PSYCHOLOGICAL EXPERTISE IN COURT

Psychology in the Courtroom, Volume II

Edited by
Daniel A. Krauss
Joel D. Lieberman

PSYCHOLOGY, CRIME AND LAW
Series Editor: David Canter

PSYCHOLOGICAL EXPERTISE IN COURT

PSYCHOLOGY, CRIME AND LAW SERIES

Series Editor: David Canter
International Centre for Investigative Psychology
The University of Liverpool, UK

Other titles in the Series

Jury Psychology: Social Aspects of Trial Processes
Psychology in the Courtroom, Volume I
Joel D. Lieberman and Daniel A. Krauss

Principles of Geographical Offender Profiling
David Canter and Donna Youngs

Applications of Geographical Offender Profiling
David Canter and Donna Youngs

Psychology and Law: Bridging the Gap
David Canter and Rita Žukauskienė

Safer Sex in the City *(forthcoming)*
David Canter, Maria Ioannou and Donna Youngs

Behavioural Analysis of Crime:
New Directions in Offender Profiling *(forthcoming)*
Donna Youngs

The Investigative Psychology of Serial Killing *(forthcoming)*
David Canter and Donna Youngs

Profiling Violent Crime *(forthcoming)*
David Canter and Donna Youngs

Psychological Expertise in Court

Psychology in the Courtroom, Volume II

Edited by

DANIEL A. KRAUSS
Claremont McKenna College, USA

JOEL D. LIEBERMAN
University of Nevada – Las Vegas, USA

ASHGATE

Published by
Ashgate Publishing Limited
Wey Court East
Union Road
Farnham
Surrey, GU9 7PT
England

Ashgate Publishing Company
Suite 420
101 Cherry Street
Burlington
VT 05401-4405
USA

www.ashgate.com

British Library Cataloguing in Publication Data
Psychological expertise in court. - (Psychology in the courtroom ; v. 2) (Psychology, crime and law series)
1. Evidence, Expert 2. Forensic psychology
I. Lieberman, Joel D. II. Krauss, Daniel (Daniel A.)
347'.067

Library of Congress Control Number: 2009924670

ISBN: 978-0-7546-7687-4 (hardback)
ISBN: 978-0-7546-9380-2 (ebook)

Printed and bound in Great Britain by
MPG Books Ltd, Bodmin, Cornwall.

Contents

Notes on Contributors

Desiree Cassar received her MA in psychology from John Jay College of Criminal Justice in 2000, and her PhD in social psychology from the Graduate School and University Center, at the City University of New York in 2005. Her areas of research include jury decision-making in capital cases and the effects of expert testimony on jury behavior. She has taught psychology at Berkley College in New York City, Claremont McKenna College in California and is now residing in Canada.

Mark Costanzo is Professor of Psychology and co-director of the Center for Applied Psychological Research at Claremont McKenna College. He received his MS and PhD from the University of California at Santa Cruz. He has published research on a variety of law-related topics including police interrogation, false confessions, jury decision-making, sexual harassment, attorney argumentation, alternative dispute resolution, and the death penalty. He is the author of *Psychology Applied to Law* (Wadsworth, 2004), and *Just Revenge: Costs and Consequences of the Death Penalty* (St. Martin's Press, 1997); and is co-editor (with Dan Krauss and Kathy Pezdek) of *Expert Psychological Testimony for the Courts* (Erlbaum, 2007).

Solomon Fulero is both a practicing attorney and a psychologist. Dr Fulero received his PhD in social psychology and his law degree from the University of Oregon in August 1979 and December 1979 respectively, and a respecialization certificate in clinical psychology from Wright State University in June 1988. He is Professor and former Chair of Psychology at Sinclair College in Dayton, Ohio and Clinical Professor of Psychology and Psychiatry at Wright State University in Dayton. He is the co-author of the Wadsworth/Thomson textbook *Forensic Psychology* (third edition), published in April 2008, as well as numerous scholarly articles in both psychology journals and law reviews. Dr Fulero is a Fellow of the American Psychological Association. He has been on the Executive Committee of the American Psychology-Law Society (APLS), was the APLS representative to the governing Council of Representatives of the American Psychological Association in 1999–2002, was a member of the APA Committee on Legal Issues from 2000 to 2003 (chair in 2002–3), and was President of APLS in 2003–4.

Sarah M. Greathouse received her MS in Legal Psychology from Florida International University and is currently working on her PhD in Forensic Psychology at the City University of New York. To date, her research has focused on issues relevant to jury decision-making and eyewitness identifications. She serves as the editorial assistant for *Law and Human Behavior*, the official journal of the American Psychology-Law Society.

Susan R. Hall, JD, PhD received both degrees for the University of Arizona joint degree program in psychology and law. She received her BA from Georgetown University. Her scholarship aims to help clinicians better understand the law, and examines the clinical and forensic needs of children and youth exposed to violence and maltreatment. For example, her newest book, *Courtroom Modifications for Child Witnesses: Law and Science in Forensic Evaluations* (with Sales, 2008), describes the steps a mental health professional should take in order to conduct forensic assessments of allegedly abused child witnesses regarding their needs for courtroom modifications. Other relevant publications include *Child Abuse Reporting Laws and Attorney-Client Privilege: Ethical Dilemmas and Practical Suggestions for the Forensic Psychologist* (2006), and *Child Therapy and the Law* (with Sales, DeKraai, and Duvall, 2007).

M. Alexis Kennedy practiced law in Vancouver, Canada before returning to complete her PhD in forensic psychology at the University of British Columbia. Her areas of research include child abuse, sexual assault, body image, and prostitution. Her doctoral dissertation on cross-cultural perceptions of child abuse won two American Psychological Association awards (Divisions 37 and 41). Dr Kennedy teaches at the University of Nevada, Las Vegas and is active on local boards and task forces addressing human trafficking, sexual exploitation, and juvenile justice issues.

Margaret Bull Kovera is Professor of Psychology at John Jay College of Criminal Justice, City University of New York. She received her PhD from the University of Minnesota and her BA from Northwestern University. She is a Fellow of the American Psychological Association, the American Psychology-Law Society (APLS), and the Society for the Psychological Study of Social Issues and is the Past-President of APLS. She received the Saleem Shah Award for Early Career Achievement in Psychology and Law and the APLS Outstanding Teacher and Mentor in Psychology and Law Award from APLS. For over a decade, she has had continuous funding from the National Science Foundation for her research on jury decision-making and eyewitness identification.

Daniel A. Krauss received a joint PhD in psychology (clinical psychology and psychology, policy and law) and law degree from the University of Arizona, and he received his BA from The Johns Hopkins University. He is currently Associate Professor of Psychology at Claremont McKenna College. His research explores the

use of psychological expertise in the legal system. Dr Krauss is the author of over 30 article and chapters related to the intersection of law and psychology, was a member of the APA Committee on Legal Issues from 2004 to 2007, and was the United States Supreme Court Fellow at United States Sentencing Commission in 2002–3.

Richard A. Leo is an Associate Professor of Law at the University of San Francisco School of Law. He received both his JD and PhD from the University of California at Berkeley. Dr Leo has written four books and more than 50 articles and book chapters on police interrogation practices, false confessions, and wrongful convictions. He is the author of *Police Interrogation and American Justice* (Harvard University Press, 2008), and (with Tom Wells) *The Wrong Guys: Murder, False Confessions and the Norfolk 4* (2008), and (with George Thomas) *The Miranda Debate: Law, Justice and Policing* (1998). Leo has won awards for research excellence and distinction from many organizations, including the American Society of Criminology, the American Psychological Association, and the American Academy of Forensic Psychology.

Lora M. Levett, PhD is an Assistant Professor in the Department of Sociology and Criminology and Law at the University of Florida. She received her PhD in Psychology from Florida International University in 2005. The goal of her research is to use psychological theory and methods to identify solutions to practical legal questions while contributing to basic knowledge about human behavior.

Joel D. Lieberman, PhD is Associate Professor and Chair of the Criminal Justice Department at the University of Nevada, Las Vegas. He received his PhD from the University of Arizona in Psychology. His work focuses on the application of social psychological theories to criminal justice issues. This research has led to numerous published articles and book chapters involving the identification of psychological factors that explain why jurors have difficulty comprehending judicial instructions, and why jurors rely on inadmissible evidence, defendant characteristics, and less reliable expert witness testimony, in their decision-making. In addition, he has authored a recent book on *Scientific Jury Selection* (with Bruce Sales).

Gianni Pirelli received his MA in psychology from The Graduate Center at John Jay College of Criminal Justice (CUNY) and is currently a fifth-year doctoral student in their Forensic Psychology PhD program (Clinical track). His areas of research include standards of practice in forensic psychology, forensic mental health assessment, psychometric properties of traditional and forensic assessment instruments, and clinical judgment and decision-making. He is currently Chair of the Student Section of the American Psychology-Law Society (AP-LS), Division 41 of the American Psychological Association (APA).

Netta Shaked-Schroer is a doctoral student in Applied Social Psychology at Claremont Graduate University. Her research focuses on jury decision-making

as it relates to the death penalty, false confessions, and most recently, forensic identification. In addition to her research, she is an instructor at Irvine Valley College.

Allison Strother is a recent graduate of Claremont McKenna College. She was a double major in Government and Economics-Accounting. She plans to attend law school in the fall.

Patricia A. Zapf, PhD received her doctoral degree in Clinical Forensic Psychology from Simon Fraser University in Canada in 1999. She is currently Associate Professor in the Department of Psychology at John Jay College of Criminal Justice, the City University of New York. She is an Associate Editor for *Law and Human Behavior* and was appointed Fellow of the American Psychological Association in 2006 for outstanding contributions to the field of law and psychology. She has published over 60 articles and chapters, mainly on the assessment of insanity and criminal competencies.

Tina Zottoli is graduate student in the Forensic Psychology program at the John Jay College of Criminal Justice, the City University of New York. Her research is focused on juvenile justice issues, with a primary emphasis on neuron-cognitive development and adolescent decision-making in legal contexts.

Series Preface

Over recent years many aspects of law enforcement and related legal and judicial processes have been influenced by psychological theories and research. In turn concerns that derive from investigation, prosecution and defence of criminals are influencing the topics and methodologies of psychology and neighbouring social sciences. Everything, for example, from the detection of deception to the treatment of sex offenders, by way of offender profiling and prison management, has become part of the domain of a growing army of academic and other professional psychologists.

This is generating a growing discipline that crosses many boundaries and international frontiers. What was once the poor relation of applied psychology, populated by people whose pursuits were regarded as weak and arcane, is now becoming a major area of interest for many students and practitioners from high school through to postgraduate study and beyond.

The interest spreads far beyond the limits of conventional psychology to disciplines such as Criminology, Socio-Legal Studies and the Sociology of Crime as well as many aspects of the law itself including a growing number of courses for police officers, and those associated with the police such as crime analysts or forensic scientists.

There is therefore a need for wide-ranging publications that deal with all aspects of these interdisciplinary pursuits. Such publications must be cross-national and interdisciplinary if they are to reflect the many strands of this burgeoning field of teaching, research and professional practice. The *Psychology, Crime and Law* series has been established to meet this need for up to date accounts of the work within this area, presented in a way that will be accessible to the many different disciplines involved.

In editing this series I am alert to the fact that this is a lively new domain in which very little has been determined with any certainty. The books therefore capture the debates inherent in any intellectually animated pursuit. They reveal areas of agreement as well as approaches and topics on which experts currently differ. Throughout the series the many gaps in our knowledge and present-day understanding are revealed.

The series is thus of interest to anyone who wishes to gain an up-to-date understanding of the interplays between psychology, crime and the law.

Professor David Canter

This page has been left blank intentionally

Preface to the Two-Volume Set

Joel D. Lieberman and Daniel A. Krauss

Over the past three decades there has been dramatic growth in the amount of research conducted that focuses on the application of psychology to legal issues. During that time period, researchers have examined issues related to evidence, testimony, and courtroom procedures with perspectives guided by their training in a variety of sub-disciplines within the field of psychology, including the areas of social, clinical, and cognitive psychology.

This two-volume set is designed to review the major areas in which psychology has been applied to court proceedings, to discuss relevant problems identified by psychological research, and to offer some perspective on how courts within and outside the United States endeavor to handle these problems. We have called upon a well-respected group of scholars to discuss psychological issues related to the procedures used by courts as well to specific types of evidence, and to specific types of trials.

The material presented in these two volumes is intended to offer an authoritative and comprehensive treatment of psychology in the courtroom. Recognizing that subdividing the chapters into two volumes has substantial benefits over one large tome, we are able to provide more expansive treatment of a wide variety of important topics by known experts in their respective fields. Such depth of coverage would not be possible in a single volume. Yet, at the same time, the two volumes afford the reader the opportunity to choose a single volume that offers exceptional depth on the topics in which he or she may be most interested.

Book Structure

The first volume, *Jury Psychology*, focuses on broad issues related to how courts and juries make decisions across a wide variety of topics with special emphasis placed on the application of psychological evidence and theory to factors that occur before and after the presentation of evidence in a standard trial, and to the nature of the trial itself. Issues from procedural justice to the psychology of jury instructions are examined in this volume. The second volume, *Psychological*

Expertise in Court, focuses on the use of psychological expert testimony and evidence in a diverse group of areas, from false confessions to sexual harassment.

Moving Beyond the Traditional Localized Focus of Research Application

In the process of conceptualizing the structure of these volumes, we realized that much of the research in psychology and law focuses on issues relevant to the United States. For example, research on jury instructions typically examines U.S.-based jury instructions; research on juror decisions in death penalty cases typically evaluates the impact of capital trial procedures in the U.S.; research in the area of defendant competency or insanity typically reviews U.S.-based criteria for such determinations. To some extent this ethnocentric focus is not surprising. Much of the relevant research has been conducted in the United States, thus researchers are simply more familiar (and concerned) with U.S.-based procedures and standards.

However, at the same time, the work in this area is generally not driven by a desire to understand U.S.-based procedures specifically, but rather from an interest in identifying problems and solutions in legal proceedings and decision-making from a psychological perspective. Psychology is based on cognitive and behavioral processes and responses, and not on national differences. Thus, many of the principles identified in U.S.-based research may be highly applicable to the legal systems of other countries as well.

We asked all of the contributors to these volumes to consider how the findings and conclusions they were writing about had international applicability. It should be noted that this volume set is not designed to provide a full international comparative analysis of the issues. Such a task would be truly daunting given the large number of legal systems that exist throughout the world. We believed that above all else, the main focus of the chapters should be on discussing the psychological principles relevant to the topic being written on. However, we felt that by asking authors to contemplate the international relevancy of their topics we could advance the application of existing research to legal systems outside the United States. In addition, we are hopeful that some international consideration of procedures used by legal systems in other countries might stimulate research in the United States, leading to the identification and application of new methods for increasing the overall fairness of the trial process.

Volume Integration

As previously noted, we felt that the topics covered in this two-volume set naturally fit into the areas covered by each volume, and that the two-volume approach allowed us (and the contributing authors) to provide greater coverage and depth for the topics than would have been feasible in a single volume. However, it is important to avoid

viewing the two volumes as separate works, or to think of the chapters within the volumes as entirely distinct. Rather, the material reviewed in various chapters is inherently inter-related. For example, to understand the impact of a specific type of evidence (e.g., sexual harassment, syndrome, insanity) one must first consider the admissibility of such information, and the likelihood that such information will be presented in court. One should then consider the type of trial in which testimony will be used (e.g., civil versus criminal, or perhaps even one where capital punishment may be considered), and special challenges that jurors may face within that type of trial. Further, it is important to understand how that testimony will impact the jurors, and how jurors will combine that testimony with information they are exposed to outside the courtroom, with previous relevant experiences they may have had, or with factors related to any important individual differences between jurors. One must then consider the instructions that a judge will provide jurors with regard to how such testimony is to be used, as well as the general instructions provided in that type of trial. Finally, overall theories of jury decision-making should be applied when attempting to determine how jurors will combine all of the information.

Thus, a true understanding of the issues involved in these volumes necessitates that readers approach the chapters in each book with a motivation to integrate the various topics. Doing so is critical for advancing one's understanding of the topics, as well as advancing future research in the field, and ultimately developing procedures that can be implemented to improve legal decision-making.

This page has been left blank intentionally

Preface to Volume II—Psychological Expertise in Court

Daniel A. Krauss and Joel D. Lieberman

Psychological Expertise in Court: Psychology in the Courtroom focuses on legal areas where psychologists or other mental health professionals have been called upon to offer expert testimony to the courts. In this volume, a number of global trial contexts are explored in chapters on interrogations and false confessions, psychological syndromes, and child sexual abuse. In addition, this volume is primarily concerned with psychological evidence presented during the trial process—although some of the chapters have important discussions of how certain types of evidence are and should be collected (e.g., eyewitness identification, confessions, and reports of child sexual abuse). In contrast, psychological issues that occur before or after the presentation of evidence such as pretrial publicity, jury selection, and jury instruction are covered in *Volume I—Jury Psychology: Social Aspects of Trial Processes*.

While some of the topics in this volume have a long and controversial history (e.g., insanity evaluations), others (e.g., sexual harassment) have only recently seen psychological expert testimony allowed in court proceedings. Similarly, although psychological expert testimony was historically primarily the province of clinical psychologists, more recent psychological research outlining the problems of police investigative techniques and the international attention afforded to defendants who were wrongly convicted has led to more social and cognitive psychologists entering the courtroom as expert witnesses. The knowledge these psychological experts provide and the effects they have on all aspects of legal proceedings from evidence collection to jury decisions is the main focus of this volume. As with the first volume in this two-volume set, we asked authors to pay particular attention to how courts in a variety of different countries handled different types of expert testimony offered by psychologists.

Before expert psychological evidence can influence trial decisions, however, it must be admitted into court. As a result, this volume begins with a discussion of the admissibility of expert testimony. Daniel Krauss, Desiree Cassar, and Allison Strother provide a historical overview of the development of admissibility standards in the United States, beginning with the *Frye* test,

through the development and revision of the *Daubert* standard. Krauss et al. consider whether judges in the United States have the necessary scientific background to correctly uphold the admissibility threshold created by *Daubert*. In addition, the authors compare the standards used by courts in the United States with those used in a variety of other nations. This opening chapter provides the reader with a framework for understanding how judges decide what types of expert testimony should be admitted to (and excluded from) the courtroom. Specific types of expert testimony are then discussed in Chapters 2 through 7.

In Chapter 2, Richard Leo, Mark Costanzo, and Netta Shaked-Schroer discuss the psychological issues and cultural aspects that contribute to suspects falsely confessing to crimes. Interrogation procedures in a variety of countries are examined. The authors review techniques used by law enforcement agencies that may put undue pressure on suspects. Unfortunately, jurors are often not exposed to the interrogation process, but rather only to the final confession produced by that process. Leo et al. discuss the powerful influence of confession evidence on jurors, as well as procedural and policy solutions that may be implemented to either reduce false confessions, or to make jurors more aware of the extent of pressure placed on suspects during the interrogation process.

Chapter 3 focuses on the issue of eyewitness identification and testimony. Solomon Fulero identifies factors that contribute to eyewitness errors from the perspective of system variables (factors that are controllable by the criminal justice system) and estimator variables (factors beyond the control of the criminal justice system). The frequency that expert witnesses may testify regarding the accuracy of eyewitness testimony is discussed, as well as the impact that eyewitness testimony has on jurors. Fulero concludes that in the absence of expert testimony, jurors are not likely to be fully aware of the problems associated with relying on eyewitness memory.

Chapter 4 examines an area where psychology has been relevant to the courtroom for perhaps the longest period of time—the use of the insanity defense. Patricia Zapf, Tina Zottoli, and Gianni Pirelli discuss the issues of both criminal responsibility and competency to stand trial. The historical developments of insanity standards are reviewed, and the frequency with which the defense is used is discussed. The authors include a review of research investigating the characteristics of the defendant and the crimes that are more likely to be associated with a successful insanity-based defense. Zapf et al. also review the standards for Competency to Stand Trial determinations, how such evaluations are made, and recommendations for policy and procedural solutions in this area.

In certain cases, insanity determinations may be based upon determinations that a defendant was suffering from a particular "syndrome." Alexis Kennedy reviews a variety of types of syndrome evidence in Chapter 5. Among the primary types of syndrome evidence that are reviewed are rape trauma syndrome, post-traumatic stress disorder, and parental alienation syndrome.

Chapter 6 focuses on children's testimony in the courtroom. Susan Hall notes that during the 1980s, courts in the United States relaxed the rules governing children's testimony at trials. As a result, the frequency that such testimony was presented in court dramatically increased. These changes led to a marked increase in the amount of research conducted regarding the reliability of children's testimony. Hall reviews that research and identifies factors that may increase the suggestibility of child witnesses.

