 484, 493, 494
- Reitzel, L. R., 781
- Relman, A. S., 333
- Rennie, C. E., 407
- Repo-Tiihonen, E., 807
- Reschly, D., 209
- Reschly, D. J., 198, 199,
207–210, 214, 215
- Resick, P., 839
- Resick, P. A., 47
- Resnick, R. J., 17, 24
- Ressler, R. K., 43
- Revenstorff, D., 823
- Reyes, A., 406
- Reyes, C., 747
- Reyna, V. F., 470, 572–575
- Rhodes, M. G., 525
- Rice, M., 116, 254
- Rice, M. E., 386, 392, 394, 395,
397, 405, 408, 426, 846,
847, 855
- Rich, P., 850
- Richards, B., 767
- Richards, H. J., 800
- Richardson, D. C., 636
- Richardson, G., 636
- Rickards, L., 819
- Riddlesberger, M. M., 566, 569,
571, 576
- Rimm, D. C., 764
- Ringland, G., 415
- Riordan, R. J., 797
- Rios, P., 485
- Roach, E. S., 715
- Robbenolt, J. K., 470–473,
475–478, 480, 481
- Robbins, P. C., 41, 115, 128,
355, 367
- Roberts, B., 229, 230
- Roberts, C. F., 328, 337, 339,
341–344
- Roberts, J. V., 764
- Roberts, L. W., 229, 230
- Roberts, M., 449
- Roberts, M. R., 448
- Robertson, H. C., 297
- Robey, A., 293, 712
- Robins, P. C., 336, 337
- Robinson, D., 802
- Robinson, E. S., 13
- Robinson, P., 341–343
- Robinson, R., 290
- Robinson, R. J., 493, 495
- Rodrigues dos Santos, B.,
597
- Rodriguez, M. C., 203
- Roe, C., 572, 573, 581
- Roebbers, C. M., 581
- Roediger, H. L., 572

- Roesch, R., 50, 281–283,
286–291, 294–297, 303,
305, 307, 316, 318, 328,
335
- Roff, M., 12
- Rogers, J. L., 322
- Rogers, L., 585, 594, 595
- Rogers, O., 13
- Rogers, R., 61, 188, 297–300,
303, 315, 324, 327–329,
336
- Rogers, R. R., 282
- Rogosch, F., 237–239,
241–243, 246, 255
- Rogstad, J. E., 338
- Rohrer, D., 215
- Rolfe, A., 805
- Romanczyk, A., 589
- Romani, C. J., 777, 797
- Ronan, K. A., 288, 291
- Ronay, K., 260
- Rooney, J., 802, 817
- Rorschach, H., 157
- Rosanbalm, K., 246, 254, 256
- Rose, J. L., 819
- Rose, M., 479, 485, 489, 490,
499, 500
- Rose, M. R., 485, 490, 491
- Rose, S., 528
- Rose, V. G., 470, 472
- Rose, V. R., 36
- Roseby, V., 46
- Rosen, G. M., 188
- Rosen, R. C., 844
- Rosenberg, C. E., 320
- Rosenfeld, B., 221, 224, 263,
427
- Rosenfeld, J. P., 631, 632, 634,
636–642
- Rosenfield, B., 804
- Rosenhan, D., 690
- Rosenhan, D. L., 493, 495
- Rosenthal, R., 537
- Ross, D. F., 525, 563, 589, 595
- Ross, L., 596
- Ross, R., 48
- Ross, R. R., 764, 777, 778, 802,
851
- Ross, S. J., 517, 529, 530
- Rostow, C. D., 449
- Roth, L., 229, 230, 392
- Rothman, D. J., 761, 763
- Rotman, E., 761–763
- Rotolo, T., 812
- Rottman, D., 476, 481
- Rovosky, F. A., 228
- Rowe, D., 243
- Rowland, E., 19
- Rozycki, A. T., 797, 798
- Rubin, D. C., 571
- Rubin, J. S., 797
- Rudy, L., 562, 563, 568, 572,
575, 576, 593
- Ruiz, S., 563, 569
- Runyan, D., 238, 239, 246, 254,
256, 260
- Ruparel, K., 644
- Rush, E. B., 568, 569
- Russano, M., 527
- Russano, M. B., 538, 738
- Russell, D., 850
- Russell, E. J., 529
- Russell, L., 289
- Russo, G., 237
- Rustad, M., 471, 476
- Rutter, M., 240–242, 264
- Ryan, A., 634, 638–641
- Ryan, E. P., 811
- Ryan, S. L., 541
- Ryan, V., 526
- Ryba, N. L., 287, 303
- Rychlec, M., 818
- Sacco, D. F., 525
- Sachdev, A., 470
- Sachsenmaier, T., 596
- Sacks, M., 331
- Sadler, C. M., 774
- Sadoff, R., 692
- Sadoff, R. L., 747
- Saeed, S. A., 814
- Saks, M., 420
- Saks, M. J., 471, 473, 476, 480,
481, 648, 650, 678
- Salekin, R. T., 303
- Salerno, J. M., 585
- Sales, B., 471
- Sales, B. D., 43, 46, 111,
115–117, 120, 128, 130,
131, 471, 593, 595, 711,
840
- Sales, J. M., 567
- Samenow, S. E., 773
- Sampson, R., 821
- Sampson, R. J., 821
- Samuels, D. J., 621
- Sand, L. B., 484, 493, 494
- Sandberg, D. A., 388
- Sanders, J., 420, 479, 648, 650,
678
- Sanders, M., 243
- Sandhei, M., 587
- Sang, J., 258
- Sapsford, R. J., 767
- Sargent, E. L., 339, 342
- Sarma, S., 471
- Sasaki, M., 639
- Sauerland, M., 542
- Saum, C. A., 799, 800
- Savenije, W., 407
- Savitz, B., 800
- Savolainen, J., 821
- Sawrie, S., 226
- Saxe-Clifford, S., 450
- Saxena, S., 197
- Sayfan, L., 565, 584, 587
- Saylor, W., 768
- Sayre, F. B., 317
- Saywitz, K., 266, 267
- Saywitz, K. J., 564, 569, 597
- Sbarra, D., 46
- Scalia, M. B., 762
- Scalora, M., 406
- Scaramella, L., 240, 241
- Scarpitti, F. R., 800
- Schaaf, J. M., 576, 579
- Schaap, G., 406
- Schacter, D. L., 675
- Schaeffer, C. M., 856
- Schalock, R. L., 210
- Schefflin, A. W., 659, 660, 663,
666, 668, 669, 672, 673,
675, 679–684, 688, 689,
691–694
- Scheibe, K. G., 780
- Schepard, A., 163
- Scherf, K. S., 587
- Scherner, G., 256
- Schipper, L. J., 716
- Schkade, D., 471, 473, 477, 478
- Schkade, D. A., 471, 475
- Schlauch, C., 818
- Schlenker, B. R., 850
- Schlesinger, L. B., 385, 386, 408
- Schmid, K., 239, 246
- Schmidt, F., 407
- Schmidt, F. L., 759, 780
- Schmod, J., 625, 650
- Schmucker, M., 855, 856
- Schmuckler, E., 456
- Schneider, D., 340
- Schneider, W., 581
- Schnider, R., 143
- Schnur, A. C., 770, 771
- Schoendorf, K., 156

- Schoenwald, S. K., 810
 Schopp, R. F., 116
 Schouten, P. G. W., 846
 Schreiber, J., 287, 290, 294–296
 Schreiber, N., 530
 Schreiber, T. A., 572
 Schreiber Compo, N., 527, 528
 Schrenck-Notzing, A., 8
 Schroeder, L., 643
 Schuller, R., 470, 472
 Schulz-Heik, R., 244
 Schuman, D., 61
 Schumann, E. L., 479
 Schuth, K., 245
 Schutz, B. M., 157
 Schwab, S. J., 481
 Schwartz, R., 303, 305
 Schwartz, R. G., 303
 Schwartz-Kenney, B. M., 562, 563, 572, 575, 585
 Schweighofer, A., 337
 Scoboria, A., 688
 Scott, A., 316, 358
 Scott, E., 303, 305
 Scott, G. D., 778
 Scott, L. S., 587
 Scott, P. D., 408, 411
 Scott, T., 820
 Scott, T. H., 780
 Scrivner, E., 450, 456, 461
 Scrivner, E. M., 25, 448, 459
 Scully, J., 487
 Seager, J. A., 799
 Searight, H. R., 233
 Sedler, M., 331, 332
 Segal, Z. V., 853
 Seifert, C. M., 634, 635
 Seligman, M. E. P., 852
 Selke, T. M., 394
 Sellbom, M. O., 50
 Sellers, A. H., 462
 Sellin, T., 760
 Seman, W., 328
 Semmler, C., 527
 Sergent-Marshall, S., 572
 Serin, R. C., 808
 Serna, C., 807
 Serran, G., 843
 Serran, G. A., 845, 850–852, 854, 855
 Seto, M. C., 802, 844, 855, 856
 Sewell, J. D., 456
 Sewell, K. W., 188, 298–300, 328
 Seying, R., 591
 Seymour, T. L., 619, 634, 635
 Sgroi, S., 252
 Shady, T., 259
 Shaffer, D. R., 339, 340
 Shaffer, M., 517
 Shafir, Z., 593
 Shafto, M. G., 634, 635
 Shah, S., 111, 120, 127
 Shah, S. A., 111, 385
 Shahinfar, A., 818
 Shandera, A. L., 716
 Shaner, A., 819
 Shanley, M., 471
 Shapiro, C., 589, 596
 Shapiro, D. L., 715, 721
 Shapiro, L., 659, 668, 692, 693
 Sharma, A., 578
 Shartle, C. L., 18, 20
 Shatin, L., 294
 Shaw, E., 417
 Shaw, E. C., 399, 400
 Shaw, L., 798, 800
 Shearer, C., 242
 Shearer, R. A., 800
 Sheehan, D. C., 462
 Sheehan, P. W., 659, 671, 687, 689, 691, 692
 Sheidow, A. J., 810
 Shen, F. X., 645
 Shen, Z., 642
 Shepherd, J., 527
 Sherman, D., 247
 Sherrod, C. B., 300, 301
 Shi, J., 806
 Shilhadeh, E. S., 798
 Shobe, K. K., 571
 Showalter, C. R., 326, 327
 Shuchman, M., 232
 Shuman, D., 119, 642
 Shuman, D. W., 92, 95, 116, 176, 177, 179, 315, 324, 327–329, 583, 711, 712
 Sicola, A., 226
 Siegel, A. M., 289
 Siegel, J., 243
 Sigall, H., 650
 Silten, P. R., 282
 Silver, E., 115, 128, 333, 334, 336, 338, 341, 388, 392, 424
 Silver, S. B., 338
 Silverman, M., 303
 Silvern, L., 244
 Simcock, G., 567
 Simon, R., 238, 239, 241
 Simon, R. J., 335, 339
 Simon, T., 809
 Simon, W. T., 846
 Simone, M., 597
 Simone, S., 426
 Simourd, D., 48
 Simpson, D. D., 799, 800
 Simpson, J., 768
 Simpson, J. T., 478
 Simpson, L., 473
 Simpson, M., 781, 796
 Sindberg, R., 797
 Singer, J. D., 212
 Singer, P. A., 232
 Singh, J. P., 403–405, 427
 Single, J., 151
 Sirois, F., 715
 Sjöstedt, G., 406
 Skeem, J., 283, 286–288, 290, 291, 293
 Skeem, J. L., 336, 341, 343, 395, 397, 426, 760, 806, 817
 Skelton, A., 847
 Skiba, R. J., 208, 209
 Sklansky, J., 342
 Skogstad, P., 798
 Skuse, D., 586
 Sladden, B., 525, 527
 Slater, S., 852
 Slesinger, D., 5, 12
 Slis, V., 260
 Sloan, A. E., 76
 Sloat, L. M., 341
 Slobogin, C., 36, 39, 61, 112, 137, 142, 171, 172, 180, 187, 221, 224, 225, 229, 281, 288, 294, 315, 324, 326–329, 336, 337, 355–357, 359, 735
 Slomen, D., 289
 Small, M., 538
 Smeets, T., 565
 Smith, A., 474
 Smith, A. C., 474
 Smith, B., 595
 Smith, D., 446
 Smith, F. V., 767
 Smith, G. P., 188
 Smith, J. L., 809
 Smith, J. P., 471
 Smith, M., 245
 Smith, P., 759, 760, 765–768, 771, 772, 774, 777–779, 781, 782, 817