In the final chapter, Chapter 7, Sarah Greathouse, Lora Levett, and Margaret Bull Kovera cover the topic of sexual harassment. Sexual harassment law is reviewed followed by a discussion of social science research on the topic, including an examination of antecedents and consequences of such harassment. Much of the chapter is devoted to a discussion of juror decision-making in sexual harassment cases, and in particular jurors' ability to apply sexual harassment standards when considering the evidence they are presented with.

It was our intention to provide both breadth in the areas in which expert psychological testimony may influence court decisions and depth within these various legal contexts, and we believe that this book has accomplished this goal. We hope that readers are especially stimulated by this book's focus on how a variety of different courts from around the world approach these difficult and often controversial issues as well as intrigued by the concrete solutions offered by our authors on how legal practices can be better informed by existing psychological research. While there is still a long way to go before courts worldwide are using expert psychological testimony more effectively, we hope this book's content takes a first step in the right direction.

This page has been left blank intentionally

Psychology in the Courtroom

This two-volume set explores the major areas in which psychology has been applied to court proceedings. A renowned group of psychological and legal scholars explore relevant problems that are created by or influence courtroom procedure and trial outcome from a psychological perspective. The authors discuss how courts within and outside the United States endeavor to handle these problems, and present empirically based potential policy solutions for these issues.

Readers may also be interested in the accompanying volume: *Jury Psychology: Social Aspects of Trial Processes*, edited by Joel D. Lieberman and Daniel A. Krauss. In this volume, a wide range of topics are covered including, pretrial publicity and inadmissible evidence, jury selection, jury instruction, death penalty cases, as well as decision-making in civil trials. In addition, a number of global issues are discussed, including procedural justice issues and theoretical models of juror decision-making. A consideration of methodological issues relevant to the study of juror behavior is provided. All of these topics are discussed from a psychological perspective. Throughout the volume the authors make a variety of recommendations for improving trial procedures where jurors are involved. The authors also discuss how the problems and potential solutions are relevant to courts around the world.

To Trina—DAK

To Celia—JDL

To Bruce Sales, a mentor and a friend—DAK and JDL

Chapter 1
The Admissibility of Expert Testimony in the United States, the Commonwealth, and Elsewhere

Daniel A. Krauss, Desiree Cassar and Allison Strother

Synopsis

As experts have become more essential to the determination of important legal decisions, courts have attempted both to allow important evidence to be available to legal decision-makers and to limit the "junk science" which enters the courtroom. This chapter explores the differences that exist in divergent legal systems' attempts to grapple with the admissibility of expert testimony. It offers a critique of existing legal standards in the United States, Canada, Australia, England, and the Netherlands. In the end, it argues that while legal standards that control these issues differ considerably from country to country, factors outside the standards themselves often determine whether expert testimony in a particular case is admitted or rejected.

The Admissibility of Expert Testimony in the United States, the Commonwealth, and Elsewhere

As the twentieth century drew to a close, both advances in technology (e.g., DNA testing) and the complexity of litigation issues (e.g., whether certain work practices caused emotional distress in an employee) led to more frequent use of and greater importance being placed on expert witnesses in the legal system. For example, two surveys of United States federal judges performed by Krafka et al. (2002) indicated that experts were present in approximately 95 percent of surveyed judges' civil cases, and that these cases averaged over four experts per trial. There was also growing concern that too much "junk science" was entering the courts, and that judges and jurors were being unfairly influenced by the cloak of science that surrounded unfounded expert testimony. For example,

in the United States, the 1991 publication of Peter Huber's *Galileo's Revenge: Junk Science in the Courtroom* was seen as a watershed moment in this growing debate. Huber's book offers a scathing attack on junk science-based expert testimony and discusses its damaging consequences, citing examples from a wide variety of less-than-scientific fields, including astrology, alchemy, and homeopathy. The book also documents expert testimony based on these pseudo-sciences serving as the pivotal link to liability judgments in a substantial number of civil cases.

Around the same time as the publication of Huber's book, the proliferation and success of large scale mass tort lawsuits in the U.S. (e.g., asbestos and tobacco litigation) was garnering substantial media attention. In many mass tort cases, the causal link between the putative harm visited on a large group of plaintiffs by a corporation's or an entity's business practices is based almost exclusively on expert testimony. Without an expert to bridge the gap between the defendant's conduct and the plaintiff's injury, these cases have little chance of success. Bernstein (1996) suggests Huber's critique as well as the media storm created by a number of astronomical monetary judgments in these cases raised the importance and profile of expert testimony admissibility decisions both within and outside the legal system.

The Role of Expert Testimony in the Legal System

Before discussing the substance of expert admissibility decisions by courts, it is first necessary to understand the unique role experts and their testimony play in the legal system. Experts are able to offer interpretations of events that they themselves have not witnessed. In contrast, ordinary witnesses are usually restricted to events they have directly experienced. Further, experts, based on their knowledge, experience, or research, are also allowed and called upon to offer opinions to courts about myriad topics while ordinary witnesses are restricted to factual statements. As a result, the admittance of expert testimony is often crucial to success in both civil and criminal trials. Yet, the very attribute that makes expert testimony most useful to the legal system, namely that it brings evidence and information to bear that is beyond the knowledge of the average jury or judge, also has made it increasingly controversial. It remains unclear whether sound, relevant expert testimony is aiding legal decision-makers in making better decisions or if overly persuasive but inaccurate expert testimony is confusing and misleading judges and jurors. In other words, the question remains whether the increasing use of experts is leading to more just outcomes or more injustice.

An Overview of Issues Surrounding the Admissibility of Expert Testimony

Courts world-wide are divided as to the appropriate standard for evaluating the admissibility of such testimony. Some courts have chosen to restrict greatly the amount of expert testimony that reaches the ears of the jury or fact-finders while

others impose fewer restrictions on such testimony. Those courts opting for less stringency in their admissibility decisions assume that the fact-finders will be able to distinguish successfully between high and low quality expert testimony. The structure of legal systems themselves, however, may affect whether judges and jurors can accomplish this task. Adversarial systems, like that of the United States, rely on cross-examination and competing experts to educate the judge and jury about the weaknesses of expert testimony. In inquisitorial systems, the expert is often appointed from a pre-approved panel of the court and often little attempt is made to point out weaknesses of the expert testimony to the ultimate fact-finder, the judge.

Further, courts in different countries vary on how uniformly they apply their evidentiary admissibility standards. Some courts apply the same standard to all expert testimony regardless of whether the testimony is based upon particle physics or handwriting analysis. Other courts have applied different standards to evidence they believe is novel scientifically (e.g., DNA evidence) as opposed to expert testimony that has traditionally been admitted (e.g., foot print analysis) or is non-scientific in nature (e.g., the cost of medical expenses for injury). Yet, even in jurisdictions that apply the same uniform standard to all forms of expert testimony, substantial differences in decisions on the same evidence occur because these standards are applied divergently by human actors (i.e., judges).

In addition, certain forms of expert testimony, including forensic science evidence and behavioral and social science expert testimony, have engendered greater legal controversy than others. To date though, psychological expert testimony has played a relatively minor role in evidentiary admissibility jurisprudence. The heart of this chapter will outline the legal foundations that underlie evidentiary admissibility standards so that their application to psychological expert testimony can be understood in a broader context. This chapter will not provide a convenient reference guide as to what forms of psychological expert testimony will be admitted in a particular jurisdiction because it is simply not possible to easily predict how a judge will apply a particular standard. Rather, this chapter is intended to provide a framework so that the tradeoffs and controversies surrounding different types of expert testimony admissibility decisions can be better understood.

Recent Developments in Evidentiary Admissibility Standards

In the U.S., the significance of and controversy surrounding expert testimony in large-scale tort court cases paved the way for changes in expert testimony admissibility standards. Though this sort of litigation does not exist outside the U.S., Bernstein (1996) argues that evidentiary admissibility standards abroad faced a similar re-examination due to other legal developments. At the crux of the evidentiary controversy abroad were disputes involving forensic science expert testimony (e.g., DNA evidence, blood, hair and fiber analysis) and behavioral and social science expert testimony (e.g., Battered Women Syndrome, Rape Trauma Syndrome). For

example, in an unusual Australian case involving the death of an infant, considerable controversy surrounded the admissibility of forensic evidence that purported to determine whether a child was killed by his mother rather than eaten by a dingo. This highly-publicized case, *R. v. Chamberlain* (1983), raised concern related to the admissibility of unscientific expert evidence at trial.

In the 1990s, the U.S. was the first country to enact substantial changes in its expert evidentiary admissibility practices. According to both Bernstein (1996) and Gatowski et al. (1996) the U.S.'s actions served as a significant catalyst for a rethinking of these same practices internationally. This chapter will examine: a) the factors that led U.S. courts to reevaluate their expert testimony admissibility procedures; b) the strengths and weaknesses of various evidentiary admissibility tests used in the U.S.; c) how these expert testimony standards have specifically affected the admissibility of psychological expert testimony; and d) how courts in Canada, the United Kingdom, Australia, and the Netherlands are currently approaching these issues.

A Brief History of Expert Testimony Admissibility Standards in the United States

The Frye Test

For 70 years in the U.S., the admissibility of novel scientific evidence and expert testimony was largely controlled by the standard announced in *Frye v. United States*. This 1923 District of Columbia Court of Appeals decision barred the admissibility of an early form of polygraph evidence. The part of this decision that became known as the *Frye* test required the judge to determine the admissibility of novel scientific information based upon whether it "… was sufficiently established to have gained *general acceptance* in the field in which it belongs" (*Frye*, 1923, p. 1014; emphasis added). This test is notable in a number of regards. First, while it required the judge to make the eventual admissibility decision, it deferred to the scientific community as to whether the evidence proffered was generally accepted. On the one hand, this delegation of authority kept judges from acting as amateur scientists. On the other hand, however, it prohibited scientifically sound, groundbreaking research from reaching jurors because it had not yet become generally accepted by the scientific establishment. Likewise, mainstream but inaccurate scientific evidence could be admitted by judges under this standard because the discipline had failed to recognize the errors associated with generally accepted thinking on a particular issue. For example, under this standard, courts in the mid-nineteenth century might have admitted expert testimony suggesting that the shape of an individual's head controlled his/her dispositions (i.e., phrenology) because this was a widely held or generally accepted belief among scholars at the time.

Second, the range of expert testimony that the *Frye* standard applied to was initially unclear. Because the *Frye* test was adopted in a decision involving novel scientific evidence and expert testimony based on that evidence, it did not provide guidance as to whether this test applied to all forms of expert testimony or only those that involved new forms of scientific evidence. Subsequently, although some jurisdictions applied *Frye* to all expert testimony, the more common rule among U.S. jurisdictions was that it only applied to novel forms of scientific evidence. As Bernstein and Jackson (2004) argue, most U.S. courts adjudicated expert testimony which was not novel or scientific using other standards.

Third, the *Frye* test was criticized for its vagueness and likelihood of leading to inconsistent decision-making within and between jurisdictions. The lack of clarity in terms such as "the field to which it belongs" led to non-uniform application of this test by U.S. courts. For example, the interpretation of the general acceptance of polygraph evidence could be decided differently by two different courts depending on whether the court determines that the field it belongs to is polygraphy or science. Polygraphers would easily view such evidence as generally accepted in their field, while scientists would be more skeptical of the general acceptance of successful lie detection in their discipline.

The General Relevancy Approach

For the next 50 years the *Frye* test received much scholarly criticism, but as Sanders et al. (2002) point out, it was not until shortly after the adoption of Federal Rules of Evidence (FRE) in 1975 that controversy fully returned to this area of law. According to Bernstein and Jackson (2004) and Sales and Shuman (2007), the Federal Rules of Evidence were an attempt to codify a set of uniform evidentiary rules for the federal courts, but also were adopted almost in their entirety by every state court in the United States. Groscup et al. (2002) and Krauss and Sales (1999), among others, have argued that generally, the FREs have a "liberal thrust" towards the admissibility of evidence, allowing more evidence to reach the jury than common law rules would have allowed. The FREs also contain a number of rules that specifically govern the admissibility of expert testimony. Rule 403 states that any relevant evidence "… may be excluded if its probative value [is] substantially outweighed by danger of unfair prejudice, confusion of issues, or misleading the jury …", and Rule 702 specifies that expert testimony is admissible if an appropriately qualified expert offers "… specialized knowledge that will assist the trier of fact [the judge or jury] to understand the evidence or to determine a fact in issue." Since the FREs do not specifically mention the *Frye* test, it was unclear whether these new rules of evidence replaced the *Frye* general acceptance test (Sales and Shuman, 2005; Shuman and Sales, 1999).

Some U.S. courts held that FREs did replace the *Frye* test, and began applying a more relevancy-based approach to the admissibility of expert testimony (Bernstein 1996; Sanders et al. 2002). Under this standard, expert testimony or scientific

evidence did not have to be generally accepted by the scientific community; it simply had to be relevant to the question being asked by the court. This approach, in keeping with the liberal admissibility elements of the FREs, adjudicated whether novel or other forms of expert testimony were not prejudicial according to Rule 403 and whether they could "assist the trier of fact" according to Rule 702. In order to assist the trier of fact, the court generally concluded that the proffered testimony had to be beyond the normal understanding of laypersons. Further, courts adopting this standard focused more on evaluating the credentials of an expert to determine if he/she was qualified to offer an expert opinion than they did on the scientific quality or general acceptance of the testimony being offered. Courts adopting this rule assumed that: a) jurors were capable of understanding and weighing even complex scientific information; b) jurors would not be unduly persuaded by a dubious expert in a particular case; and c) existing adversary procedures (e.g., cross-examination and competing experts) could ameliorate any bias caused by such an expert testimony. Unfortunately, the weight of empirical research that exists on this subject clearly suggests that many of these assumptions are incorrect.

Psychological research demonstrates that judges and jurors are often unaware of the limitations of expert testimony, and tend to rely on heuristic cues or cognitive short cuts rather than the content and quality of the testimony in evaluating and using expert testimony in their decision-making. An examination of the psychological decision-making theories relevant to this issue and the research based upon them, is beyond the scope of this chapter but is covered thoroughly elsewhere (see Chapter 3, in Volume I). In the end, the relevancy approach outlined by the FREs is best viewed as a generally less stringent admissibility standard than the *Frye* test.

Daubert and a Reliability Framework

Not all U.S. courts adopted a general relevancy approach following the promulgation of the FREs. Bernstein and Jackson (2004) suggest that some courts continued to apply the *Frye* test to novel scientific evidence while applying a general relevancy evaluation to other forms of expert testimony. As Bernstein (1996) and Sanders et al. (2002) note, still other courts began evaluating the reliability of the evidence that the expert offered. This latter standard would be the one adopted by the United States Supreme Court in 1993, in *Daubert v. Merrell Dow Pharmaceuticals*, when the Court attempted to clarify the confusion that surrounded expert testimony admissibility.

In *Daubert*, a group of individuals sued Merrell Dow Pharmaceuticals because they believed that the nausea drug that their mothers had been administered during pregnancy had caused their birth defects. At issue in the case was the admissibility of certain expert testimony, which was based on a re-evaluation of existing data using new methodology which provided proof of a link between the plaintiffs' harm and the Merrell Dow's drug. Originally, this expert testimony was not admitted by the trial court because the court reasoned the expert testimony offered did not meet the requirements of the *Frye* general acceptance test. Basing its decision on FRE 702,

the Supreme Court reversed, holding that the lower courts had misapplied the rules governing the admissibility of expert testimony. They reasoned that the *Frye* test had been specifically superseded by the adoption of the FREs. Further, the decision offered a new evidentiary reliability framework for courts to use in evaluating the admissibility of expert testimony. The majority opinion did not, however, limit their interpretation of this standard solely to the wording of FRE 702, but rather the Court held that the scientific validity of the *content* of the expert testimony was the main issue for the court to use in determining admissibility (O'Connor and Krauss, 2001). Subsequent to the *Daubert* case, FRE 702 was rewritten to reflect more accurately the decision. The current wording of FRE 702 is "If the scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, *if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case*" (note – italicized text differs from the original Rule 702).

Additionally, the Court noted that the trial judge was responsible for making these determinations, and offered a non-exhaustive flexible list of factors that judges could use in "… determining whether a theory or technique is scientific knowledge that will assist the trier of fact:"

1. Has the theory or technique been tested or is subject to being tested?
2. Has the theory or technique been subjected to peer review and publication?
3. What is the known or potential rate of error in applying the particular scientific theory or technique?
4. To what extent has the theory or technique received general acceptance in the relevant scientific community?

This judicial inquiry was not intended to occur in a vacuum, but rather required that the proffered expert testimony be evidentiary reliable for the specific question being asked by the court.

This "fit" requirement explicitly acknowledges that scientific validity for one purpose is not always scientific validity for another, and consequently, expert testimony that would be admitted in one case may be deemed inadmissible in another. For example, in a 2005 case, *U.S. v. Dixon*, the U.S. 5th Circuit Court recently upheld a trial court's exclusion of expert testimony regarding Battered Women's Syndrome because of the type of defense that was offered. The expert testimony in question focused on the defendant's "subjective perceptions of danger," rendering the testimony inadmissible since the duress defense used in the case required an "objective inquiry" into the defendant's conduct. The expert's abilities, data, and methodology were not in question; rather, the evidence offered was simply not valid in this particular case due to its subjective nature. Had the defendant offered another

defense, or had the case been different (e.g., a civil matter rather than a criminal trial of illegal firearms possession), this expert testimony may have been admitted.

Clarification of the Daubert decision In two subsequent cases over the next six years, *General Electric v. Joiner*, 118 S.Ct. 512 (1997), and *Kumho Tire v. Carmichael*, 526 U.S. 137 (1999) (together with *Daubert*, these cases are often referred to as the *Daubert* trilogy), the United States Supreme Court further clarified the standard adopted in *Daubert*. First, the Court stated in *Kumho* that *Daubert's* evidentiary admissibility framework applied equally to all forms of expert testimony. Some courts had been making distinctions between scientific and technical expert testimony because *Daubert* specifically involved the admissibility of scientific expert testimony, and FRE Rule 702 refers to "scientific, technical, or other specialized knowledge." The *Kumho* decision clearly indicated that these distinctions were not to be made.

Secondly, the Court held that the factors listed in the original *Daubert* decision did not have to be used to make this determination. It held, however, that some attempt needed to be made by the trial judge to assess the evidentiary reliability of expert testimony proffered. As Justice Scalia explicitly notes in his concurrence in *Kumho*, judicial discretion in this matter, "… is not discretion to perform the function inadequately. Rather, it is discretion to choose among reasonable means of excluding expertise that is fausse and science that is junky." (*Kumho Tire*, 1999. pp. 158–9). Additionally, the Court determined that there would be limited appellate review of trial courts' decisions on these matters. In *Joiner*, the Court held that the expert evidentiary admissibility decision of a trial court judge could be overturned on appeal only if the judge abused his or her discretion in making the decision or the decision was "manifestly erroneous." This extremely deferential standard of review suggests that evidentiary admissibility decisions are unlikely to be useful grounds for appeal (Krauss and Sales, 1999).

It is important to recognize that the *Daubert* trilogy of cases only controls the evidentiary admissibility standards of federal courts. Although many state courts have adopted the *Daubert* reasoning, they have not adopted all of the trilogy's elements. Further, many of the most populated states in the U.S., including California, New York, and Florida, have explicitly chosen not to adopt *Daubert* and have retained the *Frye* test.

It is also important to understand that while the *Daubert* standard is commonly referred to as a reliability standard by legal commentators, it is actually meant to be an examination of the scientific validity of expert testimony. The *Daubert* court intended to use the term "reliability" in the legal sense and not in the scientific one, as they highlight in a footnote of the opinion "in a case involving scientific evidence, evidentiary reliability will be based upon scientific validity" (*Daubert*, 1993, p. 591). The Court's semantic imprecision has, however, generated some problems for lawyers' and scientists' communication on the differences between scientific reliability and validity and evidentiary reliability.

Strengths and weaknesses of the Daubert trilogy Perhaps the greatest strength of the *Daubert* trilogy of cases is that it requires admissibility decisions to be based upon the scientific validity of the expert testimony being offered and its fit to the pertinent legal question. This, however, may also be their greatest weakness. Judges simply may not be equipped to evaluate the scientific validity of expert testimony. No lesser authority than the Chief Justice noted this problem in his dissent in *Daubert*, acknowledging that "I defer to no one in my confidence in federal judges; but I am at a loss to know what is meant when it is said that the scientific status of a theory depends on its 'falsifiability,' and I suspect some of them will be, too" (*Daubert*, 1993, p. 600). Gatowski et al.'s (2001) empirical survey of 400 state court judges supports Chief Justice Rehnquist's claim and suggests a grave difficulty for the U.S. legal system. The survey found that only 5 percent of judges could correctly define falsifiability from a scientific perspective and only 4 percent possessed an adequate understanding of error rate (Gatowski et al., 2001). Survey results also indicated, however, that judicial understanding of the two other *Daubert* factors was higher, with 71 percent of judges showing clear understanding of peer review and 82 percent demonstrating an understanding of general acceptance. While the ability of judges to define two of the so-called *Daubert* criteria adequately outside the context of an actual trial and its actors may not be the fairest test of their ability to make admissibility decisions, it does at least suggest the possibility of a significant problem.