- Smith, S., 479, 485, 489–491
 Smith, S. R., 726
 Smith, S. S., 636–638
 Smith, V. L., 480
 Smithberger, A., 96
 Smoot, S. M., 450, 456
 Smythe, B., 148
 Snee, T. J., 12
 Snook, B., 43
 Snyder, S., 226
 Snyder, T. D., 210, 212
 Sobel, N. R., 539
 Sokol, K., 148
 Sollman, M. J., 716
 Solomon, J., 148, 268
 Solomon, P., 807
 Solomon, P. L., 238
 Sonnega, A., 181
 Soskins, M., 634, 637–641
 Spanos, N. P., 592, 689
 Sparrow, S. S., 840
 Spataro, J., 47
 Spence, S. A., 643
 Spencer, L. D., 286, 295
 Spicer, J., 798
 Spiegel, D., 670, 682, 687, 694
 Spiegel, H., 680, 682, 687, 691
 Spiegler, M. D., 762
 Spindler, S., 584
 Spitzer, R., 331
 Sporer, S. L., 5, 7, 10, 525, 528,
 541, 542, 587
 Springman, R. E., 102
 Sprung, H., 9
 Sprung, L., 9
 Srivastava, S., 823
 Staebler, C. R., 474
 Stafford, K. P., 50, 281
 Stahl, P., 137, 143, 148, 149,
 151, 163, 259, 260, 265,
 266, 268
 Stahl, T., 625, 650
 Stalans, L., 340
 Stalvey, T., 226
 Stamm, B. H., 812
 Stanley, B., 229
 Stanley, M., 229
 Starkey, M. L., 715
 Starnes, H., 47
 Starzyk, K. B., 852
 Staton, M., 798
 Steadman, H., 336, 392
 Steadman, H. D., 41
 Steadman, H. J., 41, 91, 115,
 128, 286, 288, 330, 333,
 334, 336, 337, 341, 355,
 367, 389, 797, 798
 Steblay, N., 540
 Steblay, N. K., 541
 Steblay, N. M., 690
 Steele, C. M., 528
 Steffan, J. S., 796–798, 800, 801,
 804, 806
 Steffen, L. J., 163
 Stein, D., 331
 Stein, M. B., 181
 Steinberg, L., 303, 305
 Steiner, B., 768
 Steiner, H., 303
 Steiner, R., 800
 Stephen, A., 755
 Stephenson, G. M., 562
 Stermac, L., 289
 Stern, L. W., 6–8, 12
 Sternbach, K. O., 819
 Sternberg, K., 147, 148, 266,
 267
 Sternberg, K. J., 594
 Stevens, K., 186
 Stevenson, M. C., 589–591
 Stewart, C. O., 447, 449, 450
 Stewart, D., 805
 Stewart, E., 238, 239
 Stewart, E. A., 770, 771
 Stimpson, S., 372
 Stinson, J. D., 840
 Stock, H. V., 290
 Stockdale, K., 402
 Stockdale, K. C., 819
 Stoeter, P., 646
 Stoll, F., 331
 Stone, M., 290
 Storey, J. E., 428, 848
 Stose, S., 643
 Stouthamer-Loeber, M., 258
 Strachan, A. M., 156
 Strand, S., 428
 Strasburger, L. H., 104
 Stratton, J. G., 451
 Stratton, J. R., 770
 Strauss, J., 331
 Stredny, R. V., 187, 334
 Strother, A., 116
 Stroud, D. A., 317, 318
 Strunk, J., 818
 Stucker, A., 778–780
 Studebaker, C., 470, 472
 Studebaker, C. A., 221, 473,
 480, 737
 Sturm, K., 778–780
 Sturmey, P., 414
 Su, M., 245–247
 Suboski, M. D., 765
 Suedfeld, P., 780
 Suhr, J. A., 102
 Sullivan, C. J., 764
 Sullivan, H. S., 324
 Sullivan, M., 47, 149
 Sullivan, M. J., 47
 Sulzer-Azaroff, B., 214
 Summers, R., 818
 Sundt, J. L., 765, 766
 Sunstein, C., 471, 473, 477, 478
 Super, J. T., 447, 458
 SurrIDGE, C. T., 767
 Susa, K. J., 524
 Sutherland, M., 565
 Suzuki, M., 643
 Swanson, C. R., 24
 Swanson, J., 417–419
 Swartz, K., 760
 Swartz, M., 417–419
 Swijtink, Z., 518
 Swope, M., 813
 Sydlik, B. L., 47
 Sykes, G., 766
 Symonds, B. D., 820
 Szasz, T., 365, 366
 Taira, M., 646
 Talley, J., 814
 Talucci, V., 451, 455
 Talwar, V., 584, 585
 Tan, L. H., 643
 Tanasichuk, C. L., 767
 Tanke, T., 151
 Tapp, J., 245, 778
 Tapp, J. L., 5, 27
 Taragin, M. I., 476
 Tardiff, K., 408
 Tate, D. C., 810, 811
 Tatham, A., 589
 Taub, E. P., 593
 Taub, S., 693
 Taxman, F. S., 778
 Taylor, G. R., 839
 Taylor, I., 760
 Taylor, J., 808
 Taylor, J. L., 808
 Taylor, N. J., 151
 Taylor, P., 43, 330, 331
 Teeters, N. K., 19
 Teicher, M. H., 570
 Tellefsen, C., 338
 Tellegen, A., 157, 187, 188

- Teplin, L., 294
 Teplin, L. A., 806, 811
 Terman, L. M., 23
 Terr, L. C., 564, 566
 Terrio, L., 24
 Terry, K., 245
 Teti, D., 243
 Theodosi, E., 799
 Theonnes, N., 259, 260
 Thériault, Y., 759, 760, 772, 774, 778, 780–782
 Thomas, C., 767
 Thomas, C. W., 767
 Thomas, S. D., 43, 47
 Thomas, S. D. M., 43
 Thomas, S. F., 578, 596
 Thomas, S. N., 585
 Thompson, D. C., 68, 69
 Thompson, H. W., 334
 Thompson, J. W., 289, 334
 Thompson, L., 297
 Thompson, R., 147, 148
 Thompson, R. A., 851
 Thompson, S. R., 565
 Thompson, W., 800
 Thompson, W. C., 479, 591–593
 Thompson, W. L., 643
 Thornberry, T. P., 389
 Thornton, D., 256, 257, 392, 393, 845, 847–849, 854
 Thurber, S., 621
 Thurstone, L. L., 24
 Tiihonen, J., 807
 Tillbrook, C., 61
 Tillbrook, C. E., 299, 300, 306
 Timm, H. W., 687, 690
 Tippins, T., 163
 Tobey, A. E., 596
 Toch, H., 21, 22, 26, 27, 446, 454
 Todd, C., 542
 Toedter, L., 240
 Toglia, M. P., 525, 563
 Tollestrup, P., 528
 Tollestrup, P. A., 528
 Tolman, A. O., 426
 Tomasi, D., 238
 Tomkins, A. J., 5, 28
 Tomlinson, M. H., 763
 Tomz, J. E., 453, 454, 459
 Tong, L. S. J., 803
 Torres, C., 462
 Torres, S., 462
 Toth, S., 239, 241, 243, 246
 Toth, S. L., 570
 Traux, P., 823
 Tredoux, C. G., 517, 525, 532–534, 540
 Trgovac, M., 818
 Trickett, P., 239, 246, 258
 Trickett, P. K., 241, 246, 254, 584
 Triebel, K., 226, 227
 Tripoli, L., 490
 Tripp, A., 338
 Trompetter, P., 443
 Trope, Y., 524
 Trout, R., 476
 Troxel, N. R., 561, 593
 Truelove, R. S., 529
 Tsukiura, T., 643
 Tullis, R., 282
 Tully, E., 243, 254
 Tulving, E., 530
 Turnbull, S., 337
 Turner, R., 332
 Turner, R. C., 294
 Turvey, C. L., 795
 Tversky, A., 473
 Twain, Mark, 469
 Twentyman, C. T., 243
 Tyner, E., 818
 Ullman, D., 406
 Undeutsch, U., 27
 Upadhyaya, M., 238, 239, 260
 Useem, B., 765, 768
 Usher, J. A., 566
 Ustad, K. L., 298, 303
 Vaez-Azizi, L. M., 197
 Vaidya, C. J., 564
 Valentine, D., 243
 Valentine, T., 526, 527, 531, 545, 549
 Valentine, T. Y., 545
 Valentino, K., 570
 Valliant, P. M., 802
 van Beek, D., 406
 VandeCreek, J., 711
 VandeCreek, L., 88, 95, 385, 426
 VandeCreek, L. D., 175
 Vandenberg, B. R., 102
 Vandenboom, C., 641
 van den Brink, R. H. S., 407
 van der Heijden, K., 415
 Van der Keilen, M., 476
 van der Merwe, L., 415
 van der Schrier, J. H., 527
 Van der Veen, S., 809
 Vandewalle, M. B., 336
 van Dijk, J. J. M., 839
 Van Hasselt, V. B., 462
 van Hiel, A., 818
 Van Hoof, J. C., 636
 van Ijzendoorn, M., 241
 Vann, C. R., 291
 van Os, T. W. D. P., 407
 Van Rybroek, G. J., 338
 Van Voorhis, P., 778
 Varghese, F., 796, 801, 804, 806
 Veglia, F. V., 850
 Veiling, B., 244
 Veltkamp, L. J., 813
 Venn, J., 663
 Verbrugge, P., 764
 Verdun-Jones, S. N., 282
 Veronen, L. J., 47
 Verschuere, B., 623, 625, 631, 637, 650
 Verstijnen, I. M., 527
 Vess, J., 847
 Veysey, B. M., 797
 Victor, E., 479
 Victor, T. L., 102
 Vidmar, N., 470, 472, 479, 481, 485, 499, 500
 Viljoen, J., 297
 Viljoen, J. L., 43, 281, 294, 297, 402, 406
 Villalba, D., 528
 Villetaz, P., 765
 Vincent, G. M., 297, 407
 Vingoe, F. J., 666
 Violato, C., 839
 Viscusi, W. K., 470, 473, 481
 Vitacco, M. J., 299, 300, 338
 Viteles, M. S., 23
 Voerman, A. E., 453
 Vohs, K. D., 850
 Volanti, J., 451
 Volkow, N., 238
 Vossel, G., 646
 Vrij, A., 625, 646, 650
 Wachtell, M. F., 68, 69
 Wack, R. C., 386
 Waddell, C. W., 365
 Wade, K. A., 536
 Wadley, V., 226
 Wagenaar, W. A., 527, 545
 Waggoner, J., 847
 Wagner, A. D., 645