Yet even if judges thoroughly understood the meaning of the four *Daubert* criteria and could apply them adequately to expert testimony, other aspects of the *Daubert* trilogy decisions also pose difficulties. Most significantly, the lack of appellate review of trial courts' *Daubert* decisions necessarily means that trial courts will not receive guidance on: a) how to weigh and balance the explicitly mentioned four *Daubert* criteria in their decisions; b) how to choose additional factors to evaluate scientific validity; nor c) how to balance and weigh any newly selected factors with the appropriate *Daubert* factors in reaching a decision. As a result, the courts are likely to evidence non-uniform decision-making between and within jurisdictions with regard to admissibility. Furthermore, plaintiffs, particularly in high-cost tort cases, are likely to forum shop for favorable judges (Krauss and Sales, 1999). Forum shopping refers to litigants bringing cases in jurisdictions that have the most favorable rules and previous decisions applicable to their current position. In many civil cases, the plaintiffs will have considerable discretion in which jurisdiction to bring a case against a corporation or an entity that conducts business in multiple locations. So, while the *Daubert* trilogy of cases affords judges flexibility in choosing the appropriate factors to use in evaluating expert testimony and in the manner in which they balance them, it does so at the expense of uniformity and consistency in decision-making.

Difficulties in applying Daubert to forensic, behavioral, and social science testimony These difficulties are clearly present when the *Daubert* standard is applied to many forms of forensic science expert testimony such as blood analysis,

hair and fiber analysis, bite mark analysis, and fingerprints (e.g., see, Saks and Koehler 2005; Thompson and Cole, 2007). The link between such evidence and its identification of the perpetrator of a crime rests on a number of assumptions that to date have received limited empirical validation. For example, Saks and Kohler (2005) argue there is little that is scientific about forensic scientists' attempts to compare pairs of marks or fluids to each other (i.e., a sample from the alleged perpetrator and something found at the crime scene). They note that forensic scientists rarely possess population norms for the items that they are comparing or that "… different objects share a common set of observable attributes" (Saks and Kohler, 2005 p. 892). DNA evidence, in contrast, is recognized as a useful model and important exception to these problems. In spite of these issues, problematic forms of forensic science evidence are rarely prohibited even in *Daubert* jurisdictions because of both the long tradition of admitting this type of evidence and the importance to the prosecution of admitting such expert testimony.

The problems of *Daubert* are compounded when this standard is applied to controversial behavioral and social science expert testimony. Sales and Shuman (2007) note that much of the expert testimony in this arena (e.g., sanity opinions, future dangerousness pronouncements, best interest of child in a divorce proceeding, or syndrome testimony) does not lend itself to easy analysis through the four *Daubert* criteria. For example, how would you determine the error rate associated with expert clinical predictions (based on an expert's experience rather than on a standardized risk assessment tool) of a criminal defendant suffering from Battered Women Syndrome? How falsifiable is the clinical prediction? Has peer review been completed on the accuracy of clinical judgments of Battered Women Syndrome? Even if such empirical data could be created, it presently is unlikely to exist. These questions pose severe problems for researchers well-versed in clinical diagnosis and specializing in Battered Women Syndrome, and they present even greater problems for judges who have little training in the complexity of diagnosis, prediction, and treatment. The *Daubert* inquiry also does not necessarily answer the question of whether a clinical prediction or diagnosis is accurate in the particular case before the court.

Similar to forensic science evidence, Slobogin (2007) contends that certain forms of behavioral and social science expert testimony are commonly admitted by courts even though their scientific validity is suspect while other forms with stronger empirical basis are often rejected. For example, expert testimony on a defendant's future dangerousness or on a defendant's mental state at the time of the offense are routinely admitted by judges in the absence of any empirical research suggesting their accuracy. This is true of most mental health expert testimony that is traditionally used by the prosecution in criminal trials. In contrast, Slobogin (2007) argues well-developed social science expert testimony based upon empirical research in certain areas (e.g., the fallibility of eyewitness identification, the likelihood of false confessions) has faced greater difficulty in being admitted both because it lacks a tradition of acceptance and because it favors the defense. It is clear that some

changes in the admittance of the latter expert testimony have occurred as a result of the *Daubert* decision, but the former continues to be admitted regardless of its lack of scientific validity. Slobogin concludes admittance of behavioral and social science expert testimony, even in *Daubert* courts, has less to do with its scientific validity than it has to do with a tradition of acceptance or exclusion by the judges who control the legal system.

In the end, the reliability framework announced by the Court in the *Daubert* trilogy cases creates a number of substantial problems for judges, including: a) judges' limited ability to understand the concepts and terms used to evaluate scientific validity in the *Daubert* decision; b) limited guidance for how judges should apply and weigh the *Daubert* factors as well as other relevant factors in evaluating particular pieces of expert testimony; and c) the limited usefulness of the *Daubert* factors to many forms of evidence, such as behavioral and social science evidence and forensic science expert testimony.

The above discussion is not meant to imply that judges are incapable of applying the *Daubert* reliability test in a meaningful and effective manner. In practice, for almost every anecdotal report of a judge improperly applying aspects of the *Daubert* decision, there are complementary examples of judges performing sensible and appropriate evaluations (see Sales and Shuman, 2005 for further discussion of this issue). The discussion above is intended to highlight that the *Daubert* reliability framework is especially difficult to apply in a consistent and uniform manner for many forms of evidence.

Frye, Relevancy, and Daubert Reliability Compared

Much of the U.S. scholarship following the *Daubert* decision attempted to determine if *Daubert* was a more or less stringent standard than the *Frye* test or general relevancy approaches. This, unfortunately, is a somewhat ill-conceived question. The three standards ask and are intended to answer distinctly different questions. General relevancy focuses on the credentials of the expert offering the expert testimony and his or her ability to assist the trier of fact to determine a pertinent legal question. In contrast, *Frye* asks whether the expert testimony is mainstream thinking in the expert's field. Finally, *Daubert* examines the validity of the content of the expert's testimony and its fit to the legal question. These standards are probably best viewed as existing on a continuum in which general relevancy is likely to allow a somewhat broader range of expert testimony than *Frye*. In turn, *Frye* can lead to more expert testimony being admitted than *Daubert*. This continuum notion has been somewhat borne out in practice. Many post-*Daubert* reviews, including Bernstein's (1996), found less scientifically-based behavioral and social sciences evidence (e.g., Battered Women Syndrome) more likely to be deemed inadmissible than had been the case under previous standards. More generally, a host of empirical investigations, including those by Dixon and Gill (2002) and Krafka et al. (2002), have found that following *Daubert*, less expert testimony appears to be admitted than under *Frye* or

general relevancy standards. A study by Groscup et al. (2002), however, suggested little difference in appellate expert testimony admissibility decision pre- and post-*Daubert*.

These three standards, due to their subjective elements and lack of appellate review, do not necessarily have to be applied distinctly in practice. In fact, the different expert admissibility standards might be better envisioned as admitting evidence with substantially overlapping distributions. In other words, what evidence is admitted in a particular case may end up depending more on the manner in which the tests are applied rather than what standard is used. As a result, in specific cases and for specific types of expert testimony, it would not be unusual for seemingly counterintuitive decision-making to occur. For example, a very strictly applied general relevancy test may lead to certain expert testimony being prohibited in one case while a less stringently applied *Daubert* framework may allow the same expert testimony to be admitted. It is not difficult to imagine a case in which a judge in a general relevancy jurisdiction might legitimately decide that Battered Women Syndrome (see Chapter 5 in this volume) is not beyond the understanding of the lay juror. Alternatively, the judge may find that even if Battered Women Syndrome is beyond their understanding, such testimony will not assist the jury in understanding issues relevant to the trial. As a result of these findings, the court could prohibit its admissibility. In contrast, a different judge in a *Daubert* jurisdiction might admit such testimony because he/she determines: a) it is beyond the knowledge of the average juror; b) it is generally accepted in the scientific community; c) it has been empirically tested; and d) research on it has been published in peer review journals. What then are the factors that determine how strictly a judge applies a particular evidentiary admissibility test?

Judges' training, traditions, and values As mentioned previously, judges may have difficulty applying the *Daubert* test if they lack the scientific training and ability to understand its underlying concepts. In such situations, they could become educated in science. There are numerous resources currently available to U.S. judges to aid them in making *Daubert* decisions, including: The Reference Manual on Scientific Evidence, 2nd edition, published by the Federal Judicial Center. There are also a wide variety of continuing education programs available to judges to help them in understanding various scientific evidence topics and the principles underlying scientific validity. There is, however, no empirical evidence that judges use these resources, or if they do, that they improve the quality of their admissibility decisions.

Judges may also ask for outside help in understanding the science in question. Under FRE 706, judges have the option of appointing their own expert witness to help in their decision-making, and recently a number of expert panel services have been created in the U.S. to help judges in finding suitable experts. Unfortunately, Cecil and Willging's (1993) survey of judges indicates that judges rarely use their

power under FRE 706 to ask for or receive outside help on scientific evidence admissibility questions.

Finally, judges may rely on the *Daubert* factors that they do understand in making evidentiary admissibility decisions. There is some limited evidence to support the notion that judges have been doing the last of these possibilities. Gatowski et al.'s (2001) survey of 400 judges found that respondents were most likely to rely on the general acceptance component of the four *Daubert* factors in their decision-making. This general acceptance factor is essentially identical to the abandoned *Frye* test. Additionally, Dixon and Gill's (2002) post-*Daubert* review of court decisions also found support for the continued use of the general acceptance test among judges. This *Daubert* factor was especially predictive of decisions not to admit expert testimony, although a study by Groscup et al. (2002) suggests that judicial use of general acceptance as a factor in decision-making decreased following *Daubert*. Judicial reliance on general acceptance or *Frye* in *Daubert* jurisdictions could be the result of judges' ease in scientifically understanding and applying this factor. It could also be a result of their experience in using this standard if they previously evaluated expert testimony using *Frye*, or it could be some combination of the two.

It is not surprising that what judges have done in the past affects how they apply new standards, even if the new standards call for a somewhat different type of assessment. Thompson and Cole (2007) note that this reasoning may explain why judges have continued to admit problematic forensic science expert testimony even in *Daubert* jurisdictions. Likewise, some forms of social and behavioral expert evidence that clearly lack demonstrable scientific validity, like clinical pronouncement of what might be in a child's best interest in a divorce proceeding, have rarely been scrutinized even in *Daubert* jurisdictions (Krauss and Sales, 2000; Sales and Shuman, 2007).

Beyond ability, practical experience, and tradition, judges' evidentiary admissibility decisions may also be influenced by their values. This is particularly likely to be so for determinations made using the ambiguous *Daubert* standard, as Berger (2001) and Sales and Shuman (2007) have found. For example, if judges are former prosecutors, they are likely to value the admittance of forensic scientific evidence as central to achieving justice in criminal trials. As a consequence, they are more likely to apply *Daubert* less stringently to such forensic scientific testimony. In sum, while the three different approaches to the admissibility of expert testimony suggest different questions and different types of evaluations by judges, in practice knowledge of the evidentiary admissibility rule will not always easily determine whether expert evidence will be admitted or prohibited in a particular jurisdiction or case.

General Conclusions Regarding U.S.'s Evidentiary Admissibility Standards

The U.S. Supreme Court's decisions in the *Daubert* trilogy of cases set the stage for a possible narrowing of evidentiary admissibility determinations in jurisdictions

that subsequently adopted the Court's reasoning. The decisions required judges to evaluate the evidentiary reliability or the scientific validity and the fit of expert testimony to the legal question at issue in adjudicating admissibility. Yet, even in jurisdictions that fully adopted the Court's reasoning (a significant minority of states still follow the general acceptance test) there has been inconsistent admissibility decision-making with regard to identical forms of expert testimony. Due to several factors, this problem is particularly pronounced with regard to forensic identification and behavioral and social science expert testimony. Difficulties for the evaluation and application of this evidence include: a) ambiguity in the criteria to be used by judges; b) judicial difficulty in understanding and weighing the criteria; c) lack of appellate review of trial court decisions; and d) judicial values, traditions, and beliefs.

Evidentiary Admissibility Standards Internationally

In the next section, we turn to how legal systems in other countries have handled the problems inherent in determining the admissibility of expert testimony. We also highlight how their approaches are both similar and different from those used in the U.S. The evidentiary admissibility standards of Canada, Australia, the United Kingdom, and the Netherlands are described, and an attempt is made to place their approaches in the context of the general relevancy, *Frye's* general acceptance test, and *Daubert's* reliability approach.

Canada

Over the course of several years, the Supreme Court of Canada has addressed the issue of expert testimony admissibility in a number of cases but has yet to announce a standard for lower courts to follow. In *R. v. Beland* (1987), the Court adopted a general relevancy approach, rejecting a general acceptance test in its decision to deny polygraph evidence. The Supreme Court of Canada reiterated its adoption of a general relevancy test by admitting Battered Spouse Syndrome expert testimony as a self-defense claim for a battered woman accused of murder in *R. v. Lavallee* (1990). This ruling has influenced many lower courts to admit expert testimony specific to Battered Spouse Syndrome, but the *Beland* and *Lavallee* decisions did little to establish a standard for lower courts to use as guidance in expert evidence admissibility decisions. Some Canadian courts subsequently adopted a multilayered approach to evidentiary admissibility. In this approach, the courts consider: a) the relevancy of novel evidence; b) the potential rate of error associated with the evidence; c) the probative value of the evidence; d) whether the evidence is beyond the experience and knowledge of the trier of fact; and e) the qualifications of the expert.

In *R. v. Mohan* (1994), the Supreme Court of Canada rejected expert psychiatric testimony proffered on behalf of a pediatrician accused of sexually assaulting female

patients. The defendant called on a psychiatrist to testify that perpetrators of such offenses fit the profile of a sexual psychopath, and that the defendant did not fit this behavioral profile. In its decision, the Supreme Court of Canada established four expert testimony admissibility criteria: a) relevance (i.e., legal relevancy); b) necessity in assisting the trier of fact; c) absence of any legal rule prohibiting such evidence; and d) qualifications of the expert.

At first glance, these criteria suggest a liberal relevancy approach similar to the U.S. general relevancy standard. Yet, closer scrutiny of these criteria indicate that in practice the standard announced in *Mohan* more resembles a general acceptance test. The Court held that the trial judge should take into account whether the expert is using a "reliable indicator" of membership in a distinct group such as sexual psychopaths. Applying the criteria above, the Court did not find any evidence indicating "general acceptance" of a standardized profile of individuals who sexually assault women. The Canadian courts have subsequently held that these criteria only apply to psychiatric testimony regarding personality profiles. They do not apply more broadly to other types of expert testimony admissibility decisions, including most types of psychological or other relevant social science testimony.

As a result of these decisions, Canada does not appear to have a uniform judicial standard for the admissibility of expert evidence. Lower Canadian courts apply various admissibility tests and have broad discretion to make these decisions. Bernstein (1996) suggests that most courts apply a general relevancy test similar to FRE 702, and specifically do not apply a *Frye*-based general acceptance test. In contrast, a minority of Canadian courts have adopted a general acceptance test while an even smaller percentage of courts apply a more conservative reliability test.

Canadian courts are likely to show even greater non-uniformity in decision-making than U.S. courts because their courts lack an explicitly defined set of criteria for making these decisions. Consequently, attorneys and their respective clients may engage in forum shopping to facilitate their case. To avoid such problems, the Supreme Court of Canada may move towards adopting *Daubert* or for that matter any other uniform standard. Such an adoption would likely improve equity in decision-making, especially if explicit appellate guidance on how to apply these criteria were provided. Although adopting *Daubert*-based criteria may limit the quantity of admitted expert testimony, adequately guiding trial judges in their admissibility decisions may improve the quality of the expert testimony that is admitted.

Australia

In 1960, the High Court of Australia announced a standard in the *Clark v. Ryan* decision allowing expert testimony if the substance of the testimony is beyond the knowledge of the trier of fact (judge or jury), and if it is garnered from a structured field of study. Although the *Clark* decision suggests a general acceptance standard, the High Court of Australia did not specifically define what "a structured field of study" was, nor did it indicate how courts might apply this criterion.

In the decades following the *Clark v. Ryan* (1960) decision, Australian courts evidenced confusion regarding this standard, a trend noted by Bernstein (1996), Freckelton (1993), and Gatowski et al. (1996). This confusion was apparent following the *R. v. Gilmore* (1977) appellate decision. In this case, spectrographic voice analysis evidence was considered. Some legal commentators argued that this decision represented an adoption of a *Frye* general acceptance test; others disagreed. Subsequently, some courts applied a combination of the general acceptance and reliability tests. For example, the *Murphy v. Queen* (1989) court applied a general relevancy approach similar to FRE 403 and Canada's *Beland* decision in determining the admissibility of evidence. Similar to Canada, the *Murphy* court specified that the court must weigh both the probative and prejudicial value of evidence.

In the 1980s, a number of cases raised concern related to the admissibility of forensic science evidence because of the supposed misuse of such evidence by the government to secure convictions (e.g., *R. v. Chamberlain*, 1983). In *R. v. Chamberlain* (the Dingo Baby Case), for example, Mrs Chamberlain was convicted of killing her baby based on the forensic testimony presented at trial. James Cameron, a professor of forensic medicine, was the key witness for the government. He testified that he could distinguish the outline of a small human handprint in the bloodstains on the baby's clothing—seemingly powerful evidence of Mrs Chamberlain's guilt. As further evidence of human involvement in the killing, Cameron presented slides of the baby's clothing with a pattern that Cameron claimed was bloody fingerprints. Cameron also asserted that the bloodstain patterns indicated that the baby was killed by "a cutting instrument" used by a human. During cross-examination, Cameron was questioned about his testimony in another case that was later found to be completely erroneous. He admitted that he offered evidence against the defendants in the previous case without "correct knowledge of all the attendant circumstances." Despite the highly speculative nature of Cameron's testimony and widespread disagreement with it from other experts, the trial judge admitted the evidence and instructed jurors that even if they did not fully comprehend the complexities of the science, they could rely on "how convincing they found the experts." On appeal, two higher courts affirmed this decision, relying on an evaluation of the experts' credentials rather than focusing their attention on the problematic expert evidence.

Just prior to and following the release of Mrs Chamberlain in 1988, courts began favoring stricter evidentiary admissibility rules. Some courts turned to *Frye* while others adopted a stricter general acceptance rule than *Frye*. One such court considering scientific evidence involving bite marks applied a standard based upon a universal rather than a general acceptance of such evidence. Still other courts such as the South Australian Court of Appeals adopted a general relevancy approach and rejected *Frye* when considering scientific evidence. In general, however, as Bernstein (1996) and Gatowski et al. (1996) conclude, the majority of Australian courts appear to be applying an expert evidentiary admissibility standard that most closely resembles the general acceptance test of *Frye*.

Two years after the *Daubert* decision, the Australian parliament instituted an evidence code to guide federal courts in their evidentiary admissibility decisions. According to Section 79 of the code, expert testimony is admissible if a person "has specialized knowledge based on the person's training, study or experience," and if the testimony "is wholly or substantially based on that knowledge" (Bernstein, 1996, p. 161)." Bernstein (1996) argues this rule is vague, liberal and similar in intent to FRE 403. Although state courts are not obligated to adopt this evidence code, it is probable that some courts will adopt Section 79 of the code for their own evidence admissibility rules.

If Australian courts continue to apply a general acceptance-based test, novel scientific testimony may be excluded on the basis that it has not yet become generally accepted theory. This would limit the admissibility of cutting-edge research pertinent to deciding cases. On the other hand, if Australian courts move towards a reliability approach, judges may experience similar difficulties to that of American judges using *Daubert*. For example, like their U.S. counterparts, Australian judges may not be well equipped to evaluate the scientific validity of expert testimony being offered. Similar to Canada, the Australian High Court must also consider how best to guide lower court judges in their application of *Daubert* criteria. If a reliability test is to be effective in admitting expert testimony and excluding "junk science," judges must be directed how best to apply it to different types of evidence.

United Kingdom

The evidentiary admissibility standard in England is clearly a general relevancy approach. According to Bernstein (1996), what matters for English courts is whether the expert testimony is beyond the knowledge of jurors, particularly concerning psychiatric or psychological expert evidence. In *Preece v. H.M. Advocate* (1981), for example, judicial awareness concerning "junk science" was raised when the appellate court overturned Preece's murder conviction. The prosecution's case centered on testimony from a forensic scientist, Dr Clift, who used the results of tests on body fluids, fibers and hairs found at the crime scene to persuade the jury of Preece's guilt. The most important piece of forensic evidence presented was information about the blood stains found on the victim's clothing. The stains were determined to be of blood type A and from a secretor (a person who secretes certain blood components into other body fluids such as urine and saliva). Preece was both of blood type A and a secretor, a profile shared by less than one third of the population, so this fact seemed to be reasonably convincing evidence of Preece's culpability. In his testimony, however, Clift did not mention that the victim was also of blood type A and highly likely to be a secretor. Clift's failure to disclose this salient information demonstrates the importance of requiring a certain level of disclosure from expert witnesses, even if such disclosure may assist the other party in the case.