- Wagner, E. F., 820
Wagstaff, G. F., 692
Wahler, R., 241
Waid, W. M., 623
Wain, H. J., 694
Wakeling, H. C., 850
Walcott, D. M., 385, 426
Wales, D., 851
Wales, H. W., 318
Walker, C., 332
Walker, D. K., 212
Walker, L., 8, 40
Walker, N., 339
Walkey, F. H., 818
Wall, B. W., 289
Wallace, M., 214
Wallace-Capretta, S., 802
Walle-Olsen, A., 587
Wallerstein, J. S., 147, 149, 151
Walsh, W. A., 597
Walters, G. D., 766, 800, 818
Walters, R. H., 27
Wamboldt, A. D., 796, 806, 807
Wang, C., 237
Wang, E. W., 809
Warberg, B., 845
Ward, T., 801, 840, 852, 855
Ware, J., 850
Warmelink, L., 646
Warner, T. D., 229, 230
Warren, A. R., 595
Warren, J., 189
Warren, J. I., 287, 290
Warren, K. L., 585
Warshak, R. A., 149
Washburn, J. J., 811
Wasyliw, O. E., 328
Waterman, J., 257, 266
Waters, N. L., 471, 472, 481
Waters, T., 814
Waters, W. F., 289
Watkins, J. G., 666
Watkins, R. E., 19–22
Watson, D., 243, 254
Watson, R. J., 844
Watt, K. A., 396, 411, 416, 847, 848
Waxman, T., 341–343
Weber, N., 527, 541, 542
Webert, D. R., 531
Webster, C. D., 282, 289, 294, 296, 394–396, 403, 406, 408, 409, 411, 413, 416, 419, 771, 846, 847
Webster, D., 238, 256
Webster, J. M., 798
Webster, S. D., 850
Webster-Stratton, C., 245, 246
Weeks, J. C., 233
Weinberger, L. E., 806
Weiner, I. B., 93, 101, 103, 721, 723, 728
Weiss, B., 778
Weiss, D. S., 454
Weiss, R., 263
Weisz, J. R., 778
Weitz, S. E., 180, 185, 187, 189
Weitzenhoffer, A. M., 682
Weld, H. P., 12
Welfel, E., 426
Wells, D. L., 47
Wells, G., 41, 49
Wells, G. L., 481, 532, 533, 535, 537, 538, 540, 541, 545, 546, 588
Wells, M. T., 471, 472, 476, 481
Welsh, A., 592, 593
Welsh, B. C., 777
Welsh, W., 42
Werth, J., 426
Wessel, E., 591
Wessely, S., 330, 331
West, S. G., 823
Wester, W., 660, 661, 680, 682, 692–694
Wester, W. C., 669
Westhuis, D. J., 805
Wetmore, S. A., 517
Wettstein, R. M., 386
Wexler, H. K., 799, 805
Whalen, R., 631, 632
Wheatman, S. R., 339, 340
Wheeler, A., 587
Wheeler, M. C., 288, 291
Whipple, E. E., 245, 246
Whipple, G. M., 5, 8
Whitbeck, L., 241
White, E. K., 455
White, P. A., 297
White, S. G., 409
White, T., 238
White, T. W., 812–814
Whitehead, G. M., 482
Whitehead, J. T., 777
Whitehouse, W. G., 680, 685
Whittemore, K., 337, 339, 342
Whittemore, K. E., 281
Wickens, T. D., 518
Widom, C., 241, 254, 255
Widows, M. R., 188
Wiederanders, M. R., 335, 338
Wiener, R. L., 474
Wiersema, B., 254
Wiersma, D., 407
Wierzbicki, M., 799
Wigmore, J. H., 10, 13
Wilczek, A. P., 476
Wilde, G. J., 778
Wildman, R. W., 297
Wiley, D. C., 474
Wiley, T. R. A., 589–591
Wilhelm, F. H., 851
Wilkinson, I. D., 643
Willett, L. R., 476
Willging, T. E., 13
Williams, A. J., 590
Williams, B., 238
Williams, C., 186
Williams, C. W., 186
Williams, G., 537
Williams, J. M. G., 570
Williams, L., 243
Williams, M. W., 798
Williamson, P., 800
Willner, P., 808
Wilson, A. E., 839
Wilson, C., 43, 385
Wilson, C. M., 407
Wilson, J. C., 585
Wilson, R. J., 841
Wilson, S., 797, 798, 800
Wilson, S. C., 682
Wilson, S. K., 623
Windmann, S., 571
Wines, E. C., 760–763
Wingfield, N., 470
Winick, B. J., 281, 285, 286, 293
Winick, C., 12
Winograd, M., 641
Winter, R., 476, 478
Winters, K. C., 330
Wiocik, P., 238
Wise, E. A., 823
Wissler, R. L., 471, 473, 476, 480, 481
Wistrich, A. J., 481
Witt, P. H., 180, 185, 187, 189
Witte, T., 809
Witte, T. D., 819
Wittmann, J., 163
Wodjdac, V., 632
Wolf, A. M., 777
Wolfe, D., 242
Wolff, M. A., 484

- Wolff, N., 806
 Wollert, R., 847
 Wolters, G., 527
 Wong, F., 149
 Wong, G., 232
 Wong, K. W., 819
 Wong, S., 392, 809
 Wong, S. C. P., 402, 760, 808, 809, 817, 819
 Wood, B., 584
 Wood, J., 451, 455
 Wood, J. M., 578
 Woodman, R., 372
 Woodruff, P. W. R., 643
 Woods, P., 392
 Woody, W. D., 476, 478
 Woolard, J., 303, 305
 Woolard, J. L., 303, 305
 Wooldredge, J., 768
 Wooldredge, J. D., 767
 Wooten, A., 517
 Wooten, S. E., 796, 806, 807
 Wootton, D., 541
 Workman, S., 232
 Worling, J., 811
 Worling, J. R., 396, 406
 Wormith, J., 402
 Wormith, J. S., 20, 22, 760, 767, 768, 781, 796, 800, 815, 817
 Wormith, S., 766
 Wormith, S. J., 392, 401, 767
 Worring, R. W., 668
 Wortley, R., 765
 Wright, D. B., 525, 527, 589, 592, 593
 Wrightsman, L. S., 43, 45
 Wu, C., 241
 Wuensch, K. L., 339, 340
 Xiao, N., 587
 Yamaguchi, K., 643
 Yang, C., 246, 254, 256
 Yang, M., 402, 760
 Yasuhara, K., 407
 Yates, P., 842, 843
 Yeager, M. H., 210
 Yee, D. S., 802
 Yell, M. L., 198, 199, 207
 Yellowlees, P., 820
 Yeomans, M., 407
 Yessine, A. K., 818, 819
 Yildirim, Y., 814
 Yim, I. S., 568, 569
 Yochelson, S., 773
 Young, K., 581
 Young, S., 42, 251, 252
 Young, S. G., 525
 Youngblade, L., 243
 Younggren, J., 711
 Youngs, D., 43
 Yozwiak, J. A., 589–591
 Ysseldyke, J., 203
 Yuille, J. C., 528
 Yurgelun-Todd, D. A., 643
 Zajac, G., 42
 Zakrajsek, T., 805
 Zamble, E., 767–769, 804
 Zamrini, E., 226
 Zanis, D. A., 800
 Zapf, P., 50
 Zapf, P. A., 281, 286–288, 290, 291, 294, 296, 297, 303, 305–307, 408, 428
 Zasler, N., 186
 Zaylor, C., 812–814
 Zeig, J. K., 693
 Zeisel, H., 481
 Zelig, M., 450, 682
 Zember, E., 575
 Zeoli, S., 849
 Zheng, Y., 643
 Zhu, J., 643
 Zinger, I., 759, 773, 777, 796, 797, 804, 848, 853
 Zoder, I., 765
 Zohar, J., 181
 Zubek, J. P., 780
 Zubrycki, R. M., 763
 Zuravin, S. J., 245, 254, 260

Subject Index

- Abuse, 237
 - Academic careers, 120
 - Achenbach Child Behavior Checklist, 158
 - A Concise Restatement of Torts*, 172
 - Actuarial prediction method, 392
 - calibration/derivation, 394
 - confidence intervals, 393
 - cross-validation, 393, 394
 - definition of, 391
 - probability estimates, 395
 - sample dependence, 393
 - sample-specific risk factors, 393
 - unit (equal) weighting, 394
 - VRAG, 392, 394, 395
 - weighting issue, 394
 - Actus reus, 354–356
 - Addington v. Texas*, 400, 423
 - Addressing Sexual Offenses and Misconduct by Law Enforcement*, 461
 - Administrative segregation (AS), 778, 780–781
 - Admissibility of expert evidence, 116
 - Adults, eyewitness testimony
 - confidence–accuracy correlations, 541
 - DNA testing, 517
 - exonerations, 516
 - eyewitness identification, 516
 - informational and decisional factors
 - composites, 548
 - decision criterion, 519, 521
 - innocent suspect, 520
 - level of risk, 518
 - lineup composition, 549–551
 - lineup/photospread, 519
 - match-to-memory strength, 520
 - nonidentifications, 549
 - perpetrator, 520
 - repeated procedures, 550
 - SDT, 517, 518, 522
 - witness instructions/expectations, 552
 - witness’s decision, outcomes of, 519
 - Northrop/Davis case, 547–548
 - postevent factors
 - blind administration, 538
 - composites, 531–532
 - consistency, witness reports, 530
 - identification procedure presentation format, 538–540
 - inconsistent statements, 530
 - in-court identification, 538, 539
 - instructions to, 537–538
 - interviewing techniques, 530
 - lineup composition, 532–535
 - memory-refreshing techniques, 531
 - photospread identifications, 539
 - postevent information, 535–536
 - repeated identification procedures, 536
 - sequential lineups, multiple “laps” in, 540–541
 - “showup”, 538, 539
 - simultaneous lineup, 538
 - time delay, 529
 - practical considerations for
 - evidence of identity, 545–546
 - expert training and role, 543–545
 - nonidentifications, 546
 - sources of information, 546–547
 - witness attributes, criminal events
 - age, 526
 - alcohol/drug intoxication, 527–528
 - attentional capture, 529
 - cross-race recognition effect, 525
 - disguises, 527
 - distance, 527
 - duration, 527
 - experience with faces, 524
 - familiar/stranger faces, 524–525
 - gender, 525–526
 - lighting, 527
 - personological attributions, 524
 - stress and weapon presence, 528–529
 - transient attributes, 526–529
 - witness confidence and response latency, 541–543
- Adversarial system, 711, 712, 718

- Advocacy and hypnosis, 694–696
 AFCC, 144
 Alienated children, 149
 American Academy of Forensic Psychology, 61, 123
 American Association for Correctional Psychology (AACP), 22
 American Board of Forensic Psychology, 28, 35, 57
 American Board of Professional Psychology (ABPP), 443
 American Board of Psychiatry and Neurology, 57
American Journal of Law and Medicine, 75
American Jurisprudence (Am. Jur.), 74
 American Law Institute (ALI), 64, 356
 American Medical Association (AMA), 661, 690–691
 American Psychiatric Association, 661
 American Psychological Association (APA), 3, 61, 129, 138, 144, 446, 447, 649
 American Psychological Association, 2010, 204
 American Psychology–Law Society (AP-LS), 35, 61
 American Society of Clinical Hypnosis (ASCH) Task Force, 660, 661
 Americans with Disabilities Act (ADA), 448, 451
 Anger management, 803
 Antisocial associates, 801
 Antisocial cognitions, 801
 Antisocial personality, 801
 APA-accredited internships, 122
 APA Committee on Urban Initiatives, 446
 APA Ethics Codes
 advertising and public statements, 97
 assessment standard, 100
 beneficence and malfeasance, 87
 competence standard
 child custody, 91
 courtroom performance, 92
 cross-examination, 92
 expert testimony, 92
 forensic practitioner, 91
 forensic psychology, 90
 forensic roles, 90
 insanity, 94
 insurance policies, 90
 reasonably certain, 94
 scientific and professional knowledge, 90
 statutory law, 93
 strongly suggestive, 94
 education and training, 98
 fidelity and responsibility, 87
 human relations, 95
 integrity, 87
 justice, 87
 people's rights and dignity, 87
 privacy and confidentiality, 95
 record keeping and fees, 98
 research and publication, 99
 resolving ethical issues, 88–90
 therapy standard, 102–104
 Applied developmental training, 120
 Aptitude by treatment interactions, 215
 Assessment of Capacity to Consent to Treatment (ACCT), 233
 Association of Family and Conciliation Courts (AFCC), 46, 138
Atkins v. Virginia (2002), 59
 Attachment theory, 571
 Atypical Presentation (AP) scale, 299
 Audiotapes, 684
 Automaton defense, 355