Although this decision (along with the Australian Dingo Baby Case discussed above) led to legal discourse favoring the adoption of a standard similar to *Frye*,

it did not establish a standard of evidentiary admissibility for English courts to follow.

As Gatowski et al. (1996) suggest English courts still approach the issue of evidentiary admissibility by focusing on relevancy. In *R. v. Robb* (1991) for example, the defendant appealed the trial court's decision to admit scientific evidence. The evidence in question was voice identification expert testimony, and the defendant appealed its admission because the technique was unscientific as determined by experts in the field. In spite of the experts' agreement regarding the unreliability of the technique, however, the court affirmed the admission of the testimony. English courts advocate that evidentiary reliability is a matter that goes to the weight jurors place on the evidence presented rather than whether it is admissible in the first place.

A number of cases have sparked interest in the adoption of stricter evidentiary admissibility rules for expert testimony, such as *R. v. Maguire*, 1992 and *R. v. McIlkenny*, 1991. One recommendation provided by the legal reform group Justice was to use caution when accepting results based on unpublished methodologies and lacking peer review. Another recommendation suggested that courts verify whether the expert disqualified alternative hypotheses before presenting the results. Regardless of these recommendations, however, English courts have yet to turn their attention toward stricter evidentiary admissibility rules. Bernstein (1996) notes that instead the English legal system has attempted to improve the forensic science system. To accomplish this, they have instituted reforms to improve the quality of forensic laboratories and the training of employees, and mandated external inspections of forensic laboratories.

Improvement of forensic science, procedures, and tests may increase the quality of that testimony; however, it does little to affect the quality of behavioral and social science evidence, such as Battered Women Syndrome and Rape Trauma Syndrome. Without adequate evidentiary admissibility standards, English courts may continue to admit questionable scientific expert testimony. This in turn may lead jurors to make inappropriate decisions. Further, English courts' continued use of a general relevancy standard assumes that: a) jurors are capable of understanding and weighing even complex scientific information; b) jurors will not be unduly persuaded by unscientific expert testimony; and c) cross-examination and opposing experts will remove bias caused by such expert testimony. As mentioned earlier, the limited empirical research that exists on this subject indicates that many of these assumptions appear faulty.

The Netherlands

There are no juries in court-centered jurisdictions such as the Netherlands. Judges in these jurisdictions have an active function in appointing experts, interrogating experts, and evaluating expert evidence. Although defendants can hire experts for their case, this is a rare practice. Judicial appointment eliminates the potential for a

battle of experts at trial because only one expert is appointed. It also allows judges to exercise a modicum of quality control in that they will not re-appoint an expert who they feel is unqualified. Yet, in a one-expert system, the influence of an unscrupulous expert is likely to be greater. Without the availability of another expert to question the validity and veracity of the expert's opinions or conclusions, the sole expert is undoubtedly likely to be extremely influential.

The Netherlands does not employ evidentiary rules with respect to the admissibility of expert testimony. Instead, this legal system applies decision and argumentation rules after the presentation of expert evidence. The Dutch Code of Criminal Procedure outlines only one statutory condition for expert testimony, which is, as quoted in van Kampen (2002), that the expert's testimony must be "his opinion, made in the course of the investigation at the trial, as to what his knowledge teaches him about that which was the subject thereof." Knowledge in this court-centered system refers to any knowledge possessed by the expert, scientific or otherwise.

Prior to the late 1980s, the Dutch Supreme Court held that the admissibility of an expert was largely a trial court decision. This is similar to the rule used in U.S. jurisdictions in which there is limited appellate review of trial court decisions. In 1989, however, the Dutch Supreme Court expanded its review of expert testimony in a sexual abuse case known as the *Anatomically Correct Dolls* case (HR 28 February 1989, NJ, 748; *Anatomically Correct Dolls*). In this case, an expert used anatomically correct dolls while interviewing alleged child sexual abuse victims, an investigative method that the defense argued was unsound. Despite the defense's objections, the appellate court accepted the expert's testimony as evidence. The Supreme Court later reversed the appellate court's decision. It found that the appellate court should have explained why it relied on the expert's testimony in spite of the defense's attack on the expert's method. As it was, the appellate court simply decided to accept the expert's testimony without justifying why it disregarded the defense's objections.

The High Court again reviewed a lower court's evidentiary admissibility decision in a manslaughter case referred to as the *Shoeprint* case (HR 27 January 1998, NJ, 404; *Shoeprint*). During the appellate trial, an orthopedic technician testified about similarities between shoeprints at the crime scene and those of the defendant. Because of these similarities, the technician concluded that the prints at the crime scene must have been made by the defendant. The defense, however, argued that this testimony should be excluded as evidence, contending that the orthopedic technician was not sufficiently experienced, and thus not qualified, to draw this conclusion. Without explaining its rationale for doing so, the appellate court rejected the defense's argument. The court simply decided that the technician's testimony was acceptable, provided that it was consistent with that of another expert, and consequently could be used as evidence.

The Supreme Court reversed the appellate court's *Shoeprint* decision for multiple reasons. Even though the technician had extensive training and experience in providing orthopedic shoes, the appellate court erred in relying on the technician's testimony without explaining its rejection of the defense's argument—the same

mistake made by the appellate court in the *Anatomically Correct Dolls* case. In addition, the Supreme Court found that the appellate court's decision "did not show that the appellate court had examined whether the expert's specialized knowledge also concerned research on, and analysis of, shoeprint, and if so, which method the expert used in investigating the prints, why the expert considered the method to be reliable, as well as to what extent the expert was capable of professionally utilizing that method" (HR 27 January 1998 NJ 1998, 404 m.nt. *JR.*, *Shoeprint*).

By requiring lower courts to ensure that experts' experience and knowledge qualify them to appropriately utilize a given method, the *Shoeprint* decision expanded the criteria for acceptable expert testimony. Whereas the *Anatomically Correct Dolls* case concerned only the reliability of the expert's method, the *Shoeprint* case involved a challenge to the particular expert's qualifications and abilities in utilizing a method. Methodological similarity between experts, a point raised by the trial court as grounds for admissibility, was not a measure of reliability according to the High Court. In this respect, the *Shoeprint* decision closely resembles FRE 702 (as amended in 2000). Both require not only expertise, but also an explanation of how and why a given method or principle is reliably applicable to the case at hand.

These cases indicate that in the Netherlands, both relevancy and reliability can play roles in the admissibility of expert testimony. Yet, it is important to understand, as van Kampen (2002) observes, that there are no uniform expert testimony admissibility standards in the Netherlands.

Although similar to FRE 702 or an even more liberal admissibility approach, the court-centered system places the burden on trial court judges to address reliability issues. Judges, however, must address these issues only when the defense raises concerns about particular expert testimony. Such a system requires greater judicial objectivity than adversarial systems because the defense can only attack the reliability of the expert testimony once the evidence has been admitted at trial. This *post-hoc* decision-making may cause the evidence to influence judges unduly before they exclude it, and may affect the outcome of the case. This problem is similar to one experienced in adversarial systems when there is a bench trial (i.e., no jury). As the judge in these situations is both the one who determines admissibility of evidence and the trier of fact, any evidentiary challenge will occur after the judge could have been unduly influenced by potentially inaccurate expert testimony.

Moreover, because experts in a court-centered system are generally appointed by the judge to assist the judge and are not perceived as witnesses as they are in adversarial systems, the defendant does not have the right to confront (i.e., cross-examine) the expert. This again grants the judge substantial discretion in the appointment of an expert, and the expert substantial latitude in what he/she might offer as expertise. Van Kampen (2002) argues that these differences do not necessarily lead to worse outcomes than those in adversarial systems where evidentiary admissibility standards are employed. They do, however, at least raise the possibility that the selection of an expert will be more important and that the influence of the expert is likely to be greater in a court-centered system than in an adversarial one.

Conclusions

Clearly, the importance of expert testimony in the courts is only likely to increase in the coming decades. As a result, legal systems will undoubtedly continue to refine their attempts to prevent dubious expert testimony from influencing legal decisions. There will, however, always be trade-offs between admitting cutting-edge expertise and prohibiting less than well-established evidence from entering the courtroom. Yet, some important lessons can already be gleaned from the experiences of courts in the U.S., Canada, Australia, the United Kingdom, and the Netherlands and their application of their evidentiary admissibility standards. First, there is limited empirical support for the proposition that juries will not be unfairly influenced by expert testimony regardless of its scientific validity. As a consequence, some level of judicial or professional screening of evidence is necessary. Reliability-based standards, such as *Daubert* in the U.S., potentially offer the most stringent scrutiny of expert testimony. Other standards such as *Frye* or general acceptance are likely to be less stringent than reliability framework but more stringent than general relevancy approaches. With the exception of the Netherlands, the United Kingdom's general relevancy framework requires the most of the trier of fact in evaluating expert testimony. This is because it is the most likely to admit questionable expert testimony and leave it to the trier of fact to determine how much weight to afford such testimony. The Netherlands' court-centered approach may require even more of the judge in evaluating expertise, in that he/she is responsible for selecting the expert, evaluating the quality of the expert's expertise, and deciding how much weight to give that evidence in reaching a decision.

Second, uniform rules for the admissibility of expert evidence, like those in the U.S. and the United Kingdom and unlike those in Canada and Australia, are likely to lead to greater consistency in similar cases. Similarly, more explicit standards and appellate review and guidance on how to apply and balance various elements of standards are likely to lead to greater consistency in judicial admissibility decision-making.

Third, reliability based expert testimony admissibility standards, like the U.S.'s *Daubert* criteria, have the potential of limiting more "junk science" from entering the courtroom. However, reliability standards require either that judges undertake substantial training in scientific methodology or that judges have impartial scientists available with whom to consult. Otherwise, reliability standards' practical use may not result in different outcomes from other evidentiary standards. Further, expert testimony regarding social and behavioral science and forensic science presents significant difficulties for reliability approaches. This expert testimony is problematic because it is not easily evaluated using commonly applied legal criteria. Moreover, even if such evaluations were possible, relevant data do not currently exist to easily answer many of the concerns of existing legal standards.

Finally, judges' beliefs, traditions, and values will continue to play an important role in adjudicating the admissibility of expert testimony unless more explicit

guidelines for how to interpret and apply existing admissibility standards are provided. This is especially problematic for admissibility standards, such as the U.S.'s *Daubert* criteria, that are more ambiguous and flexible in interpretation and application. Further, judicial beliefs, traditions, and values are likely to have their greatest influence when judges apply these standards to areas that are the most ill-suited to these criteria. As a result, the admissibility and influence of social and behavioral science expert testimony will continue to be problematic for the legal system.

References

Berger, M. (2001). Upsetting the balance between adverse interests: The impact of Supreme Court's trilogy on expert testimony in toxic tort cases. *Law and Contemporary Problems*, 64, 286–329.

Bernstein, D. (1996). Junk science in the United States and the Commonwealth. *Yale Journal of International Law*, 21, 123–82.

Bernstein, D., and Jackson, J. (2004). The *Daubert* trilogy in the states. *Jurimetrics*, 44, 351–66.

Cecil, J., and Willging, T. (1993). *Court-Appointed Experts: Defining the Role of Experts Appointed under Federal Rule of Evidence 706*. Washington, DC: Federal Judicial Center.

Chamberlain v. R., 72 F.L.R. 1, 30 (Fed. Ct. N. Terr. 1983).

Chamberlain v. R., 153 C.L.R. 521, 569 (1984).

Clark v. Ryan, 103 CLR 486 (1960).

Daubert v. Merrell Dow Pharmaceuticals Inc., 507 U.S. 579, 113 S.Ct. 2786 (1993).

Dixon, L., and 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.

Dorland, M., and Krauss, D. (2005). The danger of dangerousness in capital sentencing: Exacerbating the problem of arbitrary and capricious decision-making in capital sentencing. *Law and Psychology Review*, 29, 63–104.

Faigman, D., and Wright, A. (1997). The battered women syndrome in the age of science. *Arizona Law Review*, 39, 67–115.

Freckelton, I. (1993). Science and the legal culture. *Expert Evidence*, 2(3), 107–14.

Frye v. United States, 292 F. 1013 (D.C. Cir. 1923).

Gatowski, S., Dobbin, S., Richardson, J., Ginsburg, G., Merlino, M., and Dahir, V. (2001). Asking the gatekeepers: A national survey of judges of judging expert evidence in the post-*Daubert* world. *Law and Human Behavior*, 25, 433–58.

Gatowski, S., Dobbin, S., Richardson, J., Nowlin, C., and Ginsburg, G. (1996). The diffusion of scientific evidence: A comparative analysis of admissibility standards

in Australia, Canada, England, and the United States, and their impact on social and behavioural sciences. *Expert Evidence*, 4, 86–92.

General Electric v. Joiner, 522 U.S. 136, 118 S.Ct. 512 (1997).

Groscup, J., Penrod, S., Studebaker, C., Huss, M., and O'Neil, K. (2002). The effects of *Daubert* on the admissibility of expert testimony in state and federal criminal cases. *Psychology, Public Policy, and Law*, 8, 339–72.

HR 28 February 1989, NJ, 748; *Anatomically Correct Dolls*.

HR 27 January 1998, NJ, 404; *Shoeprint*.

Kovera, M., and 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, 574–86.

Krafka, C., Dunn, M., Johnson, M., Cecil, J., and Miletich, D. (2002). Judge and attorney experiences, practices, and concerns regarding expert testimony in Federal civil trials. *Psychology, Public Policy, and Law*, 8, 309–31.

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, 731–50.

Krauss, D. (2006). Evaluating science outside the trial box: Applying *Daubert* to the Federal Sentencing Guidelines' Criminal History Score. *International Journal of Law and Psychiatry*, 29, 289–305.

Krauss, D.**,** and Sales, B.D. (1999). The problem of helpfulness in applying *Daubert* to expert testimony: Child custody determination in family law as an exemplar. Special Issue: *Daubert* and Social Science Research. *Psychology, Public Policy, and Law*, 5, 78–100.

Krauss, D., and Sales, B.D. (2000). Legal standards, expertise, and experts in the resolution of contested child custody cases. *Psychology, Public Policy, and Law*, 6, 843–79.

Krauss, D., and Sales, B.D. (2001). The effects of clinical and scientific expert testimony on juror decision-making in capital sentencing. *Psychology, Public Policy, and Law*, 7, 267–310.

Kumho Tire v. Carmichael, 526 U.S. 137 (S.Ct. 1999).

Murphy v. The Queen, 167 CLR 94 (1989).

O'Connor, M., and Krauss, D. (2001). Legal update: New developments in Rule 702. American Psychology-Law Society, Division 41, *American Psychological Association Newsletter*, 21, 1, 1–5.

Preece v. H.M. Advocate, (unreported case) recorded in Crim. L.R. 783 (1981).

R. v. Beland, 2 S.C.R. 398 (1987).

R. v. Chamberlain, 51 A.L.R. 493, 505 (1983).

R. v. Gilmore, 2 N.S.W.L.R. 935 (1977).

R. v. Johnston, 69 C.C.C.3d, 395 (Ont. Ct. Gen. Div. 1992).

R. v. Lavallee, 55 C.C.C.3d 97 (1990).

R. v. Lewis, 29 A. Crim. R. 267 (S.C.N.T. 1987).

R. v. Maguire, Q.B. 936 (1992).

R. v. McIlkenny, 93 Crim. App. 287 (1991).
R. v. Melaragni, 73 C.C.C.3d 348 (Ont. Ct. Gen. Div. 1992).
R. v. Mohan, 89 C.C.C.3d 402 (1994).
R. v. Robb, 93 Crim. App. 161 (1991).
R. v. Runjanic, 53 A Crim R 362 (S. Aus. Ct. of App. 1991).
R. v. Turner, 60 Crim. App. 80 (1974).
R. v. Ward, 1 W.L.R. 619 (1993).
Rose v. R., 69 Austl. Crim. Rep. 1, 10 (1993).
Saks, M.J., and Koehler, J.J. (2005). The coming paradigm shift in forensic identification science. *Science*, 309, 892–5.
Sales, B., and Shuman, D. (2005). *Experts in the Court*. Washington, DC: American Psychological Association.
Sales, B., and Shuman, D. (2007). Science, experts, and the law: Reflections on the past and the future. In M. Costanzo, D. Krauss, and K. Pezdek (eds) *Expert Psychological Testimony for the Courts*. Mawah, NJ: Lawrence Erlbaum. pp. 9–31.
Sanders, J., Diamond, S., and Vidmar, N. (2002). Legal perceptions of science and expert knowledge. *Psychology, Public Policy, and Law*, 8, 139–53.
Shuman, D., and Sales, B. (1999). The impact of *Daubert* and its progeny on the admissibility of behavioral and social science evidence. *Psychology, Public Policy, and Law*, 5, 3–15.
Slobogin, C. (2007). *Proving the Unprovable: The Role of Law, Science, and Speculation in Adjudicating Culpability and Dangerousness*. Oxford: Oxford University Press.
Thompson, W.C., and Cole, S.A. (2007). Psychological aspects of forensic identification evidence. In M. Costanzo, D. Krauss, and K. Pezdek (eds), *Testimony for the Court*. Mawah, NJ: Lawrence Erlbaum. pp. 31–68.
U.S. v. Dixon, 413 F.3d 520 (5th cir. 2005).
Van Kampen, P.T.C. (2002). Expert evidence. The state of the law in the Netherlands and the United States. In P.J. van Koppen and S.D. Penrod (eds), *Adversarial versus Inquisitorial Justice: Psychological Perspectives on Criminal Justice Systems*. New York: Plenum. pp. 209–34.

Chapter 2

Psychological and Cultural Aspects of Interrogations and False Confessions: Using Research to Inform Legal Decision-Making[1]

Richard A. Leo, Mark Costanzo, and Netta Shaked-Schroer

Synopsis

False confessions are a major cause of wrongful convictions. In many countries, physical abuse and torture are still used to extract confessions from criminal suspects. Cultural orientations such as collectivism and power distance may influence the tendency to confess, and a suspect's past experience in a country that uses physical abuse during interrogations may render suspects fearful and more prone to falsely confess. After looking at interrogations outside the United States, we examine the issue of why false confessions sometime occur in the U.S. legal system. We provide an overview of the stages of a typical interrogation and provide a psychological analysis of the full array of tactics used by police interrogators. Finally, we describe several reforms that hold the potential to dramatically reduce the risk of false confessions.

Psychological and Cultural Aspects of Interrogations and False Confessions: Using Research to Inform Legal Decision-Making

Every day, police interrogate suspects. Many of these interrogations culminate in confessions to crimes. We do not know—and it may be impossible to know with precision—how many of these confessions are false (Leo and Ofshe, 1998b). By most accounts, the great majority of criminal confessions are true. We know

1 We thank David Johnson for helpful suggestions on an earlier draft of this chapter.

that most confessions are true because there is often strong corroborating evidence suggesting that the confessor actually committed the crime. However, we also know that false confessions occur with some frequency. According to the Innocence Project (2006), false confessions are the fourth most common cause (after mistaken eyewitness identification, faulty forensic evidence, and false informant testimony) of wrongful convictions (see Garrett, 2008). Also, in recent years, researchers have identified approximately 300 proven false confessions that have occurred in the post-*Miranda* era (Bedau and Radelet, 1987; Leo and Ofshe, 1998a; Warden, 2003; Drizin and Leo, 2004; Gross et al., 2005). All of these false confession cases involved DNA exonerations or other indisputable evidence of innocence (e.g., the alleged murder victim turns up alive). However, in addition to known true confessions and known false confessions, there are also a significant number of ambiguous confessions. In these ambiguous cases, a criminal suspect may claim that he or she confessed falsely, but there is no decisive evidence to corroborate or contradict the disputed confessions.

Although we do not know how many false confessions occur each year, social scientists have learned an enormous amount about the psychology of the interrogation process and about the causes of false confessions. This chapter will examine this body of research and suggest steps that could be taken by the legal system to reduce the number of false confessions, and to improve its ability to distinguish true from false confessions. Nearly all the social scientific research on interrogations has been conducted in the United States and in the United Kingdom (see Kassin and Gudjonsson, 2004). Before turning to our examination of this research, we will briefly summarize interrogation practices in other countries and explore the influence of culture on the process and outcome of interrogation.

Interrogations outside the United States

Interrogation practices and national laws governing the conduct of interrogations vary widely. Some countries place few restraints on police interrogators while others require adherence to a strict set of procedures for the treatment of criminal suspects. Reliable data on interrogation practices and criminal confessions are scarce for many countries—particularly for non-democratic nations. Worldwide, the use of physical torture is widespread as a means of extracting information from people suspected of serious crimes (Conroy, 2000). We describe interrogation procedures from five countries below.