Barefoot v. Estelle (1983), 92, 128
Barefoot v. Estelle, 420
Barfield v. State, 665
 Basic mental health services, 795
Baxstrom v. Herold, 388
 Behavioral deep freeze model
 correctional climate, 768–769
 crowding, 768
 prison victimization, 769
 Behavioral modification, 803
Behavioral Sciences & the Law, 28
Best Practices for Forensic Mental Health Assessments, 74
 Binding authority, 69–70
BMW of North America v. Gore, 480
 Board certification in forensic psychology, 129
 Board Certified Behavior Analyst, 202
 Bootstrapping method, 633, 635, 638
 Boucher, R.J., 386–388
 Boulder model, 113, 126
 Brief Symptom Inventory (BSI), 400
 British Medical Association, 661
Brown v. Board of Education (1954), 15, 80
 Bush, Jack, 773

 California, diminished capacity doctrine in, 358–366
 California Rule of Court 5.220, 160
 California Senate Bill 54, 366
 California Senate Bill 590, 367
 Camberwell Family Interview, 781
 Canadian correctional system, 20
 Capacity evaluations
 assessment techniques and tools, 223–224
 barriers-conducting evaluations, 223

- civil law
 - doctrine *see* Informed consent doctrine, 228
 - financial capacity, 225–227
 - guardianship, 224–225
 - health-care proxy, 231
 - MacCAT-T, 232, 233
 - testamentary capacity, 227
 - unstructured clinical judgment, 232
- legal standards, 223
- Capacity to Consent to Treatment Instrument (CCTI), 232
- Caregiver evaluation
 - child relationship, 259
 - children evaluation, 265–267
 - child visitation, 259–260
 - clinical interview, 251–254
 - external validity, 250–251
 - informed consent, 249
 - interview content
 - misperceptions about a child, 255
 - physical abuse, 254–256
 - sexual abuse, 256–257
 - interviewing children, 267–268
 - neglectful parents, 257
 - observations and home visits, 260
 - parental self-report data, 261
 - protective factors, 258–259
 - psychological measures, 261–263, 268–269
 - records and collateral reports, 263–265
 - risk assessment, 254
 - social isolation, 260
- Case conference model, 717
- Case law, 63
- Central Intelligence Agency (CIA), 666
- Child/children
 - abuse and neglect, 570
 - children's lie-telling, development of, 585
 - eyewitness testimony *see* Children's eyewitness testimony, 561
 - forensic interviewers, 572
 - intelligence, 580–581
 - maltreatment
 - PTSD, 571
 - maltreatment
 - attachment theory, 571
 - dissociation, 570
 - McMartin case interviews, 578
 - memory accuracy, 566
 - memory development, 562–563
 - memory for traumatic and stressful events, 564
 - parent–child conversation, 567
 - verbal ability, 581–582
 - witnesses, 563
- Child advocacy centers (CACs), 596–597
- Child bystanders, 586
- Child custody and parenting evaluations
 - alienation, 137
 - attachment, 141
 - best interests of the child, 139
 - child abuse, 137
 - child development, 137, 139
 - children, 137
 - child's voice, 142
 - coercive control dynamics, 140
 - collateral witnesses, 143
 - conflict dynamic, 137
 - conflict resolution, 140
 - court, 141
 - domestic violence, 137, 140
 - estrangement, 142
 - ethical practice
 - AFCC Model Standard 5.1 states, 146
 - AFCC Model Standard 5.5, 145
 - APA Guidelines, 146
 - conflict, 148
 - conflicts of interest, 144
 - decision making, 148
 - divorce research, 147
 - informed consent, 145
 - legal custody, 147, 148
 - mental illness, 148
 - overnight, 148
 - parenting coordinators, 149
 - parenting plans, 147
 - physical custody, 147
 - record keeping, 147
 - residential schedule, 148
 - shared, 50–50, or sole physical custody, 148
 - specialized competence, 144
 - training, 144
 - transparency, 147
 - visitation, 148
 - ethical practice APA Guidelines, 145
 - ethical practice multiple relationships, 144
 - family, 142
 - friendly parent provision, 139
 - interviews, 137
 - licensing complaints, 139
 - loyalty conflict, 142
 - mediation, 142
 - mental health problems, 142
 - Model Standards state, 138
 - overnight, 143
 - parenting plan, 139
 - personality dynamic, 137
 - process
 - attorneys, 152
 - children's suggestibility, 155
 - collateral information, 158–159
 - coparenting, 157

Child custody and parenting evaluations

(continued)

- interviews with children, 154
 - interviews with parents, 152–154
 - parents and children together, 155
 - psychological testing, 156–158
 - psychological evaluations, 138
 - psychological testing, 137
 - quality of parenting, 140
 - rebuttable presumption, 140
 - relocation law, 137
 - report writing, 139
 - analysis, 162
 - children, 161
 - collateral information, 162
 - parent, 161
 - procedures, 161
 - recommendations, 163
 - Specialty Guidelines state, 160
 - research, 139
 - alienated children, 149
 - domestic violence, 149
 - intimate partner violence, 149
 - power and control, 149
 - relocation evaluations, 150
 - restrictive gatekeeper, 151
 - risk and protective factors, 151
 - sexual abuse, 150
 - sexual abuse, 139
 - shared parenting, 143
 - substance use/abuse issues, 140
 - time-sharing plans, 138
 - traits, 140
- Child maltreatment
- biological and genetic aspects, 238
 - child protective service investigation, 247
 - criminology, 238
 - developmental and psychological factors
 - child factors, 244
 - contextual factors, 244–247
 - parental factors, 240–244
 - social support, 240
 - economic and social impact, 239
 - etiology and impact, 237
 - evaluation methodology
 - care and protection evaluations, 248
 - caregiver *see* Caregiver evaluation, 248
 - multimodal, 248
 - genetic, 238
 - intervention method, 238
 - legal codes/social service systems, 238
 - macrosocial factors outweigh supports, 239
 - multidetermined nature, 247
 - multimethod approaches, 270

- no single/uniform solution, 239
 - potentiating microsocial, 239
 - protective factors, 238
 - psychology, 238
 - public health problem, 237
 - risk factors, 238
 - severity and chronicity, 247
 - sociology, 238
 - static and dynamic causal agents, 237
 - taxometric structures, 239
 - United States, 237
- Child molesters, 840–842
- Child protective service agencies, 120
- Child protective service records, 263
- Child protective service system, 258
- Children's eyewitness testimony
- accommodations for
 - child advocacy centers, 596–597
 - child witness trauma, 593
 - out-of-court testimony, 593–595
 - disclosure of abuse
 - emotional expression, 584
 - factors affecting, 583
 - face recognition and identifications
 - contact hypothesis, 587
 - cross-racial effect, 587
 - face processing, 586–587
 - lineup fairness, 587–589
 - "mugshot commitment", 586
 - own-race bias, 587
 - "unconscious transference", 586
 - individual differences in
 - intelligence, 580–581
 - verbal ability, 581–582
 - jurors' reactions to, 589–591
 - child-victim characteristics, 589
 - child abuse cases, expert witnesses in, 591–593
 - child victim gender, 589
 - perceptions of children's honesty, 589
 - lying, 584–585
 - maltreatment effects and trauma-related
 - psychopathology, 570–571
 - memory development, 562–563
 - suggestibility, false reports, and false memory, 571–572
 - fuzzy-trace theory, 573
 - misinformation, 573
 - misleading questions, 576–579
 - "Mr. Science", 574
 - repeated interviews, 579–580
 - "Sam Stone", 576
 - source monitoring theory, 574
 - theoretical issues, 572–575
 - younger and older children, 575

- trauma, stress, and memory, 563–564
 attachment-avoidant individuals, 565
 central *vs.* peripheral details, 568
 children's age, 566–567
 “defensive exclusion”, 565
 infantile amnesia, 566
 language and parental communication, 567
 parenting style, 567
 participant *vs.* bystander child witnesses,
 568–569
 physiological distress, 569–570
 repeated events, 569
 survival-based processing, 565
 theoretical issues, 564–566
- Child suggestibility, 121
- Civil capacity
 competence, construct of, 221–222
 evaluations *see* Capacity evaluations, 222
 history, 220–221
 mental health practitioners, 219
- Civil litigation, 472
 expert witness, 175
 process
 burden of proof, 174
 discovery, 173
 expert disclosure, 174
 interrogatories, 173
 preponderance of the evidence, 174
 standard of proof, 174
 trier of fact, 175
- Clinical psychology, 120
- Closed circuit television (CCTV), 595–596
 6th and 14th Amendments, 595
 vividness effect, 596
- Coalition to Save Our Children v. State Board of
 Education*, 1995, 208
- Coca-Cola Company v. Chero-Cola Company*, 13
- Cognitive-behavioral theory (CBT), 119, 773,
 802–803
- Cognitive Interview, 530
- Cognitive psychology program, 121
- Cognitive restructuring, 803
- Committee for the Recognition of Specialties
 and Proficiencies in Professional
 Psychology, 35
- Commonwealth v. Rogers* (1844), 319
- Commonwealth v. Urena*, 486, 487
- Commonwealth v. White*, 66
- Compensatory damages, 471
- Competence Assessment for Standing Trial for
 Defendants with Mental Retardation
 (CAST-MR), 303–304
- Competency Assessment Instrument
 (CAI), 295
- Competency Screening Test (CST), 294
- Competency to stand trial
 adequate screening procedures, 287
 CAST-MR, 303–304
 community-based settings, 286
 competency assessment tools
 CAI, 295
 CST, 294
 ECST-R, 299
 FIT, 296
 GCCT, 297
 IFI, 295
 ILK, 300
 MacCAT-CA, 298
 concurrent validity, 291
 criminal proceedings, 281
 criterion-related validity, 291
 examiner–judge agreement, 291
 functional evaluation approach, 292–293
 JACI, 304–305
 legal issue, 288
 medication, 289
 mental health contacts, 305
 mental health professionals, 282
 meta-analysis, 290
 populations
 criminal justice system, 301
 intellectual disabilities, 301, 302
 juvenile competency, 303
 MR, 301, 302
 psycholegal abilities, 283
 psycholegal deficits, 290
 rational manner, 291
 screening instruments, 286
 sell decision, 289
 Supreme Court decisions, 283
 vignette methodology, 291
- Complex trial protocol (CTP), 641, 642
- Confabulation, 682, 687, 688
- Confidence interval (CI), 533
- Congregate system, 761, 762
- Consult with Counsel (CC) scale, 299
- Contingency management programs, 772
- Contingent negative variation (CNV), 642
- Control question technique (CQT), 615, 622
 computerized recording systems, 618
 global and numerical chart evaluation, 618
 physiological recordings, 617
 posttest phase, 618
 pretest interrogation, 616–617
 validity of, 625–626
- Core Correctional Practices (CCPs), 854
- Corpus Juris Secundum* (C.J.S.), 74
- Correctional climate (CC), 768–769
- Correctional Program Assessment
 Inventory–2010© (CPAI-2010©), 759

- Correctional psychology, 4
 AACP, 22
 administrative segregation, 780–781
 Canadian correctional system, 20
 classification of inmates, 19
 “congregate” system, 761
 effects of imprisonment
 behavioral deep freeze, 767–769
 deterrence theory, 764–766
 schools of crime theory, 766–767
 imprisonment, history of, 760–761
 Jesness Classification System, 21
 knowledge cumulation process, 780
 meta-analysis, 772
 Minnesota Multiphasic Personality
 Inventory (MMPI), 21
 North America, developments in,
 761–762
 nothing-works proponents, 778
 perceptual adaptation, 780
 prediction and treatment
 contingency management programs, 772
 CPAI-2010©, 772
 HCR-20, 771
 LSI-R, 770
 misconducts, 770–772
 PCL-R, 771
 recidivism, 770–772
 segregation rates, 772
 prison management
 confinement, special conditions of,
 778–779
 correctional residential treatment programs,
 774
 Ohio officials, 774
 Rideau Correctional Centre, 773
 “routine practice” programs, 777
 Vermont program, 773–774
 prison psychologists, 19, 20
 rehabilitation, 21
 rehabilitative ideal, 762–763
 “separate” regime, 761
 trenches, 22
 unit management, 22
 Correctional supervision, 795
 Corrections and Law Enforcement Family
 Support (CLEF) program, 453, 454
 Corsini’s estimation, 19
 Counseling, 129
 Counseling psychology, 121
 Courtroom testimony, 7–9
Crawford et al. v. Honig, 201
 Crime victims, 47
Criminal Behavior and Mental Health, 28
Criminal Justice and Behavior, 27
 Criminal offenders
 correctional interventions, effectiveness of,
 795–797
 criminal associates, 803
 evidence-based practices
 antisocial associates, 801
 antisocial cognitions, 801
 antisocial personality pattern, 801
 cognitive-behavioral theory, 802–803
 heuristics, 803
 homework, 803–804
 intensiveness, 804–805
 risk factors, 801
 structure, 804
 therapist cultural competence, 805
 victim impact, 803
 jails/prisons, 795
 risk assessment
 actuarial estimate, 815
 antisocial personality, 815
 criminogenic needs, 815
 dynamic risk factors, 815
 effective correctional intervention,
 816
 integrated-actuarial approach, 815
 static risk factors, 815
 special populations
 intellectual disabilities, 808
 juvenile offenders, 805
 mental illness, 805–808
 substance abuse problems, 805
 violent offenders, 805, 808–809
 telehealth, 811–814
 therapeutic change
 behavioral indicators, 819
 clinical significance testing, 822–823
 disciplinary infractions, 819
 effect sizes, 822
 multiconstruct approach, 819–820
 multimethod approach, 817–819
 multi-outcomes approach, 821
 multi-time approach, 820–821
 parole/probation, 819
 pre–post significance testing,
 821–822
 reliability of change, 823
 security classification, 819
 therapeutic alliance, 820
 treatment
 attrition, 799
 motivation and therapeutic resistance,
 800–801
 service utilization, 797–799
 treatment dropouts, 799–800
 Criminal psychology, 4, 26–27

- Criminal responsibility, 723, 724, 727
 empirical developments
 judicial instruction research, 339–340
 NGRI verdicts *see* Not guilty by reason of
 insanity (NGRI) verdicts, 335
 research *see* Jury and Juror decision making,
 340
- insanity defense
 Insanity Defense Reform Act (IDRA), 367,
 368
 reform and fate of diminished capacity,
 366–372
- insanity standards and construal
 fairness and justice, 316
 GBMI, 322
 legal standards evolution, 319–322
 mens rea, 316–319
 social interactions, 316
 MSO *see* Mental state at the time of the offense
 (MSO), 324
- Criminogenic needs, 48
- Cross-examination, 751–755
- Cross-race recognition effect, 525
- Crum v. State*, 493
- Curriculum vitae, 747
- Daubert v. Merrell Dow Pharmaceuticals, Inc.*
 (1993), 60, 116, 648, 677, 736, 738, 748
- Davis, Larry, 514, 515, 547–548
- Day v. Kilgore*, 490, 491
- Degree and nondegree training opportunities
 continuing education programs, 123
 internship training, 121
 Master's degree in forensic psychology, 124
 PhD/PsyD programs, 121
 psychology and joint degree programs, 122
 respecialization and postdoctoral training, 123
- Denis v. Commonwealth*, 665
- Densen v. Stanley*, 494
- Dereliction, 172
- Desistence, 795
- Deterrence theory, 764–766
- Developmental psychology program, 121
- Diana v. State Board of Education* (1970), 200
- Didactic and experiential experiences, 121
- Didactic and experiential training components,
 125
- Digital videoconferencing programs, 811
- Diminished capacity, 353
 actuality of intent, 376–377
 actus reus and mens rea, 354–356
 California, 358–366
 forensic evaluations of, 372–379
 genuine cases of, 374
 history of, 353
 insanity defense reform and, 366–372
 intent, general and specific, 356–358
 intoxication defense, 356
 no opinion/negative opinion, 377–378
 premeditation and deliberation, 363
- Diminished responsibility, 364
- Directed lie technique (DLT), 618–619, 626
- Direct examination, 750–751
- District courts, 67
- Divorce research, 147
- Dixon v. Attorney General*, 389
- Dred Scott v. Sandford*, 79
- Drope v. Missouri*, 288
- Durham v. United States* (1954), 320
- Dusky v. United States* (1960), 282, 296
- Dynamism, 741–742
- Early intervention system (EIS), 450
- Educational disabilities
 compliance monitoring, 206
 federal law, 198
 litigation
 placement bias litigation, 200–201
 right to education, 199
 school psychologists, 197
 second-generation litigation and legislation,
 206
 state and federal legislation
 due process procedural safeguards, 203–204
 evaluations and reevaluations, 204–205
 free appropriate education, 201–202
 individualized education program, 202
 LRE, 202
 SWD, 209
 system reform trends, 214
- Education of the Handicapped Act (EHA), 198
- Effect size
 Cohen's d, 822
 partial eta squared statistics, 822
 treatment impact, 822
- Eigenface models, 531
- Eighth Amendment, 59, 70
- Element of a tort, 172
- Emmerich v. Philadelphia Center for Human
 Development*, 426
- Empirical research on forensic psychology
 training, 131
- Employee Polygraph Protection Act (EPPA), 614
- Employment, screening/evaluation
 relevant/irrelevant technique, 620–621
 test for espionage and sabotage, 621
- Estelle v. Smith* (1981), 288
- Estimate of Risk of Adolescent Sexual Offense
 Recidivism (ERASOR), 406

- Ethical Principles of Psychologists and Code of Conduct (EPPCC), 86, 143, 176
- Ethics, 120
- Ethics Codes
- APA, 87
 - aspirational in nature, 86
 - bligatory rules, 86
 - constraining and the supportive features, 86
 - ethical obligation, 86
 - minimize practitioners' vulnerability, 108–109
 - personal values, 104–105
 - professional practitioners, 85
 - professional responsibility, 106–108
 - professions derive, 85
 - psychological practice, 86
 - research, 86
 - teaching, 86
- Evaluation of Competency to Stand Trial—Revised (ECST-R), 299
- Event-related potentials (ERP), 627
- Bayesian method, 635, 638
 - bootstrapping method, 633, 635, 638
 - CNV, 642
 - CTP, 641, 642
 - impact of countermeasures, 639–642
 - MERMER method, 636, 637
 - P300-based GKT procedure, 630–642
 - reaction time measure, 634
- Exhibitionism, 843
- Experiential learning opportunities, 126
- Expert testimony
- admissibility of, 421–423, 737
 - cross-examination, 748
 - doll experiments, 16
- Federal Rules of Evidence
- Daubert v. Merrell Dow Pharmaceuticals*, 736
 - FRE 104, 736
 - FRE 609, 753
 - FRE 613, 752
 - FRE 615, 746
 - FRE 701, 734
 - FRE 702, 734
 - FRE 703, 735
 - FRE 704, 735
 - FRE 705, 735
 - FRE 803(18), 753, 754
 - Kumho Tire Co. v. Carmichael*, 736
 - Frye v. United States*, 736
 - guardianship proceeding, 748
 - influential decision, 14
 - juvenile courts, 13
 - “learned treatises”, 753
 - legal decision makers, 38–39
 - mental health professions, 14
 - redirect and recross examination, 754–755
 - Reference Manual on Scientific Evidence*, 738
 - school segregation, 15
 - social science brief, 15
 - strategies and approaches
 - cross-examination, 751–755
 - direct examination, 750–751
 - sworn testimony, 739
 - credibility and trustworthiness, 741
 - depositions, 742–744
 - dynamism, 741–742
 - effective communication, 739–740
 - expertise, 740
 - trials, 744
 - testifying in court
 - board certifications, 747
 - curriculum vitae, 747, 748
 - licensure, 747
 - pretrial meeting, 744
 - qualification colloquy, 749
- Expert witness, 175, 591–593
- Externships, 126
- Exxon Shipping Co. v. Baker*, 471
- Eyewitness identifications, 121
- Eyewitness testimony
- of adults *see* Adults, eyewitness testimony, 513
 - of children *see* Children's eyewitness testimony, 563
- Face composite, 514
- Face recognition, 585–589
- FACES, 531
- Factual Understanding (FU) scale, 299
- Faculty members, 124
- Family dynamics, 142
- Family Education Rights and Privacy Act, 203
- Federal Bureau of Investigation (FBI), 445, 668, 669
- Federal Rules of Civil Procedure (FRCP), 173, 719
- Federal Rules of Evidence (FRE), 60, 368
- FRE 104, 736
 - FRE 609, 753
 - FRE 613, 752
 - FRE 615, 746
 - FRE 702, 39, 175, 734
 - FRE 703, 175, 735
 - FRE 704, 735
 - FRE 705, 735
 - FRE 803(18), 753, 754
- Fischer v. Fischer*, 497, 498
- Fitness-for-duty evaluations (FFDE), 450–451
- Fitness Interview Test (FIT), 296
- Forensically informed manner, 157

- Forensic assessment instruments (FAIs), 287
 - MES, 326
 - R-CRAS, 327
- Forensic clinical programs, 120
- Forensic clinical services, 124
- Forensic mental health assessment (FMHA)
 - capital mitigation, 59
 - child custody, 59
 - legal concepts, 58
 - mental health professionals, 57, 60, 62, 74
 - methodological approach, 58
 - professional standards, 57
 - scientific approaches, 62
 - shaping conclusions and communication, 60
- Forensic psychological assessment, 121
- Forensic psychological ethics, 127
- Forensic psychology, 111
 - academic, 38
 - adversarial method, 37
 - American Psychology–Law Society (AP-LS), 35
 - cognitive and personality assessment
 - court clinics, 18
 - court psychologist, 18
 - intelligence testing, 18
 - performance measures, 17
 - correctional, 4, 19
 - criminal psychology, 4, 26–27
 - descriptive, 37
 - education and training programs, 28
 - experimental method, 37
 - forensic psychiatry, 5
 - graduate programs, 120
 - interdisciplinary, 28
 - legal psychology, 4, 5
 - literature and research explosion, 27
 - nomothetic, 37
 - police, 22
 - police psychology, 4
 - prescriptive law, 37
 - proactive, 38
 - probabilistic evidence, 37
 - professional practice, 4
 - research endeavor, 4
 - Specialty Guidelines for Forensic Psychology, 4
 - specialty training programs, 120
 - stare decisis, 36
 - trainees, 120
 - training, 131
- Forensic reports, writing
 - addressing future circumstances, 724–725
 - addressing past circumstances, 723
 - addressing present circumstances, 722–723
 - being clear, 725–726
 - being defensible, 728–730
 - being informative, 727–728
 - being relevant, 726–727
 - case conference model, 717
 - client, preferences of, 713
 - medical malpractice, 713–714
 - molestation, 714–715
 - psychological evaluation, 715
 - confidentiality, 718
 - consultations
 - child custody, 724
 - closed head injury, 722
 - competency, 722
 - confession, 720
 - criminal responsibility, 723, 727
 - diminished capacity, 720
 - parole, 724
 - personal injury, 722, 727
 - probation, 724
 - schizophrenia, 720
 - temporary insanity/diminished capacity, 723
 - visitation rights, 724
 - custody cases, 727
 - ethical principles and professional realities
 - considerations in practicing law, 716–717
 - considerations in practicing psychology, 717–718
 - expert opinions, quality of, 715–716
 - posttraumatic stress disorder, 718
 - experts' written reports, 719
 - Federal Rules of Civil Procedure, 719
 - focus of, 719–725
 - Specialty Guidelines for Forensic Psychology, 720
 - “work product”, 713
- Forensic rotation, 122
- Forensic skills, 120
- Formal legal research, 71–73
- Fourteenth Amendment, 69
- Fourth Amendment, 69
- Fraser v. U.S.*, 426
- Frye v. United States*, 11, 60, 421, 647, 671, 736, 748
- Full-Scale Intelligence Quotient (FSIQ), 722
- Functional capacity, 179
- Functional magnetic resonance imaging (fMRI), 643–645
- Fuzzy-trace theory (FTT), 573
- Galvanic skin response (GSR), 614
- Gender, witness's report, 525–526
- Generalist clinical, 129
- General Electric Co. v. Joiner*, 1997, 116
- General personality and cognitive social learning (GPCSL), 414