England and Wales

Interrogations procedures in England and Wales have often been held up as a model of humane treatment. In 1986, in the wake of several high-profile false confessions, the Police and Criminal Evidence Act (PACE) became law (Home

Office, 2005). PACE prohibits police from lying to suspects about the existence of incriminating evidence (a common practice in the United States). Additionally, PACE outlaws threats of physical harm, and the use of torture or other forms of cruel, inhuman, or degrading treatment. In cases where the interrogation involves vulnerable suspects—for example, juveniles or people with mental disorders—an independent and responsible adult must be present. Finally, all interviews occurring at a police station must be audiotaped and made accessible to lawyers, judges, juries, and experts. These procedures aim to minimize police coercion, protect suspects from potential abuses, and protect police from false claims of impropriety. Although research suggests that the number of confessions has not decreased significantly since the introduction of the PACE reforms, the use of coercive tactics in interrogations has dropped substantially (Gudjonsson, 2003).

Russia

The Russian Federation adopted a new Criminal Procedure Code in 2001 and implemented the new code in 2002. In response to widespread complaints of torture and mistreatment of criminal suspects, the new codes focused on expanding the rights of suspects. According to the new procedures, suspects have the right to counsel and cannot be detained for longer than 48 hours without a court order. Suspects now enjoy a presumption of innocence, and the use of torture or other harsh treatment is prohibited. In cases where confessions are obtained, corroborating evidence is necessary to charge the suspect with a crime. Collectively, these steps ought to reduce the rate of torture and, consequently, false confessions. Nevertheless, claims of police mistreatment during interrogation remain numerous and human rights organizations have argued that police too often fail to adhere to the new laws (Amnesty International, 2006). Further, there appear to be few repercussions for the mistreatment of criminal suspects. Finally, if suspects want a medical examination to offer evidence of torture, they must first obtain a referral from an investigator or judge. This requirement complicates suspects' ability to gather proof and may discourage suspects from trying.

China

Perhaps the most infamous case of a false confession in China is that of She Xianglin, who was falsely imprisoned for 11 years after confessing to the murder of his wife. Before and after his confession in police custody, he maintained his innocence and claimed that his confession was only given in response to torture. However, it was not until his wife was found alive that his claims of police brutality were believed (Sterling, 2005). Although torture was once accepted as a means to obtain a confession in China, considerable changes have been made during the last decade. In 1996, criminal procedures were revised to enhance defendants' legal rights. It became illegal to use physical coercion to obtain confessions and defendants now

had a right to counsel. Nevertheless, suspects could be detained for up to 40 days and confessions resulting from physical coercion remained admissible in court. Research has demonstrated that although confessions are still present in about 66 percent of criminal cases, the rate of confessions has decreased since the 1996 rulings (Lu and Miethe, 2003).

In 1997, additional changes were made. The courts ruled that voluntary and sincere confessions made prior to any criminal investigation could serve as a mitigating factor in sentencing decisions. Although torture had already been ruled illegal, it was not until 2005 that the courts decided confessions obtained as a result of torture or coercion were not admissible as evidence. It was also held that suspects must be questioned by prosecutors to assure they were not mistreated by authorities. Finally in 2006, China ruled that all interrogations of job-related crimes must be videotaped to protect against coercion and corruption. Although coercion persists, these legal changes indicate greater concern for defendant rights (Russ, 2005).

Japan

Japanese criminal law affords suspects a presumption of innocence, the right to remain silent, and the right to counsel. Torture is rarely reported, with figures decreasing yearly (U.S. Department of State, 2005a). Nonetheless, some criminal justice practices exert psychological pressure on suspects that elicits false confessions (Johnson, 2002; Leo, 2002). Suspects may be held for up to 23 days prior to indictment and are unlikely to receive bail if they maintain their innocence or remain silent (Foote, 1992). Although the law offers suspects the right to counsel, prosecutors can limit access to an attorney prior to indictment (U.S. Department of State, 2005a). If suspects are unable to afford a lawyer, they will not receive one until after their indictment, although the relatively new "duty lawyer" system gives some such suspects access to an attorney. Additionally, the right to remain silent carries with it a duty to endure questioning. Concerns have been raised because approximately 90 percent of criminal cases in Japan involve confessions (Ramseyer and Rasmusen, 2001). This high rate of confessions probably results from a combination of cultural expectations, police pressure, or a tendency to prosecute only cases where strong evidence exists against the defendant (Johnson, 2002).

Mexico

In Mexico in 1990, Manuel Manriquez was detained by police in connection with a murder, without an arrest warrant. He was brutally beaten, burned, and had gas and chili peppers placed in his nose. Fearing that his survival depended on it, he confessed to the murder. He later withdrew this confession, maintaining that it was false and only given as a result of torture. Although the confession was the only evidence in this case, Manriquez was convicted (CPIHR, 2000). Similar stories from cities throughout Mexico have been reported by international human rights groups.

Threats, physical torture, beatings, burning of the skin, and suffocation have been used to obtain confessions from criminal suspects in Mexico. In 1991, Mexican law was changed to prohibit the use of torture to obtain confessions and prevented confessions obtained through the use of torture from being admissible in federal courts throughout Mexico and in any court in Mexico City. However, this did not curb the use of torture during interrogations, and confessions remained the primary evidence in most cases (U.S. Department of State, 2005b). In 1995, the Supreme Court of Mexico held that confessions had to be corroborated by additional evidence to prove guilt. Unfortunately, judges did not always conform to this ruling and, in many cases, the corroborating evidence was the testimony of the arresting officer. A decade later, Mexican law changed again to allow suspects the right to counsel from the time they come into contact with the prosecutor. However, police can still hold suspects of "serious" crimes in jail until trial. Judges do not have the discretion to release suspects prior to trial even if they pose little to no risk. As a result, 40 percent of prisoners being held in Mexico are yet to be convicted (Human Rights Watch, 2006).

The Role of Culture in Interrogations

Culture has been defined as "the set of attitudes, values, beliefs, and behaviors shared by a group of people and communicated from one generation to the next" (Matsumoto, 1997, p. 4). Some culturally-based "attitudes, values, beliefs, and behaviors" are likely to impinge on the process of interrogations and potentially increase the risk of false confessions.

One important way of distinguishing between cultures is by locating them along a continuum ranging from individualistic to collectivistic. Individualist societies, such as the United States and the United Kingdom, function within a framework of individuality, personal identity, and individual freedom (Triandis, 1994). In these societies, independence is encouraged, and individuals emphasize their personal goals while deemphasizing the wellbeing of their group. In contrast, most countries in Asia and Latin America hold collectivist values. In collectivist societies, priority is placed on the welfare of the group, rather than on the needs of the individual (Triandis, 1994; Westwood et al., 1992). Collectivist cultures exhibit a strong orientation towards interdependence and an emphasis on community relationships. A suspect's individualistic or collectivistic orientation may influence his behavior during interrogation. So might the related cultural trait of "power distance." According to Hofstede (1980), power distance refers to the degree to which individuals recognize and accept hierarchies and unequal distributions of power. In low power distance cultures, authorities tend to endorse equality and mutual respect. Conversely, high power distance cultures expect individuals to comply with the demands of people who hold positions of authority and to embrace the role of subordinate. High power distance cultures tend to place great trust in authorities, believing that their decisions

are made for the benefit of the entire group. In his analysis of 40 countries, Hofstede (1980) demonstrated that low power distance cultures tended to be individualistic, while high power distance cultures tended to be collectivistic. Extending these findings, Kemmelmeier et al. (2003) identified a strong relationship between collectivism, conventionalism, and submission to authority.

Although there has been surprisingly little research conducted directly examining the influence of culture on confessions, the separate research literatures on cultural attributes and false confessions, in combination with case examples, strongly suggest that culture plays a significant role in the psychology of confession. There is some research suggesting that people from collectivist cultures are particularly likely to waive their *Miranda* rights due to lack of familiarity with the American legal system combined with values that require obedience towards police (Einesman, 1999). Research indicates that people who endorse collectivist values are more likely to exhibit behaviors of obedience and trust towards authorities (Foote, 1992), and there is direct evidence that people who score high on measures of compliance to authority are at higher risk for false confession (Gudjonsson, 2003). Taken together, these findings suggest that people from collectivist cultures may be more obedient to police officers—the authority figure in the interrogation room.

The relationship between cultural values and confessions was at issue in the case of *Liu v. State* (1993). The defendant maintained he was both unaware of his right to remain silent, and that his behavior was guided by Chinese cultural values which required that he fully cooperate with police and not question the behavior or intentions of people in authority. Although he was informed of his *Miranda* rights, Liu maintained that his cultural background prevented him from invoking these rights. The courts considered this claim, but ruled that his confession was admissible in court.

Just as cultural values influence the psychology of confessions, relevant past experience in a different culture or country may cause suspects to make damaging assumptions in the interrogation room. Suspects may be unable to disregard their negative experiences or perceptions of interrogations by police in their home country. Specifically, suspects may assume that interrogation practices from their country of origin—which may include physical abuse, deprivation, or torture—are typical or at least possible in other nations. Consequently, suspects may acquiesce to police demands to confess (even if the confession is false) in order to avoid the physical harm they judge to be a likely consequence of failure to cooperate. This type of reaction was observed in a Vietnamese defendant in *Le v. State* (1997). Le's beliefs and experiences with interrogations in Vietnam led him to conclude that denials or remaining silent would inevitably result in some form of torture. He worried that if he refused to speak, the interrogators would turn him upside down and push fish salt up his nose.

Latin American defendants have reported similar experiences. For example, a defendant from Mexico complied with all police requests for an admission of guilt because he assumed that unquestioning compliance was necessary to avoid abuse

or torture (*United States v. Zapata*, 1993). Similarly, a man from Cuba waived his *Miranda* rights because he presumed that failure to do so would result in a physical beating (*People v Merrero*, 1984). Similar to *Liu v. State* (1993), the courts did not accept these claims, and the defendants were subsequently convicted. In sum, the connection between physical torture and interrogation in many countries may lead suspects from these countries to comply with police requests and admit guilt. Fear of abuse or torture, combined with cultural norms of obedience to authority are likely to increase proneness to confess. Further research in this important area would give us a fuller understanding of how culture shapes the behavior of police and suspects during interrogations. Such research could also help determine the best procedures for protecting suspects' rights, and help to reduce the likelihood of false confessions.

Research on these issues is critical in the U.S. because, beginning in 1965, Congress passed immigration laws that increased the ethnic, racial, religious, and cultural diversity of the population. By 2003, there were 33.5 million foreign-born people in the United States (Larsen, 2004). In addition to retaining their native language, many of these immigrants did not abandon the values and beliefs associated with their country of origin. As a result, the United States, although described as an individualist society, is inhabited by people of multiple cultures. Researchers have continuously argued that attitudes and behaviors cannot be fully understood without consideration of their cultural context (e.g., Triandis, 1994). The issue of culture values and past experience is likely to become a growing concern for the U.S. legal system.

False Confessions in the U.S. Legal System

There is no worse error in the criminal justice system than the wrongful conviction of an innocent person. And, in theory, there is no outcome that the U.S. legal system is more engineered to prevent. The error of convicting an innocent person produces at least three tragedies: first, an innocent person suffers the pains of prosecution, conviction, and imprisonment. Second, the true perpetrator remains free to commit additional crimes. Third, decision-makers in the legal system—police officers, prosecutors, judges, and jurors—must live with the knowledge that they were responsible for the first two tragedies.

In the U.S. legal system, there are four types of police error that lead to false confessions: 1) misclassifying an innocent suspect as guilty and then subjecting him to a guilt-presumptive interrogation; 2) using coercive interrogation techniques to overcome resistance, gain compliance, and elicit a false admission; 3) failing to understand the vulnerabilities of a suspect; and 4) contaminating the innocent suspect's knowledge of the non-public crime facts (thereby making a confession seem reliable because the suspect "knows facts only the true perpetrator would know").

The first step in the process of eliciting a false confession occurs when a detective erroneously decides that an innocent person is guilty. Whether to interrogate or not is therefore a critical decision point in the investigative process. If police did not interrogate innocent people, they would never elicit false confessions. This "misclassification error" is typically a consequence of an interrogators' mistaken belief that he can tell when a suspect is being deceptive (Davis and O'Donahue, 2003). This faulty judgment—often based on an interrogator's intuition or a suspect's nonverbal behavior (Meissner and Kassin, 2004)—then leads to a coercive interrogation designed to make the presumably deceptive suspect stop lying and admit to involvement in a crime (the mythology of the interrogator as human lie detector is discussed later in this chapter).

Once detectives have misclassified an innocent person as guilty, they sometimes resort to coercive interrogation techniques (Leo, 2008). Because no credible evidence is likely to exist against an innocent but misclassified suspect, obtaining a confession becomes unusually important—detectives may desperately need a confession to successfully build a case against the suspect, especially in homicide cases (Gross, 1996). Once interrogation commences, the primary cause of police-induced false confession is psychologically coercive interrogation methods. Cumulatively, these methods (discussed in detail below) cause a suspect to perceive that he has no choice but to comply with the interrogators' demands. Regretfully, most interrogation training manuals—including the widely used and influential manual by Inbau et al. (2001)—give little thought to how the methods they advocate sometimes lead the innocent to confess. Instead, police trainers and interrogators, in the face of abundant empirical evidence to the contrary (Kassin and Gudjonsson, 2004; Ofshe and Leo, 1997), assert that their methods will only produce confessions from the guilty.

Even though psychological coercion is the primary cause of police-induced false confession, suspects clearly differ in their ability to withstand interrogation pressure and thus in their susceptibility to making false confessions. Two personality traits (suggestibility and compliance) and three individual characteristics (retardation, youth, and mental illness) have been shown to increase the risk of false confessions. Individuals who are highly suggestible tend to accept and repeat back information that is suggested to them, and also tend to have poor memories, higher levels of anxiety, and tend to be less assertive. Compliant individuals tend to be conflict avoidant, acquiescent, and eager to please authority figures (Gudjonsson, 2003). These two personality traits exert an especially strong influence on behavior when a suspect is under stress, sleep deprived, or suffering from drug or alcohol withdrawal (Kassin and Gudjonsson, 2004; Blagrove, 1996).

The largest study of proven false confessions to date has revealed that youth is the personal characteristic that puts a suspect at highest risk for false confession (Drizin and Leo, 2004). In their sample of 125 proven false confessions, Drizin and Leo found that 63 percent of false confessors were under the age of 25. Juveniles and young adults tend to be intellectually immature, naively trusting of authority, acquiescent, and relatively eager to please adults in positions of authority. These

traits predispose them to be submissive when questioned by police. Young suspects typically lack the cognitive maturity to fully understand the nature or the gravity of an interrogation or the long-term consequences of their responses to police questions. Juveniles are less capable of withstanding interpersonal stress and thus more likely to perceive aversive interrogation as intolerable. They may tell an interrogator what he or she wants to hear, merely to escape the situation (Redlich, 2004). Mentally retarded suspects have shorter attention spans, poorer memory, reduced conceptual skills, and are more easily confused. They are less likely to understand the long-term consequences of making or agreeing to an incriminating statement, and they are not likely to understand that the police detective who appears friendly is really their adversary (Ellis and Luckasson, 1985; Cloud et al., 2002). Similarly, the mentally ill possess psychiatric symptoms that make them likely to conflate true and false information. These symptoms include faulty reality monitoring, distorted perceptions and beliefs, impaired ability to distinguish fact from fantasy, proneness to feelings of guilt, anxiety, mood disturbances, and lack of self control (Kassin and Gudjonsson, 2004).

As we have argued elsewhere (Costanzo and Leo, 2007), it is useful to conceptualize false confessions as the products of situation-person interactions: For intellectually normal adults—the vast majority of criminal suspects—it is the coerciveness of the interrogation that generates a false confession. However, for highly vulnerable suspects (juveniles, the mentally ill or retarded) less coercive interrogations can lead to false confessions. Thus, there are two sets of risk factors: coercive interrogation techniques and individual vulnerabilities. The idea of a risk factor is, of course, probabilistic—the presence of coercive interrogation techniques or individual vulnerabilities makes false confessions more likely, but not certain.

The best way to prove that a confession is false is to have DNA evidence that exonerates the confessor. Unfortunately, the biological evidence that permits DNA testing is only available in a minority of criminal cases (Scheck et al., 2000). If, aside from the confession, there is no strong independent evidence that implicates or excludes the confessor, we must rely on a close analysis of both the interrogation process itself and of what is called the "post-admission narrative" (PAN). A confession is more than an "I did it" statement. It also consists of a subsequent narrative that contextualizes and attempts to explain the "I did it" statement. In false confessions, the PAN is the story that gets wrapped around the admission and thus makes it appear, at least on its face, to be a compelling account of the suspect's guilt.

Unless an innocent suspect has learned about the crime facts from community gossip or the media, he or she will not know either the mundane or the dramatic details of the crime (Leo and Ofshe, 1998a). Consequently, the false confessor's PAN will likely be replete with errors and will be either inconsistent with or contradicted by the objective case evidence (Leo et al., 2006). That is, *unless*, the answers are implied, suggested, or explicitly provided to the suspect by the interrogator. The problem of police leaking information about the crime to the suspect is known as *contamination* of the PAN. Contamination does occur, whether advertently or

inadvertently, in many false confession cases (Leo and Ofshe, 1998a; Leo et al., 2006). For example, sometimes interrogators will accuse a suspect of lying by telling him that the details of his denial are incorrect and then tell the suspect what the correct details are. Sometimes interrogators will even show evidence to a suspect (e.g., victim or crime scene photographs) thereby educating him about how and why the police believe the crime occurred.

The reliability of a suspect's confession can be evaluated by analyzing the fit—or lack of fit—between the PAN and the crime facts (Leo et al., 2006). The goal is to determine whether the suspect's PAN reveals the presence (or absence) of guilty knowledge and whether it is corroborated (or disconfirmed) by objective evidence. As Leo and Ofshe (1998a) have pointed out:

> There are at least three indicia of reliability that can be evaluated to reach a conclusion about the trustworthiness of a confession. Does the statement (1) lead to the discovery of evidence unknown to the police? (e.g., location of a missing weapon that can be proven to have been used in the crime, location of missing loot that can be proven to have been taken from the crime scene, etc.); (2) include identification of highly unusual elements of the crime that have not been made public? (e.g., an unlikely method of killing, mutilation of a certain type, use of a particular device to silence the victim, etc.); or (3) include an accurate description of the mundane details of the crime which are not easily guessed and have not been reported publicly? (e.g., how the victim was clothed, disarray of certain furniture pieces, presence or absence of particular objects at the crime scene, etc.). (p. 438)

The absence of such corroboration, however, does not necessarily mean that a confession is unreliable. Such evidence may not exist in every case—bloody clothes may be burned, weapons may be disposed of, and money can be spent (Leo et al., 2006).

Stages in the Process of Psychological Interrogation

The short answer to the question, "why do suspects confess?" is "police interrogation." Suspects seldom confess spontaneously but virtually always confess in response to some amount of pressure to stop denying and start admitting. As Inbau et al. (1986) note, "it is impractical to expect any but very few confessions to result from a guilty conscience unprovoked by an interrogation" (p. 112). The conditions and techniques of interrogation are inherently distressing; they are designed to induce anxiety in order to break down the suspect's anticipated resistance and motivate compliance (Inbau et al., 1986). Accusatorial interrogation may also induce fear, confusion, physical exhaustion, and mental fatigue. American interrogation is a multi-stage process that relies on basic principles of social influence, deception, and coercion to manipulate a suspect's perception of his situation, and his available options.

There are four distinct stages that comprise the interrogation process in the U.S. (See Leo, 2008 for a fuller analysis): 1) softening up the suspect; 2) moving past the *Miranda* warnings; 3) moving the suspect from denial to admission through the use of psychological techniques; and 4) construction of the post-admission narrative.

Stage 1: Softening Up the Suspect

Police interrogations in the U.S. often begin with an ambush. If the suspect is not already in custody, detectives will typically call and ask him to "voluntarily" come into the police station for some questioning, offering to escort him if necessary. Investigators will tell the suspect that they need only to ask him a few questions and/or that they need his help in solving a crime, invariably promising not to take much of his time. The detective strategically misrepresents the nature and purpose of the questioning at this stage in order to disarm the suspect and catch him off guard. During the softening up phase, the interrogator seeks to hide the fact that he has prejudged the suspect to be guilty and that his singular goal is to incriminate the suspect and assist the state in successfully prosecuting and convicting him.

Once in the interrogation room, the suspect is disconnected from familiar environments, friends, family, or any other source of social support that might psychologically empower the suspect to resist the interrogation process. The sometimes surreal, isolated world of the interrogation room allows detectives to control the pace and tactics of the interrogation, and to maximize the suspect's vulnerability to police pressure. One or more detectives dressed in plain clothes typically try to establish rapport with the suspect by asking background questions and engaging in small talk. Interrogators may even flatter or ingratiate the suspect to create the illusion of a non-threatening, non-adversarial environment and encounter. Detectives attempt to give the impression that they and the suspect will be engaged in a simple information exchange designed to help the investigator understand and solve the crime.