- Georgia Court Competency Test (GCCT), 297
Godínez v. Moran (1993), 283, 295
 Good Law, 78–81
 Good Lives Model (GLM), 855
 Governmental agency, 124
Government of the Virgin Islands v. Charles, 284
 Government policy analyst, 120
Gratz et al. v. Bollinger et al., 207
 Grisso's model, 221
Gutter v. Bollinger et al., 2003, 207
Guadalupe Organization v. Tempe Elementary School District No. 3 (1972), 200
 Guardianship, 224–225
 Guardianship law, 126
 Guardianship proceeding, 749
 Guidelines for Child Custody Evaluations in Family Law Proceedings, 86, 138
 Guidelines for Parenting Coordination, 47
 Guilty knowledge test (GKT), 619–620, 623, 627–628
- Harding v. State*, 669
Harrington v. State of Iowa, 631
 Health-care proxy, 231–232
 Hearsay evidence, 592, 594–595
 Heuristics, 803
Hidden v. Mutual Life Insurance Co., 15
 Hinckley case, 367
 Historical-Clinical-Risk Management–20 (HCR-20), 401–407, 771
 Homework compliance, 804
 Hypnosis, 531, 671
 advocacy, 694–696
 antisocial conduct, 659
 crimes committed by hypnotized subjects, 664–666
 crimes committed on hypnotized subjects, 663–664
 CIA, 666
 courtroom advocacy, 660
 defendants
 audiotapes, 684
Daubert v. Merrell Dow Pharmaceuticals, Inc., 677
 fairness, 683–684
Leyra v. Demmo, 674
Newman v. Hopkins, 676
 nonhypnotizable subjects, 685
People v. Ritchie, 675
 per se exclusion rule, 678–681
Rock v. Arkansas, 674, 675
Rowland v. Commonwealth, 678
 sexual seduction, 684
State v. Papp, 675
 therapeutic hypnosis, 684
 time delays, 684
 totality of the circumstances rule, 681–683
United States v. Scheffer, 676
 hypnotic memory enhancement, 685–688
 “concreting/hardening” effect, 689–690
 confabulation, 687–689
 motivation, 687
 pseudomemories, 687–689
 “involuntary”, 666
 judicial rulings, 668–669, 673–674
 guidelines test, 673–674
 open admissibility rule, 669
 per se exclusion rule, 669–673
 “totality of the circumstances” test, 673–674
 for memory recall, 659, 666–668
 mental health care, legal standard of, 660, 692–694
 Model Penal Code 2.01(2)(c), 665\$
 Montana Crimes Code 45–2–101(33)(c), 665\$
People v. Ebanks, 667
 regulation of, 659
 avocational/vocational self-improvement, 662
 Business & Professions Code 2908, 662\$
Hohe v. San Diego Unified School District, 663
Johnson v. Gerrish, 662
 licensed health-care professional, 662
Masters v. State, 661
People v. Cantor, 662
 retroactive hallucinations, 667
 scientific community
 American Medical Association, 690–691
 expert opinion, 691–692
- Identikit, 531
 Importation *see* Behavioral deep freeze model, 767
 Imprisonment
 effects of
 behavioral deep freeze model, 767–769
 deterrence theory, 764–766
 schools of crime theory, 766–767
 history of, 760–761
Indiana v. Edwards, 285
 Individuals with Disabilities Education Act (IDEA), 198
 Informal legal research, 73–76
 Informed consent doctrine
 adequate disclosure, 229–230
 competency, 230–231
 litigation and subsequent case law, 228
 OHRP, 229
 unethical and ethically troubling practices, 228
 voluntariness, 230

- Injury severity, 476
In re L.A.M., 2001, 112
 Insanity Defense Reform Act (IDRA), 367, 369, 371
 Intellectual disabilities (IDs), 808
 Intensive aftercare programs (IAPs), 811
 Intent, 353
 actuality of, 376–377
 general and specific, 356–358
 larceny, 357
 malice, 357, 358
 rape, 357
 Intentional/negligent, 173
 Interdisciplinary Fitness Interview (IFI), 295
 International Association of Chiefs of Police (IACP), 43, 446, 448, 452, 453, 461
 Internships, 126
 Interpersonal problem-solving skills, 803
 Intoxication defense, 356
 Intrafamilial trauma, 252
 Inventory of Legal Knowledge (ILK), 300
 Irresistible impulse test, 358
- Jackson v. Indiana* (1972), 288
Jenkins v. United States, 16, 734
 Jesness Classification System, 21
Johnson v. Gerrish, 662
Journal of Forensic Psychology Practice, 28
Journal of Psychiatry and Law, 75
 Judiciary Act of 1789, 68
 Judiciary Act of 1869, 68
 Jury(ies)
 decision making, 470–473
 “anchoring and adjustment”, 473
 “goal managers”, 473
 jurors’ case-relevant attitudes, 341–342
 jurors’ case construals, 342
 jurors’ implicit theories/prototypes, 343
 parallel constraint satisfaction, 473
 punitive damage awards, 470
 “story model”, 473
 hindsight bias, 475
 influences on decisions
 complex civil trials, 478–479
 conduct of defendant/plaintiff, 474–475
 “corporate identity”, 478
 deep-pockets effect, 478
 defendant’s wealth, 478
 individual characteristics of decision makers, 479–480
 injury severity, 476
 jurors versus judges, 480–481
 outcome severity, 476–477
 legal reform efforts, 471–472
 Wisconsin study, 486
- Jury decision-making aids, 482–483
 access to trial transcripts, 501–502
 note taking
 advantages of, 494–496
 disadvantages of, 496–499
 empirical research, 493
 procedures, 493–494
 predebate discussions, 483, 499–501
 right to question witnesses, 484–492
 advantages of, 487–488
 alert trial counsel, 488–489
 disadvantages of, 489–492
 empirical research, 484–487
 written witness summaries, 502–503
 Juvenile justice system, 811
 Juvenile measures of competency (JACI), 304–305
 Juvenile offenders, criminal offenders
 intensive aftercare programs, 811
 juvenile justice system, 811
 multisystemic therapy, 810
 Juvenile offending, 120
- Kansas v. Crane*, 421
Kansas v. Hendricks, 421
KeyCiting, 79
Kumho Tire Co. v. Carmichael, 736
- Larceny, 357
Larry P. v. Riles, 200
 Law
 administrative code, 58
 binding vs. persuasive authority, 69–70
 case law, 58
 clinician ignorance, 61
 Eighth Amendment’s prohibition, 59
 FMHA *see* Forensic mental health assessment (FMHA), 60
 forensic mental health professional, 59
 FRE, 60
 functional legal capacities, 59
 idiographic data, 59
 insufficient and incredible testimony, 62
 mental health professional, 59
 nomothetic data, 59
 primary sources
 case law, 63
 constitutions, 63
 statutory law, 63
 Tenth Amendment, 62
 reading and interpreting, 76–78
 secondary sources, 64
 statutes code, 58
 synthesizing, 81

- Law (*continued*)
 ultimate legal issue, 60, 61
 United States courts, 65–69
Law and Human Behavior, 28
Law and Mental Health Professionals, 74
Law and Psychology Review, 27, 75
Law and the Lawyers, 13
Law and the Social Sciences, 13
 Law enforcement, psychology and
 assessment domain
 fitness-for-duty evaluations, 450–451
 LEAA, 448
 M-PULSE, 449
 POST approach, 449
 preemployment screening, 448, 449
 psychological suitability, 448
 validation data, 449
 clinical intervention domain
 family services, 453–454
 Montreal Police Service, 452
 “role constriction theory”, 451
 stress in, 454–455
 “suicide by cop”, 452
 training, 455–456
 trauma, 452–453
 confidentiality, 460
 culture shift, key events signifying,
 445–447
 APA, 446, 447
 APA-sponsored Police Chiefs Roundtable
 Series, 446
 FBI Training Academy, 445
 IACP-PPSS, 446
 NIJ, 446
 ethical dilemmas, 460
 evolution and gaining acceptance,
 444–445
 new and emerging trends
 CLEF funds, 462
 family orientation programs, 462
 IACP, 461
 LAPD, 461
 LEOSS, 462
 Police Chief Roundtables, 461
 PSHP, 463
 operational support, 457–458
 organizational consulting, 456–457
 service delivery models, 458–459
 Law Enforcement Assistance Administration
 (LEAA), 448
 Law Enforcement Officer Stress Survey
 (LEOSS), 462
 Learning heuristics, 803
 Least restrictive environment (LRE), 202
Lee v. State, 422
 Legal actors
 attorneys, 43–45
 crime victims, 47
 criminal profiling, 43
 jury consultants, 45
 jury selection, 44
 law enforcement, 42–43
 offenders, 47
 quasi-judicial roles
 mediation, 45
 parent coordination, 46–47
 trial consultants, 44
 voir dire, 44
Legal and Criminological Psychology, 28
 Legal psychology, 4
 European origins
 courtroom testimony, 7–9
 lie detection, 9
 psychological expert testimony, 9
 psychology of testimony, 7
 expert testimony
 doll experiments, 16
 influential decision, 14
 juvenile courts, 13
 mental health professions, 14
 plaintiff’s mental condition, 15
 school segregation, 15
 social science brief, 15
 United States
 academic psychological literature, 13
 adversarial process, 12
 applied psychology, 9, 10, 12
 criminal justice system, 12
 human endeavor, 9
 hypnosis and lie detection, 10
 jury system, 12
 origins, 5
 polygraph, 11
 psychological landscape, 9
 yellow psychology, 10
 Legal research
 electronic legal databases, 71–73
 informal legal research, 73–76
 mental health professional, 70
 Legal system
 actors *see* Legal actors 104; 254, 42
 decision makers, 38–39
 litigation—specific research, 39–40
 research, 41–42
 psycholegal constructs, 49–50
 psychological phenomena, 49
 social authority testimony, 40
Leon v. McDonald, 423
Lessard v. Schmidt, 424