Stage 2: Moving Past the Miranda Warnings

In 1966, in the case of *Miranda v. Arizona*, the Supreme Court held that police must forewarn suspects of their rights to silence and appointed counsel before any custodial questioning can legally commence. The typical four-part *Miranda* warning thus reads as follows: 1) You have the right to remain silent; 2) Anything you say can and will be used against you in a court of law; 3) You have the right to an attorney; 4) If you cannot afford an attorney, one will be appointed to you free of charge. In addition, the Court held that the state must demonstrate that the suspect's waiver of these constitutional rights was made "voluntarily, knowingly and intelligently" (*Miranda v. Arizona*, 1966: p. 440). As a result, police interrogators were directed to follow up the fourfold *Miranda* warnings with two further questions that are

designed to elicit an explicit waiver: 1) Do you understand these rights?; and 2) Having these rights in mind, do you wish to speak to me?

Perhaps surprisingly, *Miranda*'s impact on police, on criminal suspects, on confession and conviction rates has been negligible (Leo, 2001a). This may be due to the many strategies that police interrogators use to avoid, circumvent, nullify, and sometimes even violate *Miranda* and its invocation rules in their pursuit of confession evidence. For example, sometimes police interrogators avoid having to give *Miranda* warnings by recasting the interrogation as a custodial interview. Police accomplish this by redefining the circumstances of questioning so that the suspect is not in custody—for example, telling the suspect that he is not under arrest or that he is free to leave—since *Miranda* warnings are only legally required when a suspect is under arrest or in custody (*Berkemer v. McCarty*, 1984). Other times police detectives may read the fourfold *Miranda* warnings but do not ask the suspect whether he understands these rights or wishes to waive them, but instead simply launch directly into interrogation as if the suspect's consent is a *fait accompli*. The courts have upheld this practice, claiming that such suspects have "implicitly" waived their rights (*North Carolina v. Butler*, 1979). Sometimes detectives elicit explicit waivers by minimizing, downplaying, or de-emphasizing the significance of the *Miranda* warnings, thus obscuring the adversarial relationship between the interrogators and the accused (Leo and White, 1999). And sometimes interrogators explicitly violate *Miranda* by either not reading the suspect his rights or continuing to interrogate "outside *Miranda*" once the suspect has invoked them (Weisselberg, 1998).

Stage 3: Moving the Suspect from Denial to Admission

Once the suspect has waived his *Miranda* rights, the tone, the content, and the character of the interrogation process may change dramatically. It is typically at this point that the interrogation process becomes accusatorial and that the detective shifts from asking the suspect questions to telling him the answers and imploring him to confess. Regardless of the detective's interpersonal style (which may range from friendly to aggressive), he relies on two underlying strategies: 1) creating the illusion that he and the suspect share the goal of "helping" the suspect; and 2) attempting to transform the suspect's perception of his situation such that he comes to see the act of confessing as offering the easiest escape from an otherwise hopeless situation. Although there are hundreds of specific interrogation techniques, police interrogation can be reduced to a few basic tactics that usually occur in predictable, repetitive sequences. The logic of interrogation is pithily captured in the following statement taken from a Los Angeles Police Department interrogation training manual, which exhorts detectives to tell suspects: "You did it. We know you did it. We have overwhelming evidence to prove you did it. But the reason makes a difference. So why don't you tell me about it" (Los Angeles Police Department training manual, undated).

Accusations Interrogation is a guilt-presumptive process. The purpose of interrogation is not to evaluate the suspect's denials, truth claims, alibis, or assertions of innocence, but to overcome the suspect's anticipated resistance and elicit an incriminating statement (Inbau et al., 2001). There are two general types of accusations made during interrogation: accusations of committing the offense in question (or some version of it) and accusations of lying. Interrogators typically express these accusations with unyielding confidence. A common accusation strategy, for example, is to tell the suspect that the purpose of the interrogation is not to debate whether he committed the crime, but to understand *why* he did it. Through frequent and repeated accusations police interrogators create the impression that the norms of interrogation demand that the suspect prove his innocence to the satisfaction of the interrogator before the process can terminate. The impossibility of convincing the interrogator that he is innocent helps to break down the suspect's resistance and create a perception of hopelessness, especially as the length of interrogation grows.

Attacking denials Police aggressively challenge a suspect's denials, especially early in the interrogation process before the suspect can establish a pattern of denial. The most basic method is to cut denials off at the source, interrupting the suspect and/or preventing the denials from being uttered in the first place. In addition, interrogators sometimes prevent a suspect from uttering a denial by talking over him and/or maintaining a monologue so that the suspect has no opportunity to insert denials into the conversation. Once the suspect begins to utter denials, however, interrogators may attack them in one of several different ways. Interrogators may tell the suspect his denials are simply lies, or not believable, or contradicted by the existing case evidence, or contradicted by some aspect of his previous account, or that his alibis are logically impossible. Generally, the goal is to persuade the suspect that his denials do not add up or make sense and thus no one—especially prosecutors, judges, and juries—will believe them. Here too interrogators will try to foreclose the possibility that the suspect can deny the accusations by telling him that the purpose of questioning is not to discuss whether the suspect committed the crime, but why. This strategy is an attempt to persuade the suspect that his fate has already been sealed and thereby shift the interrogation from denial to explanation. The onslaught of accusations and attacks on denials, especially in conjunction with other interrogation techniques, is intended to break down a suspect's resistance by silencing him, rendering him passive, and ultimately inducing resignation, despair, and hopelessness.

Evidence ploys An evidence ploy is any attempt by interrogators to make the suspect believe that they possess incriminating evidence against him (Ofshe and Leo, 1997). Evidence ploys may be either true or false: if interrogators possess actual evidence against a suspect (e.g., eyewitness identifications, accomplice statements, DNA, etc.), they will confront him with it as they press him to confess (Leo, 1996; Moston et al., 1992). If interrogators do not possess any actual evidence against a suspect,

they will usually pretend that they do. Most suspects do not know that police can lie about evidence during interrogation, and many are shaken when police describe evidence that seemingly proves their guilt. Evidence ploys are used to convince a suspect that the case against him is so overwhelming and conclusive that he has no choice but to admit to the accusation because (he is told or led to believe) no one will believe his assertions of innocence.

Investigators may simply tell the suspect that other people have said that he committed the crime. These incriminating statements (real or fabricated) might be attributed to victims, or eyewitnesses and bystanders, or accomplices, and/or other relevant third parties. Interrogators often represent such statements as conclusive of the suspect's guilt, shifting the informal burden onto the suspect to prove his innocence to the interrogator, but then refusing to credit the suspect's denials. Often there is no way for the suspect to dispute the veracity of the victim's accusations or identification other than to say that the victim is lying or mistaken. For statements from victims, the interrogator is likely to counter that the victim has no reason to lie. For eyewitness identifications, interrogators may tell a suspect that there were multiple eyewitnesses or that the eyewitnesses picked him out of a photo line-up. Eyewitness evidence ploys are also often buttressed by technological supports, such as hidden surveillance cameras. In multi-perpetrator crimes, interrogators might tell the suspect that his co-perpetrator(s) or accomplice(s) have identified him as participating in or masterminding the crime in question. Sometimes interrogators tell a suspect that even his family, parents, wife, child, or close friends do not believe his denials and this constitutes additional evidence of his guilt.

Other evidence ploys involve real or invented forensic, medical, and/or technological evidence. These ploys may be especially effective because they carry the weight, authority, and presumed certainty of modern science and technology. Interrogators may allege that they have print evidence from virtually any part of the body—fingerprints, palm prints, foot prints, shoe prints, head prints, teeth prints, and (as a false evidence ploy in sexual assault cases) even penis prints. This type of evidence ploy trades on the popular mythology that print evidence is unique to each individual, never wrong and therefore dispositive of the suspect's guilt (see Cole, 2001). Suspects may be confronted with other types of forensic evidence such as ballistic or scent evidence, or medical evidence, ranging from the opinions of doctors and other medical personnel to physical evidence such as hair, blood, semen, and DNA. Interrogators also often confront suspects with the false or inconclusive results of so-called "lie-detector" technologies such as the traditional polygraph and the more recent Computer Voice Stress Analyzer (CVSA). Interrogators often tell a suspect who continues to deny his guilt that the polygraph or CVSA will allow him to prove his innocence. Invariably, they then confront the suspect with the seemingly scientific results of a failed exam that they represent as indisputable "scientific" proof of his guilt.

Escalating pressure From the minute the interrogation begins, police apply psychological pressure. Interrogators sometimes, for example, invade a suspect's personal space, moving uncomfortably close to the suspect. Interrogators may alternate displays of sympathy with displays of hostility, rewarding the suspect with friendliness when he says what they want to hear but punishing him with anger when he does not. Interrogators will sometimes also try to prevent the suspect from making a response (e.g., verbalizing a denial) or engaging in an activity (e.g., smoking a cigarette, taking a break) that would allow him to release anxiety or diminish interrogative pressure until or unless he complies with their wishes. In addition to creating pressure through the barrage of repeated confrontation and interruption, investigators sometimes raise their voice, yell and/or relentlessly badger the suspect. The goal is not only to distress and wear down the suspect, but also to create the perception that the interrogation will not cease unless or until the suspect complies with the interrogators' demands to say "I did it" (Ofshe and Leo, 1997).

Time perception Another technique interrogators use to exert pressure on suspects is to manipulate their perceptions of time. One tactic is to make the suspect perceive that the interrogation works, as in high pressure sales, like a time-limited offer: the interrogators present the suspect with an "opportunity" to tell his side of the story. However, that opportunity is only available now and will expire once the interrogation ends. If the technique is successful, the suspect will think that his situation will worsen, perhaps irreversibly, if he does not act now and comply with the interrogators' wishes or demands. Or, interrogators may try to make the suspect perceive that the interrogation will go on much longer, if not indefinitely, if he does not comply with their wishes. Sometimes interrogators use time as a lure, leading a suspect to think that the interrogation is almost over, only to keep prolonging it, creating the expectation that if the suspect just complies, it will finally, at long last, be over.

Offering inducements for admissions Interrogators use incentives to persuade a suspect that he will obtain some psychological, material and/or legal benefit or reward and/or avoid a corresponding harm if he makes an incriminating statement. Ofshe and Leo (1997) have identified three general types or categories of incentives. *Low-end* incentives refer to interpersonal or moral appeals interrogators use to convince a suspect that he will feel better if he confesses. For example, interrogators may tell a suspect that the truth will set him free, or that confessing will relieve his anxiety or guilt, or that it is the moral or Christian thing to do, or that it will improve his standing in the eyes of the victim or the community. *Mid-range* incentives refer to appeals that interrogators use to focus the suspect's attention on the functioning of the criminal justice system in order to get the suspect to reason that his case is likely to be processed more favorably if he complies and confesses. For example, interrogators may ask the suspect how a judge and jury are likely to react, if he does not demonstrate remorse and admit guilt. Finally, *high-end* incentives (i.e., implicit

and/or explicit promises and threats) communicate the message that the suspect will receive a lighter prison sentence and/or some form of police, prosecutorial, or juror leniency if he confesses, but that the suspect will receive a longer sentence if he does not. Interrogators often imply or directly state that the alleged benefits of compliance and confession will disappear if the suspect does not take advantage of the interrogators' offer(s) during the interrogation.

Reframing the interrogation as an "opportunity" to help oneself As an opening move, the assertion that "this is your opportunity to tell your side of the story," is an attempt to lead the suspect to believe that his version of what occurred will only be entered into the official record if he speaks to police, and that he will forego this "opportunity" if he chooses not to participate. This assertion also implies that the suspect can talk his way out of the interrogation process by convincing the interrogator of his innocence. With this bait, the investigator is able to lock the suspect into an account that they can subsequently attack as implausible, inconsistent, and/or untenable as they seek to elicit incriminating statements. Even if they do not obtain an admission or confession, however, interrogators know that any account the suspect gives is likely to advantage the state because it increases the probability that the prosecution will be able to impeach him at trial if he takes the witness stand. Police seek to reframe the interrogation as an opportunity not only because they wish to divert his attention from the true purpose of interrogation (i.e., incrimination), but also because they seek to persuade the suspect that they can only help him (and thus he can only help himself) if he continues to talk to them. Interrogators may, for example, tell a suspect that in order to help him they need first to hear his side of the story.

The tactic of reframing the interrogation as an "opportunity" to help oneself simultaneously misrepresents the role of police as allies instead of adversaries, and suggests that the interrogators can and will help the suspect minimize the social or legal consequences his crime. Interrogators may explicitly tell a suspect that they can help him by how they write up their report following questioning, or by what they tell the prosecutor prior to charges being filed, or by how they testify before a judge or jury at trial. Interrogators may also explicitly offer to help a suspect in ways that extend beyond the criminal justice system, such as arranging for counseling or help from social services. However, these benefits can only be realized if the suspect complies with the interrogators' demands for an admission.

Constructing exculpating scenarios To make suspects more comfortable with making an admission, interrogators suggest reasons why, or scenarios in which, the suspect might or could have committed the crime. In interrogation manuals, this technique is known as using "themes" (Inbau et al., 2001; Senese, 2005). These "themes" are actually scenarios or rationalizations that morally, psychologically, and/or legally downplay, excuse, and/or justify the suspect's act. Interrogators use scenarios to persuade a suspect that if he makes an admission, he (with the

interrogators' help) can control how that admission is framed to other audiences (prosecutors, judges, juries, his friends and family, the victim, the victim's friends and family, and the media if it is a high-profile case). Interrogators use scenarios to persuade the suspect that he can construct or agree to an account that explains his motive for committing a crime that will not only portray him in the most sympathetic light possible, but that will also minimize his social, moral, and/or legal culpability and thus his blameworthiness for the alleged act. Scenarios can be a subtle yet powerful inducement. They work by shifting the blameworthiness from the suspect being interrogated to another person, or to the social circumstances that caused the act, or by redefining the act itself. Scenarios relocate the blame to something or someone else. For example, it is common in murder investigations for interrogators to suggest that the suspect killed the victim in self-defense. Because self-defense is not a crime, this scenario suggests that the suspect will not be charged or punished for admitting to it. It is also common for interrogators to suggest that the suspect killed the victim accidentally, as opposed to intentionally, thus lessening the criminal act and (through pragmatic implication) lowering the punishment if the suspect agrees to the interrogators' accident scenario (Ofshe and Leo, 1997; Kassin and McNall, 1991). Homicide investigators often tell the suspect that if he does not adopt the interrogators' accident or self-defense scenario, he will be perceived as committing the killing intentionally, with premeditation and in cold blood. Delivered against the backdrop of the techniques that have preceded them—isolation, accusation, attacks on denials, evidence ploys, etc.—these scenarios are intended to motivate the suspect who already feels trapped and powerless to admit to the crime because he is being offered a chance to mitigate his punishment by choosing between the better of his only two remaining choices.

The promises/threat dynamic Promises and threats often go together because they imply one another: a promise of leniency implies a threat of harsher treatment if the suspect fails to comply with the interrogators' wishes, just as a threat of harsher treatment implies a promise of leniency if the suspect does comply with the interrogators' wishes.

Kassin and McNall (1991) have demonstrated that minimization techniques communicate promises of leniency through "pragmatic implication" (i.e., indirectly or implicitly). Ofshe and Leo (1997) have illustrated how some of the scenarios widely used by American interrogators, such as the accident and self-defense scenarios mentioned above, effectively communicate and create expectations of leniency in exchange for admission. Use of the promise/threat dynamic helps serve to communicate several key points to the suspect: 1) although the existing evidence irrefutably establishes his guilt, the investigators' decision whether to continue to detain him, the prosecutor's decision whether to charge him (as well as with what and how many counts), and the judge's and/or jury's decision whether to sentence him (and for how long) have not yet been determined; 2) these decisions are negotiable, not fixed; 3) if he continues to deny, he will be treated more harshly by all these

individuals, who hold his future in their hands; and 4) if he admits to some version of the crime, he can negotiate a more favorable outcome and mitigate his culpability and/or punishment.

Stage 4: Construction of the Post-Admission Narrative (PAN)

An admission is a statement, but a confession is a story. Like all stories, some confessions are more plausible than others. In order to successfully build a case against the suspect, interrogators must elicit a believable account of the suspect's participation in and knowledge of the crime. Sometimes this involves little work for interrogators: some suspects will, with little prompting or pressure, provide a detailed narrative of how and why they committed the crime. Other suspects may only reluctantly give interrogators the account they are looking for or they may lie (or appear to be lying) about crime facts or details. Substantial additional pressure may be required if the interrogators believe that the suspect is refusing to provide details of the offense, or if the suspect's account does not match the interrogators' expectations or pre-existing beliefs about how or why the crime occurred. This need for more pressure is most acute in false confession cases—if interrogators have coerced an admission from an innocent suspect, they inevitably have difficulty getting the details from him because (absent their suggestions) he typically does not know the details. During the post-admission phase, the interrogator and the suspect jointly construct a persuasive narrative of the suspect's culpability, thereby transforming the fledgling admission into a fully-formed confession. While there are some detectives whose only goal in the post-admission phase is to obtain an uncontaminated account (Napier and Adams, 2002), the post-admission-narrative is not usually simply discovered or elicited. Interrogators seek to shape and flesh-out the narrative in order to confer legitimacy on the police procedures through which the admissions were obtained and to deflect any eventual challenges to the confession's authenticity. An effective PAN incriminates the suspect and builds a case against him that will ensure successful prosecution and conviction.

Detectives typically try to shape the suspect's PAN so that it will be maximally persuasive to the target audience—lawyers, judges, and jurors. Interrogators often seek to build into the PAN the types of statements and cues that people associate with authenticity, individual volition, and truth-telling (Costanzo and Leo, 2007). A maximally persuasive PAN should have five elements: 1) a coherent story-line; 2) motives and explanations; 3) crime knowledge; 4) expressions of emotion; and 5) acknowledgements of voluntariness.

First, to be plausible and persuasive, the confession narrative requires a coherent story line, characters, an unfolding plot (with a beginning, middle, and end), a description of actions and events in sequence, and explanations. When third-party decision-makers recognize a story line developing, they are cued to interpret subsequent information as consistent with that story (Amsterdam and Bruner, 2000).

Second, motive is an essential element of a convincing story (Pennington and Hastie, 1993). Research indicates that for a variety of case types, including capital murder (Costanzo and Peterson, 1994), rape (Olsen-Fulero and Fulero, 1997), and sexual harassment (Huntley and Costanzo, 2003), attorneys present motives for crimes and jurors seek to find motives for crimes. A motive (e.g., greed, lust, jealousy, revenge, betrayal) completes the story of the crime and increases the likelihood of a conviction.

Third, the presence of crime-related, detailed knowledge gives the suspect's confession verisimilitude. This kind of detailed knowledge makes the confession compelling, especially if the suspect's knowledge is perceived as vivid, accurate and unique (Kassin, 2006; Leo et al., 2006).

Fourth, confessions that are textured with emotion (anger, fear, sorrow, remorse, shame, and regret) and that are accompanied by behavioral displays (such as crying) humanize the admission, making it appear to be the natural product of lived experience (Katz, 1999). Interrogators recognize the probative force of emotion-laden confessions, and sometimes pressure or direct suspects to write an "apology note" to the victim (even if the victim is dead). This leads to a remorseful statement of guilt in the suspect's own handwriting and creates the appearance of an additional piece of evidence (the apology note) that seemingly corroborates the suspect's admission (even though in reality it does nothing but repeat it). In addition, investigators also impute—in their written police reports, pre-trial and trial testimony, and, in high-profile cases, interviews with the media—emotions such as relief and catharsis to suspects.

Fifth, the confession must appear to be voluntary. The investigators' goal is to portray the suspect as a rational agent who is freely choosing to participate in the interrogation and make statements to police with foreknowledge of his legal rights and options. By eliciting statements from the suspect that his confession is voluntary, interrogators divert attention from situational causes of the suspect's confession—psychologically manipulative and potentially coercive interrogation methods. Toward the end of the post-admission phase, well after the suspect's resistance has been broken down, detectives will often ask the suspect whether they made any threats or promises or if the interrogators in any way coerced him. As all seasoned interrogators know, the suspect will invariably answer "No" to this question—even if he later reports (or the record clearly indicates) that he was threatened or offered promises of leniency in exchange for his confession.

Detectives' public statements about what occurred during their (mostly unrecorded) interrogations usually follow a familiar pattern. On the prosecutor's direct examination, detectives will use the word "interview" rather than "interrogation," and will claim that their goal was merely to obtain the truth. Sometimes they will add that they made no assumptions about the suspect's guilt or innocence. Police trainers admonish interrogators to "avoid creating the impression of an investigator seeking a confession or a conviction; it is far better to fulfill the role of one who is merely seeking the truth" (Inbau et al., 1986; p. 36). On cross-examination, detectives

may acknowledge that they used some interrogation techniques, but will deny any impropriety. They will often attribute the suspect's incriminating statements to the suspect's desire to tell the truth and relieve a guilty conscience. If asked why they failed to audio- or videotape the interrogation, detectives will respond that it is not their department's policy to record or they will declare that it was not necessary: after all, the suspect confessed and signed a written statement. The goal is to shift the focus of the public narrative about interrogation away from their use of psychologically manipulative and/or coercive methods.

Procedural and Policy Solutions to the Problem of False Confessions

Interrogation is a necessary and valuable investigative activity because, even in this modern age of science and technology, oral information remains an important method of solving crimes. Especially for serious crimes such as homicide, police rely disproportionately on confession evidence to clear cases (Gross, 1996). The challenge for a democratic society with an adversarial criminal justice system is to structure the rules of procedure and the restraints on police behavior so that interrogations are conducted fairly and in a way that will produce the most reliable confession evidence possible. Although no single reform is sufficient to achieve these goals, several reforms hold promise for improving the process and outcome of interrogations.

Improve Police Interrogation Training

To reduce the number of false confessions, interrogation training needs to be significantly improved in at least three ways. First, contrary to their current training and practice, interrogators need to be taught that they cannot reliably intuit whether a suspect is innocent or guilty based on their pedestrian hunches about the meaning of a suspect's demeanor or non-verbal behavior. Regrettably, American police interrogators have created a folk psychology of human lie-detection that is based on myth and pseudo-science. Research has repeatedly and unequivocally demonstrated that interrogators' detection-deception training materials are flawed, that their detection-deception judgments are highly prone to error, and—perhaps not surprisingly—that interrogators cannot accurately assess their own lie-detection skills (Granhag and Strömwall, 2004; Memon et al., 2003; Vrij, 2008). Despite this, interrogation training in "behavioral symptom analysis" falsely increases confidence in lie-detection skills, even though it does not increase their ability to discern truth from deception (Vrij et al., 2006). This is significant because all too often—especially in high-profile cases—interrogators wrongly (but ardently) presume a suspect must be guilty merely because of his non-verbal behavior during their initial interaction (rather than because of any reasonable evidence linking the suspect to the crime). Because interrogators falsely assume that the suspect is behaving in ways that are only

indicative of guilt, they subject the innocent suspect to the manipulative methods of modern accusatorial interrogation. Once interrogators elicit a confession, they treat the fact that the suspect has confessed as confirming the initial presumption of his guilt—even if interrogators relied on psychologically coercive methods to extract the confession and even if the statement does not fit the facts of the crime.

Second, detectives need to receive better training about the existence, variety, and causes of police-induced false confessions. Contrary to current practice, interrogation trainers must be taught that their psychological interrogation techniques can and do cause innocent suspects (who are cognitively normal) to confess falsely. Third, interrogators need to receive better training about the indicia of reliable and unreliable statements and how to distinguish between them. It has long been a generally accepted principle in law enforcement (as well as among social scientists and legal scholars) that valid confessions will be supported by logic and evidence, whereas false confessions will not (Leo et al., 2006; Ayling, 1984). In practice, however, detectives virtually always treat a suspect's "I did it" statement as if it is automatically self-validating—even if it fails to be supported by logic or evidence—merely because it validates their own presumption of the suspect's guilt. As Ofshe and Leo (1997) point out, a suspect's "I did it" statement should be treated as a hypothesis to be objectively tested against the case facts (Poole and Lindsay, 1998). As one police manual put it, "An uncorroborated confession is not worth the paper that it is written on" (Oakland Police Department, 1998).

Require Probable Cause to Interrogate

As research on wrongful convictions indicates, police sometimes misclassify innocent persons as guilty based on flimsy evidence (Ofshe and Leo, 1997), such as the suspect did not behave the "right way" when first approached, the suspect is related to or knows the victim, or the suspect fits a general profile (Davis and Leo, 2006). Once police classify a person as a suspect, they presume his guilt and thus trigger confirmation bias—their goal is no longer to investigate the possibility that the suspect may be involved in the crime but instead to pressure and manipulate him into making or agreeing to incriminating statements that confirm their pre-existing belief in his guilt. Therefore, as many have pointed out, one of the best ways to prevent police from eliciting false confessions is to limit interrogation to suspects for whom there exists probable cause of guilt (Davis and Leo, 2006). By subjecting the basis for the police decision to interrogate to the independent review of a third party, a probable cause requirement could effectively prevent fishing expeditions and ill-conceived interrogations, thus screening out the kinds of cases that tend to lead to false confessions from the innocent. Of course, the only way a formal probable cause requirement will be effective is if police are required to seek probable cause whenever they interrogate, and if the penalty for police violation of the probable cause requirement was exclusion of any statements, admissions, or confessions at trial.

Videotaping of Interrogations

Electronic recording is the single most important policy reform available. Unlike some potential reforms, the recording of police interrogations is not a zero sum solution: it benefits all parties who value accurate fact-finding and informed decision-making. Moreover, without a factual record of the entire interrogation, there is no way to meaningfully ensure that the three goals of the adversary system (protection of legal rights, check on state power, and promotion of truth-finding) are being achieved. Police interrogation in the American adversary system is inherently problematic because it often involves behaviors—psychological manipulation, trickery, and deceit—that are regarded as unethical in virtually all other social contexts. Perhaps it should not be surprising, then, that interrogation remains one of the most secretive of all police activities. Electronic recording opens secretive police practices up to the possibility of external scrutiny.

Although ten states (Alaska, Minnesota, Illinois, Maine, New Mexico, Wisconsin, New Jersey, North Carolina, Maryland and Iowa) and the District of Columbia now require police to record interrogations in their entirety in some or all criminal cases, most police departments—as well as the F.B.I.—still do not record interrogations or only selectively tape the admission, not the interrogation that produced it (Slobogin, 2003; Donovan and Rhodes, 2000; Drizin and Reich, 2004).

To be fully effective, electronic recording laws should include four requirements: 1) the recording should include both audio and visual channels; 2) the entire interrogation should be recorded; 3) both the suspect and the interrogator should be visible in the recording; and 4) the recording should carry a time/date stamp. It is important for lawyers, judges, jurors, and experts to be able to evaluate not only what was said and how it was said (verbal and vocal cues), but also the rich variety of visual cues, including facial expressions, gestures, posture, personal space, and gaze. Much of the emotional meaning of an interaction is carried by nonverbal cues (Archer et al., 2001). Intimidation, fear, disorientation, exhaustion, and threat are easier to see than to hear. The second requirement—that the entire interrogation be recorded—ensures that we do not rely on partial, selective, or biased records of what transpired. Too often, police videotape only the final, damning confession. That confession may have been preceded by hours of grueling interrogation. To evaluate whether a confession is coerced or contaminated, experts need to be able to examine the interaction that culminated in a confession. The third requirement—that both the suspect and the interrogator be visible—follows directly from research by Dan Lassiter and his colleagues (Lassiter and Geers, 2004; Lassiter and Irvine, 1986). These researchers found that only video recordings that include both the interrogator and the suspect capture the totality of the encounter and minimize biased interpretations of the interrogation. The fourth requirement—including a time/date stamp on the recording—ensures that there is an accurate accounting of not only the length of the interrogation, but the length of breaks or gaps in the interrogation process.

Pre-Trial Confession Reliability Hearings

Challenges to the admissibility of confession evidence can be conducted pre trial. As in the case of voluntariness hearings, challenges to the reliability of confession evidence should commence upon filing a motion to exclude by the defense. The motion can be styled as a motion *in limine* under local rules of evidence that track Federal Rules of Evidence. Because juries often see confession evidence as dispositive of guilt, even when it is false, its prejudicial effect can be devastating to an innocent defendant (Leo, 1998). This is why Rule 402 allows judges to exclude unreliable evidence on the grounds that its probative value is outweighed by its prejudicial effect. Because the jury is the ultimate fact finder with respect to the truth or falsity of a confession, the standard for admissibility should be less than "beyond a reasonable doubt." Some scholars have proposed a "preponderance of the evidence" standard (Leo et al., 2006). A hearing on the issue of confession reliability should proceed only after any attempt to exclude a confession on involuntariness grounds. The reason for this is simple: the truth or falsity of the confession is not relevant to a determination of the voluntariness of the confession. If judges were to conduct reliability assessments first, their comparison of the contents of the confession with the corroborating evidence could color their assessment of voluntariness.

Prohibit Inherently Coercive Interrogation Techniques

It is striking that the most fundamental policy solution of all—the reform of interrogation methods—is so low on just about everyone's list of proposed solutions to the problem of false confessions. Indeed, since the disclosure of the existence of military interrogations at Abu Ghraib, Guantánamo Bay, and secret "black sites," that include the use of torture, the boundaries for which interrogation practices are considered "coercive" may have shifted (Costanzo et al., 2006). Compared to extreme techniques such as stress positions, hypothermia, and water-boarding, police interrogations of criminal suspects may seem tame. In theory, civilian criminal courts are supposed to exclude confessions that are the product of coercive or otherwise impermissible police methods. These methods include deprivations (e.g., food, sleep, water, or access to bathroom facilities), unreasonably lengthy interrogations, inducing exhaustion or extreme fatigue, and explicit threats of physical violence. In the modern era, however, promises of leniency (in exchange for confession) and threats of differential punishment (in the absence of confession)—often communicated implicitly or indirectly—are the primary causes of police-induced false confession (Leo, 2007). Appellate courts need to create an unambiguous bright-line rule explicitly prohibiting, under all circumstances, any implicit or explicit promises of leniency in exchange for an admission, as well as any implicit or explicit threat or suggestion of harm in the absence of an admission.

Appellate courts and legislatures may also wish to revisit the issue of whether (and if so to what extent) deceptive, false, dishonest, and/or misleading interrogation

techniques should be legally impermissible. Currently, police interrogation is thoroughly and intentionally deceptive: police routinely lie to and mislead suspects about their role in the interrogation, the facts of the case, and the (real or alleged) evidence against the suspect. Police also routinely lie to and manipulate suspects during the post-admission portion of the interrogation. Many scholars have argued that police lying about evidence (false evidence ploys) may heighten the risk of eliciting false or unreliable confessions (White, 1997; Kassin and Gudjonsson, 2004; Gohara, 2006).

Time Limits on Interrogations

Courts and legislatures should specify objective time limits for interrogations. Lengthy incommunicado interrogation is not only inherently unfair, but, as recent research has documented, far more common in false confession cases. On average, routine interrogations almost always last less than two hours (Leo, 1996; Cassell and Hayman, 1996), yet interrogations leading to false confessions often last longer than six hours (Drizin and Leo, 2004). Longer interrogations appear to increase the risk of false confessions by fatiguing suspects and thus impairing their ability and motivation to resist police pressures. Exhaustion, as Davis and O'Donohue (2004) have pointed out, "may lead to greater interrogative suggestibility via deficits in speed of thinking, concentration, motivation, confidence, ability to control attention, and ability to ignore irrelevant or misleading information" (p. 957). Specifying a time limit on interrogations of no more than four hours should reduce the risk of eliciting false confessions from the innocent without undermining the ability of police to elicit true confessions from the guilty (Costanzo and Leo, 2007). For as Inbau et al. (2001) point out, "rarely will a competent interrogator require more than approximately four hours to obtain a confession from an offender, even in cases of a very serious nature … Most cases require considerably fewer than four hours" (p. 597).

The "Appropriate Adult" Safeguard for Vulnerable Suspects

Because the single biggest suspect risk factor in false confession is youth, juveniles should be provided with an appropriate adult during questioning. This adult might be a lawyer or other person (independent of police) who is specially trained to fill this role. Mentally impaired suspects should also receive special treatment by interrogators. Some jurisdictions have already moved toward adopting such practices. In the wake of several false confession cases involving children and teenagers in Chicago, Illinois enacted a law requiring that all children under age 13 be provided access to attorneys before their interrogation in murder and sex offense cases. The Broward County, Florida Sheriff's Office—whose investigators had elicited notable false confessions from developmentally disabled suspects in several high-profile cases (Drizin and Leo, 2004)—requires that before questioning a developmentally

disabled suspect, detectives must notify their supervisors and make a reasonable effort to have an appropriate adult present during all questioning.

Expert Testimony on Interrogations and Confessions

The use of social science expert testimony in cases involving a disputed interrogation or confession has become increasingly common (Leo, 2001b; White, 2003; Kassin and Gudjonsson, 2004). There is now a substantial and well-accepted body of scientific research on this topic, and case law supports the admissibility of such expert testimony (Costanzo and Leo, 2007). Although there have been a few cases in which courts have not permitted expert testimony, such cases are exceptional; social psychologists have testified in hundreds of criminal and civil trials (Fulero, 2004).

If a factually disputed confession is introduced at trial, the jury will want to know how an innocent person could possibly have been made to confess falsely, especially to a heinous crime. The purpose of expert testimony at trial is to provide an overview of research on interrogation and confessions to assist the jury in making a fully informed decision about what weight to place on the defendant's confession. More specifically, social science expert witnesses can aid the jury by: 1) discussing research documenting the phenomenon of police-induced false confessions; 2) explaining how and why particular interrogation methods and strategies can cause the innocent to confess; 3) identifying the conditions that increase the risk of false confession; and 4) explaining the generally accepted principles of PAN analysis. By educating the jury about the existence, psychology, causes, and indicia of police-induced false confession, social science expert witness testimony at trial should reduce the number of confession-based wrongful convictions.

Jury Instructions

Although the effect of jury instructions is generally weak (see Chapter 6, in Volume I), cautionary instructions might help increase juror sensitivity about confession evidence and potentially reduce the risk of wrongful convictions. Two state Supreme Courts have focused on instructions as a way to reform interrogation practices and guarantee the use of more reliable confession evidence and accurate jury verdicts. In *Commonwealth v. DiGiambattista* (2004) the Supreme Judicial Court of Massachusetts ruled that any confession resulting from an unrecorded interrogation will entitle the defendant upon request to a jury instruction: "… cautioning the jury that, because of the absence of any recording in the case before them, they should weigh evidence of the defendant's alleged statement with great caution and care" (p. 534). More recently, a special committee appointed by the New Jersey Supreme Court, recommended that defendants in New Jersey be entitled to a cautionary instruction in the event that New Jersey Police fail to electronically record custodial interrogations in their entirety (Report of the Special Committee on Recordation of Custodial Interrogations, 2005). Interestingly, both Massachusetts and New Jersey

have recommended jury instructions not so much to help the jury evaluate specific evidence but as a way to motivate police and prosecutors to introduce only the most reliable form of evidence (i.e., electronically recorded interrogations and confessions) by punishing them with a cautionary instruction if they fail to do so (New Jersey subsequently enacted an electronic recording requirement for interrogations). One can, however, imagine more traditional jury instructions. Such instructions might, for example, educate a jury about the risks of particular interrogation techniques (e.g., promises of leniency, false evidence ploys, threats) and the principles of PAN analysis or the importance of external corroboration.

Conclusion

As we noted earlier, there are at least three tragedies that follow from each wrongful conviction: 1) an innocent person must suffer the pains of prosecution, conviction, and imprisonment; 2) the true perpetrator remains free to commit additional crimes; and 3) decision-makers in the legal system—police officers, prosecutors, judges, and jurors—must live with the knowledge that they were responsible for the first two tragedies. Research on false confessions—and research on wrongful convictions more generally—has the potential to reduce dramatically the number of such tragedies. It is only through careful analysis of how mistakes are made that we are able to refine procedures and practices to prevent mistakes. Police, lawyers, judges, jurors, and citizens share a common interest: to correctly identify and convict those who are guilty, and to allow the innocent to go free.

References

Amnesty International (2006, April 1). *Russian Federation: Preliminary Briefing to the UN Committee Against Torture*. Retrieved September 20, 2006, from http://web.amnesty.org/library.

Amsterdam, A., and Bruner, J. (2000). *Minding the Law*. Cambridge, MA: Harvard University Press.

Archer, D., Costanzo, M., and Akert, R. (2001). The Interpersonal Perception Task (IPT): Alternative approaches to problems of theory and design. In J. Hall and F. Bernieri (eds) *Interpersonal Sensitivity: Theory and Measurement*. Mahwah, NJ: Erlbaum. pp. 161–82.

Ayling, C. (1984). Corroborating confessions: An empirical analysis of legal safeguards against false confessions. *Wisconsin Law Review*, 1121–204.

Bedau, H., and Radelet, M. (1987). Miscarriages of justice in potentially capital cases. *Stanford Law Review*, 40, 1–79.

Berkemer v. McCarty, 468 U.S. 420 (1984).

Blagrove, M. (1996). Effects of length of sleep deprivation on interrogative suggestibility. *Journal of Experimental Psychology: Applied*, 2, 48–59.

Cassell, P.G., and Hayman, B.S. (1996). Police interrogation in the 1990s: An empirical study of the effects of *Miranda*. *UCLA Law Review*, 43, 839–931.

Cloud, M., Shepherd, G.B., Barkoff, A.N., and Shur, J.V. (2002). Words without meaning: The constitution, confessions and mentally retarded suspects. *University of Chicago Law Review*, 69, 495–624.

Cole, S. (2001). *Suspect Identities: A History of Fingerprinting and Criminal Identification*. Cambridge, MA: Harvard University Press.

Commonwealth v. DiGiambattista, 442 Mass. 423 (2004).

Conroy, J. (2000). *Unspeakable Acts, Ordinary People: The Dynamics of Torture*. New York: Knopf.

Costanzo, M., Gerrity, E., and Lykes, M.B. (2006). Psychologists and the use of torture in interrogations. *Analyses of Social Issues and Public Policy*, 6, 1–14.

Costanzo, M., and Leo, R. (2007). Research findings and expert testimony on police interrogations and confessions to crimes. In M. Costanzo, D. Krauss, and K. Pezdek (eds). *Expert Psychological Testimony for the Courts*. Mahwah, NJ: Erlbaum. pp. 69–98.

Costanzo, M., and Peterson, J. (1994). Attorney persuasion in the capital penalty phase: A content analysis of closing arguments. *Journal of Social Issues*, 50, 125–47.

CPIHR – Crowley Program in International Human Rights. (2000). *Presumed Guilty? Criminal Justice and Human Rights in Mexico*. Retrieved September 22, 2006, from http://www.crowleyprogram.org/mexicoreport.htm.

Davis, D., and Leo, R. (2006). Strategies for preventing false confessions and their consequences. In M. Kebbell and G. Davies (eds), *Practical Psychology for Forensic Investigations and Prosecutions*. New York: Wiley. pp. 121–49.

Davis, D., and O'Donohue, W. (2004). The road to perdition: Extreme influence tactics in the interrogation room. In W. O'Donahue, and E. Levensky (eds), *Handbook of Forensic Psychology*. San Diego, CA: Elsevier Academic Press. pp. 897–996.

Donovan, D., and Rhodes, J. (2000). Comes a time: The case for recording interrogations. *Montana Law Review*, 61, 223–49.

Drizin, S.A., and Leo, R.A. (2004). The problem of false confessions in the post-DNA world. *North Carolina Law Review*, 82, 891–1007.

Drizin, S.A., and Reich, M.J. (2004). Heeding the lessons of history: The need for mandatory recording of police interrogations to accurately assess the reliability and voluntariness of confessions. *Drake Law Review*, 52, 619–46.

Einesman, F. (1999). Confessions and culture: The interaction of Miranda and diversity. *The Journal of Criminal Law and Criminology*, 90, 1–48.

Ellis, J., and Luckasson, R. (1985). Mentally retarded defendants. *The George Washington Law Review*, 53, 414–93.

Foote, D.H. (1992). The benevolent paternalism of Japanese criminal justice. *California Law Review*, 80, 317–90.

Fulero, S. (2004). Expert psychological testimony on the psychology of interrogations and confessions. In G.D. Lassiter (ed.), *Interrogations, Confessions, and Entrapment*. New York: Kluwer. pp. 247–63.

Garrett, B. (2008). Judging innocence. *Columbia Law Review*, 108, 55–142.

Gohara, M. (2006). A lie for a lie: False confessions and the case for reconsidering the legality of deceptive interrogation techniques. *Fordham Urban Law Journal*, 33, 791–842.

Granhag, P.A., and Strömwall, L.A. (eds) (2004). *The Detection of Deception in Forensic Contexts*. Cambridge: Cambridge University Press.

Gross, S. (1996). The risks of death: Why erroneous convictions are common in capital cases. *Buffalo Law Review*, 44, 469–500.

Gross, S.R., Jacoby, K., Matheson, D.J., Montgomery, N., and Patel, S. (2005). Exonerations in the United States, 1989 through 2003. *Journal of Criminal Law and Criminology*, 95, 523–53.

Gudjonsson, G.H. (2003). *The Psychology of Interrogations and Confessions: A Handbook*. Chichester: Wiley.