- Level of Service/Case Management Inventory (LS/CMI), 401, 815
- Level of Service Inventory–Revised (LSI-R), 395, 770, 815
- LexisNexis, 71
- Leyra v. Denno*, 674
- Lie detectors *see* Polygraph testing, 629
- Lineup composition, 549–551
E', 534
 “effective size”, 534
 lineup bias, 533
 lineup size, 533
 mock witness evaluation, 533
- Lineup fairness, 532, 587–589
- Live testimony, 595
- Los Angeles Police Department (LAPD), 25, 461
- M’Naghten Rule, 359, 361, 362
- Münsterberg, H., 517, 734
- MacArthur Competence Assessment Tool, Treatment (MacCAT-T), 232
- MacArthur Competence Assessment Tool–Criminal Adjudication (MacCAT-CA), 298
- Machismo, 805
- Maggart v. Bell*, 490
- The Manchurian Candidate*, 666
- Mancuso v. Consolidated Edison, Co.*, 2000, 112
- Manson v. Braithwaite* (1977), 535
- Marriage of LaMusga, In re*, 141
- Mason v. Underwood*, 694
- Master’s degree in forensic psychology, 124
- Masters of Legal Studies (MLS), 122
- Masters v. State*, 661
- Matrix-Psychological Uniform Law Enforcement Selection Evaluation (M-PULSE), 449
- Matthews v. Commonwealth Edison Co.*, 497
- Mattie T. v. Johnson*, 209
- McCleskey v. Kemp*, 130
- Measures of Criminal Attitudes and Associates (MCAA), 818
- Mediation, 45
- Memory-refreshing techniques, 531
- Memory and encoding related multifaceted electroencephalographic response (MERMER), 636, 637
- Memory development, 562–563
- Mens rea*, 118, 354–356, 372–379
- Mental health care, legal standard of, 692–694
- Mental health law, 120
- Mental Health Law and Policy Journal*, 75
- Mental illness, criminal offenders
 antisocial attitudes, 806
 Changing Lives and Changing Outcomes, 807
 criminal and psychiatric recidivism, 806
 criminal thinking, levels of, 806
 criminogenic risk, 807
 juvenile offenders, 809–811
 mental health recovery, 807
 preliminary field testing, 807
 psychiatric hospitalization, 807
- Mental State at the time of offense screening evaluation (MES), 326
- Mental state at the time of the offense (MSO)
 delusions
 definition and conceptualizing, 331–332
 NGRI defendants, 330
 radical beliefs, 330
 forensic assessment instruments *see* Forensic assessment instruments, 326
 interview
 inception, 324
 psychopathology, 325
 psychotropic medications, 325
 reconciliation and termination, 326
 reconnaissance, 325
 third-party information, 329
 treatment and release of insanity acquittees, 333–335
- Metropolitan Nashville Police Department, 463
- Millon Clinical Multiaxial Inventory-III (MCMI-III), 157
- Mills v. Board of Education*, (1972), 199
- Minnesota Multiphasic Personality Inventory (MMPI), 21
- Minnesota Multiphasic Personality Inventory–2 (MMPI-2), 99, 157, 175
- Missouri v. Davis* (1983), 292
- Model Penal Code (MPC), 356, 357
- Model Standards of Practice for Child Custody Evaluation, 138
- Model Standards of Practice for Family and Divorce Mediation*, 46
- Montana v. Austad* (1982), 292
- Moral reasoning, 803
- Morrow v. Maryland* (1982), 292
- “Mugshot commitment”, 586
- Muller v. Oregon*, 15
- Multisystemic therapy (MST), 810, 811
- National Advisory Commission on Civil Disorder, 448
- National Association of School Psychologists, 204
- National Center for Credibility Assessment (NCCA), 615
- National Council of Juvenile and Family Court Judges, 149
- National Institute of Justice (NIJ), 446

- National Institute of Occupational Safety and Health (NIOSH), 447
- National Research Council (NRC), 629–630
- Neglect, 237
- Negligence, 173
- Newman v. Hopkins*, 676
- Nonclinical forensic services, 124
- Nonforensic psychological researchers, 115
- Not guilty by reason of insanity (NGRI) verdicts
characteristics, 337
frequency of, 336
rate of agreement, 336
- North Dakota v. Heger* (1982), 288
- Northrop, Alan, 514, 515, 547–548
- O'Connor v. Donaldson*, 424
- Offender Group Reconviction Scale (OGRS), 402
- Office for Human Research Protections (OHRP), 229
- Office of Special Education Programs (OSEP), 206
- Omnibus Crime Act, 446
- Omnibus Crime Control & Safe Street Act, 448
- Ontario Psychological Association, 20
- On the Witness Stand*, 734
- Open admissibility rule, 669
- Organizational consulting activities, 456–457
- Out-of-court testimony, 593–595
- Own-race bias, 587
- Own group bias (OGB), 525
- Pacific Mutual Life Ins. Co. v. Haslip*, 470
- Parent Child Relationship Inventory, 158
- Parenting Stress Index–4, 157
- Parents in Action on Special Education (PASE) v. Hannon*, 1980, 201
- Pate v. Robinson*, 288
- Peabody Picture Vocabulary Test (PPVT), 582
- Peace Officer Standards and Training Commission (POST), 449
- Peck v. The Counseling Service of Addison County*, 417
- Pedophilia, 840
definition of, 840
DSM-III-R diagnosis, 841
- Pelwani v S.*, 532
- Pennsylvania Association for Retarded Children (PARC), *v. Commonwealth of Pennsylvania*, 199
- People v. Baldi*, 665
- People v. Brown*, 66
- People v. Cantor*, 662
- People v. Cantrell*, 362
- People v. Caro*, 685
- People v. Carpenter*, 377
- People v. Conley*, 361
- People v. Drew*, 364
- People v. Ebanks*, 667
- People v. Gorshen*, 360, 361
- People v. Harris* (1983), 306
- People v. Hawthorne*, 14
- People v. Henderson*, 360
- People v. Marsh*, 665
- People v. Ritchie*, 675
- People v. Shirley*, 670, 672
- People v. Stevens*, 422
- People v. Therrian*, 422
- People v. Ward*, 421, 422
- People v. Wells*, 359
- People v. Wetmore*, 363, 366
- People v. Wheeler*, 1992, 112
- People v. Williams*, 670–671
- People v. Wolff*, 361
- People v. Worthington*, 664
- Percent of maximum possible (POMP) scores, 822
- Per se exclusion rule
defendants, 678–681
judicial rulings
People v. Shirley, 672
People v. Williams, 670–671
Rock v. Arkansas, 671
State v. Mack, 670
- Personal injury evaluations
attorneys, 183
communication
courtroom testimony, 193–194
depositions, 193
expert disclosure, 191
FRCP Rule 26, 191–192
emotional distress, 171
ethical issues
misuse, prevention of, 178
objectivity, 176
reports and testimony, 178
roles and role conflict, 176–178
examinees, 183–184
interpretation
addressing causality, 190
establishing damage, 189–190
legal context
civil litigation *see* Civil litigation, 173
tort law, 171–173
plaintiff evaluation
clinical interview, 186
collateral data, 188–189
data collection, 185

- documents and background information, 185–186
- psychological testing, 187–188
- psychological damage *see* Psychological damage, 179
- psychological injury, 171
- tort, 171
- Personality Assessment Inventory (PAI), 157, 188
- Personnel screening, 628–629
 - relevant/irrelevant technique, 620–621
 - test for espionage and sabotage, 621
- Personological attributions, 524
- Persuasive authority, 69–70
- Phallometric assessments
 - exhibitionists, 845
 - rapists, 844
 - sadists, 845
- Phillip and Angie C. v. Jefferson Board of Education*, 204
- Physiological distress, 569–570
- Plaintiff's deposition, 186
- Plessy v. Fergusson* (1896), 80
- Police Psychological Services Section (PPSS), 446, 448, 457
- Police psychology, 4
 - cognitive and aptitude screening, 23–24
 - cognitive assessment, 23
 - personality assessment, 24
 - police personality, 23
- Police stress and health program (PSHP), 463
- Policy-, law-, or litigation-driven forensic research, 116
- Polygraph testing
 - admissibility, 647–648
 - applications, 613–614
 - detection of deception
 - ERP-based detection methods, 630–643
 - neuroimaging-based detection methods, 643–646
 - thermal imaging, 646
 - voice stress analysis, 647
 - incident investigations, 615–620
 - juries evaluation, 648–649
 - personnel screening, 620–621, 628–629
 - scientific opinion, 649–650
 - techniques
 - control question technique, 615–618
 - directed lie technique, 618–619
 - guilty knowledge test, 619–620
 - relevant/irrelevant technique, 620–621
 - test for espionage and sabotage, 621
 - training/administration, 614–615
 - validity
 - chart evaluation, 621–622
 - control question technique, 625–626
 - directed lie technique, 619–626
 - field *vs.* laboratory investigations, 622–623
 - guilty knowledge test, 627–628
 - NRC's conclusion regarding lie detection accuracy, 629–630
 - problems establishing ground truth, 623–624
- Portland Digit Recognition Test, 188
- Positron emission tomography (PET), 643
- Postdoctoral clinical fellowships, 123
- Postdoctoral training, 123
- Postevent factors
 - blind administration, 538
 - composites, 531–532
 - identification procedure presentation format, 538–540
 - inconsistent statements, 530
 - instructions to, 537–538
 - interviewing techniques, 530
 - lineup composition, 532–535
 - memory-refreshing techniques, 531
 - postevent information, 535–536
 - repeated identification procedures, 536
 - sequential lineups, 540
 - time delay, 529
- Postevent misinformation, 562
- Posttraumatic stress disorder (PTSD), 174, 452, 453, 571, 718
- Practice-oriented training, 114
- Practice career, 120
- Predicting Violent Behavior: An Assessment of Clinical Techniques*, 846
- Predoctoral internship programs, 122
- Process of civil litigation, 173–175
- Professional training programs, 121
- Pseudomemory, 689
- Psychological damage
 - causality, 181–183
 - emotional functioning, 179–181
- Psychological injury, 171
- Psychological Inventory of Criminal Thinking Styles (PICTS), 818
- Psychologists in the Criminal Justice System*, 22
- Psychology, Crime, & Law*, 28
- Psychology, Public Policy, and Law*, 28, 75
- Psychology Applied to Legal Evidence and Other Constructions of Law*, 10
- Psychology for the Lawyer*, 13
- Psychology of criminal behavior, 120
- Psychology of criminal conduct (PCC), 48, 414
- Psychopathy, 49
- Psychopathy Checklist Screening Version™(PCL SV), 771

- Psychopathy Checklist–Revised (PCL-R), 401, 402
- Psychopathy Checklist–Revised Youth Version (PCL YV), 402
- Psychopathy Checklist–Revised™(PCL-R), 771
- Punitive damage awards, 470
- Quasi-judicial roles, 45–47
- R. v. Shatford*, 532
- Rape, 357
- Rapid Risk Assessment for Sex Offender Recidivism (RRASOR), 846
- Rapists, 357, 842–843
- Rational choice model, 765
- Rational understanding (RU) scale, 299
- Reaction time (RT), 634
- Reasoning and rehabilitation (R&R) program, 802
- Receiver operating curves (ROCs), 630
- Recidivism, 795, 799
- Reference Manual on Scientific Evidence*, 738
- Regina v. Whittle*, 283
- Rehabilitation, 795
- Relapse prevention skills, 803
- Relevant/irrelevant technique (RIT), 620–621
- Reliable Change Index (RCI), 823
- Repeated events, 569
- Repeated identification procedures, 536
- Respecialization in clinical psychology, 123
- Response to intervention (RTI), 214
- Retroactive hallucinations, 667
- Rideau Correctional Centre, 773
- Riggins v. Nevada* (1992), 289, 295, 325
- Risk-need-responsivity (RNR), 48, 759, 796, 801
- Ritchie v. Indiana* (1984), 292
- Rock v. Arkansas*, 671, 674, 675
- Rockwood Program for Sexual Offenders, 853–855
- Rogers Criminal Responsibility Assessment Scales (R-CRAS), 327
- Rorschach, 157
- Rowland v. Commonwealth*, 678
- Royal Canadian Mounted Police (RCMP), 622, 625
- S-1 v. Turlington*, 201
- Scenario planning, 415, 416
- Schaefer v. St. Louis & Suburban R. Co.*, 487
- Schall v. Martin*, 420
- Schizophrenia, 720
- School psychology, 121, 129
- School segregation, 15
- Schools of crime theory
- prisonization, 766
 - recidivism, 767
 - risk level, 766
- Scientist-practitioner model, 126
- Second-generation litigation and legislation
- equal results, 207
 - equal treatment, 208
 - fairness, 207–208
 - overrepresentation, 208–209
 - placement bias and overrepresentation, 207
 - right to education and LRE, 206–207
- Self-Appraisal Questionnaire (SAQ), 817, 818
- Self-hypnosis, 684
- Self-regulation, 850–851
- Sell v. United States* (2003), 289
- Seltzer v. Hogue*, 424
- Sequential lineups, 540–541, 588
- Service delivery models, 458–459
- Service utilization, 797–799
- Sex Offender Risk Assessment Guide (SORAG), 394, 401, 407
- Sex offenders, 802
- assessments
 - cognitions, 849–850
 - criminogenic factors, 848
 - relationship problems, 850
 - risk assessment, 846–848
 - self-regulation, 850–851
 - sexual factors, 851–852
 - civil commitment of, 421
 - diagnostic and assessment issues
 - child molesters, 840–842
 - exhibitionists, 843
 - rapists, 842–843
 - DSM criteria, 843–845
 - treatment
 - needs principle, 853
 - outcome evaluations, 855–856
 - responsivity principle, 853
 - risk principle, 853
 - Rockwood Program for Sexual Offenders, 853–855
- Sexual abuse, eyewitness testimony, 561
- Sexual offending, 256
- Sexual seduction, 684
- Shepardizing*, 79
- Shipley Institute of Living Scale, 720
- Short-Term Assessment of Risk and Treatability (START), 406, 407
- Shu-king*, 760
- Signal detection theory (SDT), 517, 518, 522
- Simultaneous lineups, 538, 588
- Sixth Amendment, 69
- Sligar v. Bartlett*, 492