Hofstede, G. (1980). *Culture's Consequences: International Differences in Work-related Values*. Beverly Hills, CA: Sage.

Home Office (2005). *Police and Criminal Evidence Act 1984. Codes of Practice* (rev. ed.) London: Stationery Office.

Human Rights Watch (2006). *Mexico: Lost in Translation*. New York: Human Rights Watch.

Huntley, J., and Costanzo, M. (2003). Sexual harassment stories: Testing a story-mediated model of juror decision-making in civil litigation. *Law and Human Behavior*, 27, 29–51.

Inbau, F., Reid, J.E., and Buckley, J.P. (1986). *Criminal Interrogation and Confessions* (3rd ed.). Baltimore, MD: Williams and Wilkins.

Inbau, F., Reid, J.E., Buckley, J.P., and Jayne, B.C. (2001). *Criminal Interrogation and Confessions* (4th ed.). Gaithersburg, MD: Aspen.

Innocence Project (2006). www.innocenceproject.org.

Johnson, D. (2002). *The Japanese Way of Justice: Prosecuting Crime in Japan*. Oxford: Oxford University Press.

Kassin, S. (2006). A critical appraisal of modern police interrogations. In T. Williamson (ed.), *Investigative Interviewing: Rights, Research, Regulation*. Portland, OR: Willan.

Kassin, S., and Gudjonsson, G.H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest*, 5, 33–67.

Kassin, S., and McNall, K. (1991). Police interrogations and confessions: Communicating promises and threats by pragmatic implication. *Law and Human Behavior*, 15, 233–51.

Katz, J. (1999). *How Emotions Work*. Chicago, IL: University of Chicago Press.

Kemmelmeier, M., Burnstein, E., Krumov, K., Genkova, P., Kanagawa, C., Hirshberg, M.S., et al. (2003). Individualism, collectivism, and authoritarianism in seven societies. *Journal of Cross-Cultural Psychology*, 34(3), 304–22.

Larsen, L.J. (2004). The foreign-born population in the United States: 2003. *Current Population Reports*. Washington, DC: U.S. Census Bureau.

Lassiter, G.D., and Geers, A.L. (2004). Evaluation of confession evidence: Effects of presentation format. In G.D. Lassiter (ed.), *Interrogations, Confessions, and Entrapment*. New York: Kluwer Press. pp. 197–214.

Lassiter, G.D., and Irvine, A.A. (1986). Videotaped confessions: The impact of a point of view on judgments of coercion. *Journal of Applied Social Psychology*, 16, 268–76.

Le v. State, 947 P.2d 535 (1997).

Leo, R.A. (1996). Inside the interrogation room. *Journal of Criminal Law and Criminology*, 86, 266–303.

Leo, R.A. (1998). Miranda and the problem of false confessions. In R.A. Leo and G.C. Thomas, III (eds), *The Miranda Debate: Law, Justice and Policing*. Boston, MA: Northeastern University Press. pp. 271–82.

Leo, R.A. (2001a). Questioning the relevance of *Miranda* in the twenty-first century. *The Michigan Law Review*, 99, 1000–1029.

Leo, R.A. (2001b). False confessions: Causes, consequences, and solutions. In S.D. Westervelt and J.A. Humphrey (eds), *Wrongly Convicted: Perspectives on Failed Justice*. Newark, NJ: Rutgers University Press. pp. 36–54.

Leo, R.A. (2002). *Miranda*, confessions and justice: Lessons for Japan? In M. Feeley and S. Miyazawa (eds), *The Japanese Adversary System in Context: Controversies and Comparisons*. London: Palgrave. pp. 200–219.

Leo, R.A. (2008). *Police Interrogation and American Justice*. Cambridge, MA: Harvard University Press.

Leo, R.A., Drizin, S., Neufeld, P., Hall, B., and Vatner, A. (2006). Bringing reliability back in: False confessions and legal safeguards in the twenty-first century. *Wisconsin Law Review*, 2, 479–539.

Leo, R.A., and Ofshe, R.J. (1998a). The consequences of false confessions: Deprivations of liberty and miscarriages of justice in the age of psychological interrogation. *The Journal of Criminal Law and Criminology*, 88, 429–96.

Leo, R.A., and Ofshe, R.J. (1998b). Using the innocent to scapegoat *Miranda*: Another reply to Paul Cassell. *The Journal of Criminal Law and Criminology*, 88, 557–77.

Leo, R.A., and White, W.S. (1999). Adapting to *Miranda*: Modern interrogators' strategies for dealing with the obstacles posed by *Miranda*. *Minnesota Law Review*, 84,397–472.

Liu v. State, 628 A.2d 1376 (1993).

Los Angeles Police Department (undated). Interrogation training manual.

Lu, H., and Miethe, T.D. (2003). Confessions and criminal case disposition in China. *Law and Society Review*, 37(3), 549–78.

Matsumoto, D. (1997). *Culture and Modern Life*. Pacific Grove, CA: Brooks/Cole.

Meissner, C.A., and Kassin, S.M. (2004). "You're guilty, so just confess!" Cognitive and confirmational biases in the interrogation room. In G.D. Lassiter (ed.), *Interrogations, Confessions, and Entrapment*. New York: Kluwer Academic. pp. 85–106.

Memon, A., Vrij, A., and Bull, R. (2003). *Psychology and Law: Truthfulness, Accuracy and Credibility*. London: Jossey Bass.

Miranda v. Arizona, 384 U.S. 486 (1966).

Moston, S., Stephenson, G.M., and Williamson, T.M. (1992). The effects of case characteristics on suspect behavior during questioning. *British Journal of Criminology*, 32, 23–40.

Napier, M. and Adams, S. (2002). Criminal confessions: Applying interview and interrogation techniques can help law enforcement identify suspects and obtain and safeguard confessions. *FBI Law Enforcement Bulletin*, 71, 9–15.

North Carolina v. Butler, 441 U.S. 369 (1979).

Oakland Police Department (1998). Police interrogation training manual. Unpublished.

Ofshe, R.J., and Leo, R.A. (1997). The decision to confess falsely: Rational choice and irrational action. *Denver University Law Review*, 74, 979–1122.

Olsen-Fulero, L., and Fulero, S. (1997). Commonsense rape judgments: An empathy-complexity theory of rape juror story making. *Psychology, Public Policy, and Law*, 3, 402–27.

Pennington, N., and Hastie, R. (1993). The story model for juror decision-making. In R. Hastie (ed.), *Inside the Juror*. Cambridge: Cambridge University Press. pp. 84–115.

People v. Merrero, 459 N.E.2d 1158 (1984).

Poole, D.A., and Lindsay, D.S. (1998). Assessing the accuracy of young children's reports: Lessons from the investigation of child sexual abuse. *Applied and Preventive Psychology*, 7, 1–26.

Ramseyer, J.M., and Rasmusen, E.B. (2001). Why is the Japanese conviction rate so high? *Journal of Legal Studies*, 30(1), 53–88.

Redlich, A.D. (2004). Mental illness, police interrogations, and the potential for false confession. *Law and Psychiatry*, 55, 19–21.

Report of the Special Committee on Recordation of Custodial Interrogations (2005). Special Committee Appointed by the New Jersey Supreme Court. Unpublished Report.

Russ (2005, April 14). Chinese court bars evidence obtained through torture. *Jurist Legal News and Research*. Retrieved from http://jurist.law.pitt.edu/paperchase/indexarch.php.

Scheck, B., Neufeld, P., and Dwyer, J. (2000). *Actual Innocence*. New York: Doubleday Books.

Senese, L. (2005). *Anatomy of Interrogation Themes: The Reid Technique of Interviewing and Interrogation*. Chicago, IL: John E. Reid and Associates.

Slobogin, C. (2003). Towards taping. *Ohio State Journal of Criminal Law*, 1, 309–22.

Sterling, J. (2005, May 17). China takes another step against torture. *Jurist Legal News and Research*. Retrieved from http://jurist.law.pitt.edu/paperchase/indexarch.php.

Sullivan, T.P. (2004). The police experience: Recording custodial interrogations. *The Champion*, 28, 24–7, 52.

Triandis, H.C. (1994). *Culture and Social Behavior*. New York: McGraw-Hill.

U.S. v. Zapata, 997 f.2d 751 (1993).

U.S. Department of State (2005a, March 8). *Japan: Country Reports on Human Rights Practices*. Retrieved September 20, 2006, from http://www.state.gov/g/drl/rls/hrrpt/2005/61610.htm.

U.S. Department of State (2005b, March 8). *Mexico: Country Reports on Human Rights Practices*. Retrieved September 20, 2006, from http://www.state.gov/g/drl/rls/hrrpt/2005/62736.htm.

Vrij, A. (2008). *Detecting Lies and Deceit: Pitfalls and Opportunities* (2nd ed.). London: Wiley.

Vrij, A., Mann, S., and Fisher, R.P. (2006). Information-gathering vs accusatory interview style: Individual differences in experiences. *Personality and Individual Differences*, 41, 589–99.

Warden, R. (2003). The role of false confessions in Illinois wrongful murder convictions since 1970. Center on Wrongful Convictions Research Report. www.law.northwestern.edu/depts/clinic/ (last accessed on July 15, 2005).

Weisselberg, C.D. (1998). Saving *Miranda*. *Cornell Law Review*, 84, 109–92.

Westwood, R.I., Tang, S.F.Y., and Kirkbride, P.S. (1992). Chinese conflict behavior: Cultural antecedents and behavioral consequences. *Organization Development Journal*, 10, 13–19.

White, W. (1997). False confessions and the constitution: safeguards against untrustworthy evidence. *Harvard Civil Rights-Civil Liberties Law Review*, 32, 105–57.

White, W. (2003). Confessions in capital cases. *University of Illinois Law Review*, 2003, 979–1036.

This page has been left blank intentionally

Chapter 3
System and Estimator Variables in Eyewitness Identification: A Review

Solomon M. Fulero

Synopsis

Eyewitness reliability and the factors that affect it have been the subject of study within the field of psychology for over a hundred years. This chapter reviews the research on the estimator and system variables that have been identified and studied. Estimator variables (characteristics of the witness, characteristics of the event, and characteristics of the testimony) are reviewed first. Next, system variables (the procedures used by the police and other members of the criminal justice system) are reviewed, with an eye toward improving the collection of eyewitness evidence by law enforcement officers. Finally, the increasing use of expert testimony in cases involving eyewitness evidence is discussed, along with a review of judge's decisions on admissibility of such expert testimony.

System and Estimator Variables in Eyewitness Identification: A Review

> Eyewitness identification of a person whom the witness had never seen before the crime or other incident presents a substantial risk of misidentification and increases the chance of a conviction of an innocent defendant … [S]tudies by psychologists and legal researchers … have confirmed that eyewitness testimony is often hopelessly unreliable … Thus, whenever eyewitness evidence is introduced against an accused we require the utmost protection against mistaken identification. (*Commonwealth v. Vardinski*, 2003)

The conclusion that eyewitness evidence is crucial in the outcome of cases is supported by the work of Lavrakas and Bickman (1975). These researchers surveyed 54 prosecutors regarding their opinions of "what makes a good witness." The prosecutors were asked to consider what effect a set of witness attributes would have on the outcome of a case. Ratings were made on a five-

point scale, from "this attribute is totally unrelated to the outcome" to "this attribute is very related to the outcome." Results showed that witness attributes such as race, sex, age, or socioeconomic status made virtually no difference in the prosecutor's ratings of importance. On the other hand, the victim's availability for testimony, the victim's ability to testify, and the witness's assertion of a "good memory" and clarity of recall were central to the prosecutors' ratings. Clearly, the presence of "good" and available eyewitness evidence is seen as an important determinant of case outcome.

But the importance of the eyewitness's memory in reconstructing events from the past does not end with the arrest of a suspect. At a trial, the testimony of an eyewitness who incriminates the defendant is—along with the presence of a confession—usually the most influential evidence (Lavrakas and Bickman, 1975). Alibis, circumstantial evidence, even masses of physical evidence favoring the defendant's innocence, can wither away in light of an eyewitness's courtroom identification.

Of course, the essential problem is that eyewitnesses are not infallible. We know this from a variety of sources, primarily from studies of cases of known wrongful convictions (Rattner, 1988; Huff et al., 1996). But the most important such studies have emerged from the recent availability of DNA technology to analyze claims of wrongful conviction. Wells (1993) concluded that eyewitness errors provide the single most frequent cause of wrongful convictions, and two examinations of such cases provide strong evidence for that assertion. In 1996, the United States Department of Justice published an analysis of the first 28 cases of persons in the United States who were convicted of crimes but later exonerated on the basis of DNA testing (Connors et al., 1996). Of those, 24 involved mistaken eyewitness identification, some with multiple witnesses (as many as five in one case). A later analysis found that in the first 40 of these cases, 36 (or 90 percent) were cases in which one or more eyewitnesses falsely identified the innocent person (Wells et al., 1998). Another review, extending the number of DNA cases to over 100, found that eyewitness error was involved in 84 percent of the cases of wrongful conviction (Scheck et al., 2000; see also Garrett, 2007). The number of such DNA cases as of March 1, 2009 is 232 (see www.innocenceproject.org for the current count).

In the last 30 years there has been an explosion of research on the topic of eyewitness identification, the factors that affect it, and the process of eyewitness evidence collection (see Cutler and Penrod, 1995; Wells et al., 1998). Psychologists now possess extensive information on the factors that affect reliability and also how eyewitness evidence collection can be improved in actual cases (Wells, 1993; Technical Working Group on Eyewitness Evidence, 1999; 2003; Wells et al., 2006).

Ideally, an eyewitness's identification will be a product solely of his or her memory rather than a product of the identification procedures used by the police (Technical Working Group on Eyewitness Evidence, 1999; 2003). Studies in the psychological laboratory or controlled field studies that simulate a crime and then determine the degree of accuracy of eyewitnesses confirm the fear that false identifications by bystanders occur with unfortunate frequency (Brigham et al., 1982; Buckhout, 1974; Cutler et al., 1987; Ellis et al., 1980; Leippe et al., 1978; Wells, 1984a; Wells,

Lindsay, and Ferguson, 1979). In those crime simulations in which subjects believed the crime was real and their identification would have consequences for the accused, high rates of false identification still occurred (Malpass and Devine, 1980; Murray and Wells, 1982). Studies that have looked at actual eyewitnesses in actual crimes (after the fact, of course, since crimes cannot ethically be created by researchers) have generally found similar results (see Behrman and Davey, 2001). How high a rate of inaccuracy? In some studies, as many as 90 percent of responses were false identifications; in others, only a few subjects erred. The extreme variation exemplifies the notion that the degree of accuracy can be partly determined by the specific procedures used by the police to collect eyewitness evidence during a criminal investigation.

Rather than being satisfied simply to point out that the reports of eyewitnesses are often inaccurate, we should recognize that the degree of accuracy is often influenced by the procedures used by the police and other members of the criminal justice system (Wells and Seelau, 1995). Wells (1978) referred to these as *system variables*. These include the type of questioning done by the police, the nature of the lineup or photoarrays, and the presence or absence of videotaping of procedures. These variables are *preventable errors* (Wells, 1993); and in fact, psychologists could aid in the construction of lineups and the development of interviewing procedures that reduce inaccuracy.

Estimator Variables

The other determinants of an eyewitness's accuracy—what Wells called *estimator variables*—are not controllable by the criminal justice system. These will be discussed first. Estimator variables can be sorted into three broad categories: characteristics of the witness, characteristics of the event, and characteristics of the testimony.

Characteristics of the Witness

Are members of certain groups better eyewitnesses than those of others? The empirical evidence is not overwhelming (see Wells and Olson, 2003). For example, there is no clear evidence that males and females differ significantly overall in ability to identify people from lineups. A meta-analysis by Shapiro and Penrod (1986) indicated that females might be slightly more likely to make accurate identifications but also slightly more likely to make mistaken identifications than are males (due to females being more likely to attempt an identification), thereby yielding an overall equivalent diagnosticity for males and females.

The age of the eyewitness, on the other hand, has been consistently linked to eyewitness identification performance. Young children and the elderly consistently perform significantly worse than younger adults (for children, see Lindsay et al., 1997; Pozzulo and Lindsay, 1998; as to older adults, see Brimacombe et al., 2003).

The race of the eyewitness has been examined extensively. The evidence is now quite clear that people are better able to recognize faces of their own race or ethnic group than faces of another race or ethnic group (see Meissner and Brigham, 2001; Wells and Olson, 2003). This effect is not confined to the United States, nor solely to African-Americans and Whites; the effect is also found for Asians and for Hispanics (see e.g., Luce, 1974).

Characteristics of the Event

A variety of factors affect the ability of an eyewitness to identify the culprit at a later time, including the amount of time the culprit is in view, the lighting conditions, whether the culprit wears a disguise, the presence or absence of a weapon, and the timing of knowledge that one is witnessing a crime.

Simple disguises such as covering the hair or wearing sunglasses, can result in significant impairment of eyewitness identification (Cutler et al., 1987; Hockley et al., 1999; Wells and Olson, 2003). The degree of impairment may be reduced by having the target look similar (i.e., wear sunglasses) at the time of the recognition test (Hockley et al., 1999).

Photos of criminal suspects used in police lineups are sometimes several years old. Changes in appearance that occur naturally over time and changes that are made intentionally by suspects can have quite strong effects on recognition (see Wells and Olson, 2003). Read et al. (1990) found that photos of the same people taken two years apart were less likely to be recognized as the same people when their appearance had naturally changed than when their appearance had remained largely the same.

Clearly, at very low light levels, there is a point at which a face cannot be perceived well enough to be recognized later. There is some general work on the relation between adequate lighting and distance (e.g., Wagenaar and Schrier, 1994), and this work is commonly used in stage lighting and other applied contexts.

As would be expected, the amount of time the witness has to look at a culprit's face affects the chances that the eyewitness can identify the person later (Ellis et al., 1977). However, the witness's estimate of time may be impaired by time overestimation (see e.g., Loftus, Schooler, Boone and Kline, 1987).

One factor that can signal to eyewitnesses that a crime is occurring is the presence of a weapon. A number of studies have been directed at the question of the so-called weapon-focus effect. A meta-analysis of these studies indicates that the presence of a weapon reduces the chances that the eyewitness can identify the holder of the weapon (Steblay, 1992). Loftus, Loftus, and Messo (1987) monitored eyewitnesses' eye movements, and found that weapons draw visual attention away from other things such as the culprit's face. Complicating the issue somewhat is the fact that the presence of weapons or other types of threatening stimuli can cause arousal, fear, and emotional stress.

The exact effects of such stress on memory are still being debated (see Wells and Olson, 2003; see also Deffenbacher et al., 2004 for a meta-analysis of the stress effect). Deffenbacher (1983) has suggested that very high and very low levels of arousal will impair memory. Christianson's (1992) review of the evidence relating emotional stress to memory suggests that emotional events receive preferential processing; emotional response causes a narrowing of attention (as suggested by Easterbrook, 1959) with loss of peripheral details. In a very important new study, Morgan et al. (2004) examined the eyewitness capabilities of more than 500 active-duty military personnel enrolled in a survival-school program. After 12 hours of confinement in a mock prisoner-of-war camp, participants experienced both a high-stress interrogation with real physical confrontation and a low-stress interrogation without physical confrontation. Both interrogations were 40 minutes long; they were conducted by different persons. A day after release from the camp, and having recovered from food and sleep deprivation, the participants viewed a 15-person live lineup, a 16-person photo spread, or a sequential presentation of photos of up to 16 persons. Regardless of the testing method, memory accuracy for the high-stress interrogator was much lower overall than for the low-stress interrogator.

Characteristics of Testimony

Considerable interest and research have been directed at the question of whether there are characteristics of an eyewitness's testimony that could be used to postdict whether the witness made an accurate or false identification. The bulk of this research has focused on the certainty (confidence) of the eyewitness.

Early research suggested that the certainty an eyewitness expresses in an identification is largely unrelated to the accuracy of the identification (Wells and Olson, 2003). However, several moderators of the strength of the relation have been identified. One important moderator is the overall accuracy of the eyewitnesses. When accuracy is low (e.g., from poor witnessing conditions), the certainty-accuracy relationship suffers (Bothwell et al., 1987). Later meta-analyses indicate that the certainty-accuracy relation is somewhat stronger if the analysis is restricted to those making an identification (choosers only) than if it also includes witnesses who make correct and false rejections (Sporer et al., 1995). More recent work indicates that directing eyewitnesses to reflect on their encoding and test conditions, or asking them to think about why their identification might have been mistaken, can improve the relation between accuracy and certainty, especially when calculated using calibration methods rather than the traditional correlation (Brewer et al., 2002).

Recent studies also cast doubt on the utility of eyewitness certainty in actual cases (see Wells and Olson, 2003). Eyewitness certainty has been shown to be highly malleable (Wells and Bradfield 1998; 1999; Douglass and McQuiston-Surratt, 2006; Douglass and