- Social authority testimony, 40
 Social skills, 803
 Society for Clinical and Experimental Hypnosis, 692
 Sociocultural issues in forensic psychology, 120
 Solitary confinement *see* Administrative segregation (AS), 778
 Source monitoring (SM) theory, 574
 Specialty Guidelines for Forensic Psychology (SGFP), 4, 35, 57, 86, 119, 143, 176, 281, 720
 Specialty training, 128
 Stabilization, 795
Stare decisis, 69
State of Maryland v. Gary Smith, 645
State v. Alexander, 486
State v. Armstrong, 674
State v. Driver, 14
State v. Fertig, 680
State v. Greer, 486
State v. Hurd, 673
State v. Jones, 65, 66
State v. Kendall, 488
State v. Mack, 670
State v. Mena, 687
State v. Moore, 680
State v. Papp, 675
State v. Pike (1869), 319
State v. Rameau, 695
State v. Sickles, 491
State v. Triplett, 496
States v. Sahhar, 424
 Statistical Information on Recidivism (SIR), 401, 818
 Statutory law, 63
 Stigma, 798
 Stress
 eyewitness memory, 528–529
 in law enforcement, 454–455
 Structured Assessment of Protective Factors for Violence Risk (SAPROF), 406
 Structured Assessment of Violence Risk in Youth (SAVRY), 402, 403, 406, 407, 427
 Structured Child Assessment of Relationships in Families (SCARF), 156
 Structured Interview of Reported Symptoms (SIRS), 188
 Structured Inventory of Malingered Symptomatology (SIMS), 188
 Structured professional judgment (SPJ), 396, 409–411, 847, 848
 communicate findings, 416
 criminal offenders, 407
 dynamic risk factors, 397
 forensic patients, 406
 HCR-20, 406, 407
 idiographic level, 397
 incremental validity, 406
 individual relevance, 397
 logical/rational item selection, 396
 management plans, 416
 mentally disordered offenders, 406
 nomothetic level, 397
 priori weighting, 397
 psychiatric inpatients, 406
 quantity and quality of information, 411–412
 risk factors
 decision theory framework, 414
 presence of, 412–413
 relevance of, 413–414
 risk factors, 410–411
 risk reduction strategies, 397
 SAVRY, 406, 407
 scenarios of violence, 415–416
 SORAG, 407
 START, 407
 VRAG, 407
 Students with disabilities (SWD)
 educational programs, 201
 educational services, 199
 etiology
 biologically based disabilities, 210
 biological/functional behavior, 210
 functional/behavioral deficits, 210
 high-incidence disabilities, 212
 low- and high-incidence disabilities, 210
 psychological services, 211
 SEA and LEA variations, 212–214
 federal and state legislation, 198
 federal and state litigation and legislation, 199
 IDEA, 209, 210
 individualized educational program, 202
 influenced education services, 199
 LRE, 202
 public schools, 199, 201
 response to intervention, 210
 SLD, 209
 special education services, 198
 USDE, 198
 “suicide by cop”, 452
 Suicidal ideation, 798
 Supervised practica, 126
 Supervision styles, 125
 Supreme Court, 68
 Sworn testimony
 communication
 credibility and trustworthiness, 741
 dynamism, 741–742
 effective communication, 739–740
 expertise, 740

- Sworn testimony (*continued*)
 depositions, 742–744
 trials, 744
- Sympathetic nervous system reactivity, 570
- Tarasoff v. Regents of the University of California*, 362, 425
- Target-absent lineup, 588
 Target-present lineup, 588
- Telehealth, 811
 benefits of, 812
 cost effectiveness, 813
 family counseling, 814
 female offenders, 814
 forensic mental health assessments, 813
 personality disorders, 814
 therapeutic alliance, 813
 treatment outcome research, 814
 use of, 812
- Telemedicine *see* Telehealth, 811
- Telepsychology *see* Telehealth, 812
- Tenth Amendment, 62
- Testamentary capacity, 227–228
- Test for espionage and sabotage (TES), 621, 629
- Therapeutic hypnosis, 684
- Third-party data, 188
- Thornton v. Weaver*, 496–498
- Timothy W. v. Rochester*, 201
- Tort law, 171–173
- Totality of the circumstances test
 guidelines test, 673–674
 logical errors, 681–683
- Training, law enforcement, 455–456
- Training programs
 accreditation of forensic psychological training, 129
 clinical practitioner-scientist, 114
 clinical psychology with a subspecialty in forensic psychology, 112
 clinical scientist-practitioner, 113
 credentialing in forensic psychology, 128
 degrees/dual JD–PsyD (doctor of psychology) degrees, 112
 didactic and experiential training opportunities, 125
 divergent programmatic approaches, 117
 empirical research, 131
 expert testimony, 113
 faculty as advisors and mentors, 125
 faculty expertise and student goals, 124
 forensic clinical practitioners
 clinical programs trainees, 119
 cognitive-behavioral therapy, 119
 continuing education courses, 117
 dangerousness, 119
 ethical problems, 119
 externship placements, 120
 forensic clinical services, 120
 forensic coursework, 120
 forensic psychological assessment, 118
 forensic psychological services, 119
 forensic treatment, 119
 general clinical training, 118
 insanity evaluation, 118
 licensure, 120
 on-the-job training, 117
 postdoctoral work, 117
 specialized forensic training, 119
 suicidality, 119
 forensic ethics training, 127
 forensic psychological competence, 128
 forensic psychology nonclinical practitioners, 120
 forensic psychology scientists challenges, 130
 forensic scientists, 120
 gerontological issues, 113
 JD (juris doctorate), 112
 legal assessment types, 127
 legal research and training skills, 126
 maintaining and increasing forensic psychological competence, 129
 nonclinical and clinical scientist-scholar cognitive, 114
 developmental, and social psychology, 114
 scholar, 117
 scientist, 115
 scientist-scholars, 114
 number of trainees, 125
 PhD (doctor of philosophy), 112
 PhD with a clinical specialization, 112
 PsyD, with a programmatic emphasis in clinical forensic psychology, 112
 suggestive pretrial identification procedures, 112
- Training programs degree and nondegree, 121
- Transtheoretical model of change (TTM), 819
- Trauma, law enforcement, 452–453
- Trial consultants, 44
- Trial consultation, 120
- “Twinkie defense”, 366
- U.S. Constitution’s 14th Amendment Equal Protection Clause, 130
- U.S. Constitution, 64, 68
- U.S. Supreme Court, 68
- U.S. v. Brown*, 370
- U.S. v. Cameron*, 370
- U.S. v. Childress*, 370
- U.S. v. Frisbee*, 368
- U.S. v. Gold*, 369

- U.S. v. Hood*, 368
U.S. v. Newman, 370
U.S. v. Pohlott, 369
U.S. v. Schneider, 370
U.S. v. Twine, 369
U.S. v. White, 1985, 368
 United Nations Convention on the Rights of the Child (UNCRC), 597
 United States Circuit Courts of Appeal, 68
 United States Courts, 65–69
United States v. Adams, 670
United States v. Ajmal, 484
United States v. Barnette, 423
United States v. Brawner (1972), 320
United States v. Brawner, 321
United States v. Bush, 486, 489
United States v. Callahan, 488
United States v. Darden, 492
United States v. Davis, 496, 498
United States v. Feinberg, 484, 486
United States v. Hall, 686
United States v. Johnson, 484, 489, 490
United States v. Jones, 65, 66
United States v. Lyons, 322
United States v. McCollum, 668
United States v. Norwood, 1996, 112
United States v. Phillips, 665
United States v. Polowichak, 486
United States v. Richardson, 486
United States v. Scheffer, 420, 648, 676
United States v. Stierwalt, 486
United States v. Thompson, 489
United States v. Torniero (1984), 322
United States v. Virginia (1996), 115
United States v. Wexler, 483
United States v. White, 66

 Vermont program, 773–774
 Victim impact, 803
 Victimology, 120
 Videotaped testimony, 595
 Violence Reduction Program (VRP), 809
 Violence Risk Appraisal Guide (VRAG), 392–395, 407, 846, 847
 Violence risk assessment
 “clear and convincing evidence”, 423
 and law
 admissibility of testimony, 421–423
 assessments of dangerousness, 420–421
 civil commitment decisions, 423–425
 insanity acquittees, 425
 sexual predator, 421
 Supreme Court, 420–421
 Tarasoff case and the duty to protect, 425–426
 comparative evaluative research
 clinical prediction, 398–400
 meta-analyses, 401–405
 structured professional judgment, 405–407
 comprehensive clinical risk assessments, 408–420
 contemporary models of
 actuarial prediction, 391–395
 clinical judgment, 390–391
 structured professional judgment, 395–398
 unstructured clinical discretion, 391
 research
 Baxstrom and Dixon studies, 388–390
 dangerousness, clinical judgments of, 387
 “predictions of violence”, 387
 recidivism, 388
 significant risk, 388
 research, 386–390
 summary risk ratings, 404
 Violence Risk Scale (VRS), 817, 819
 Violence Risk Scale: Sex Offender Version (VRS: SO), 819
 Violent offenders, criminal offenders
 components for, 809
 conflict resolution skills, 809
 VRP, 809
 Visitation, 259
 Voice stress analysis, 647
 Voir dire, 44

Washington v. Davis/Northrop, 547
 Weapon focus effect (WFE), 529
 Wechsler Adult Intelligence Scale–IV (WAIS-IV), 720
 Westlaw, 71
Wieter v. Settle, 282
Wilburn v. Reitman, 695
Williams v. Commonwealth, 487
Wilson v. United States (1968), 292
Witherspoon v. Illinois (1968), 116
 Witnesses
 eyewitness *see* Eyewitness testimony, 524
 juror questioning of, 470
 Working Alliance Inventory (WAI), 820
 Workplace Assessment of Violence Risk (WAVR-21), 409
 Workplace Violence Model, 447
 Writ of certiorari, 69
 Writ of habeas corpus, 69

 Yellow psychology, 10
 Youth Level of Service/Case Management Inventory (YLS/CMI), 402

Zippo Manufacturing v. Rogers Imports, 40

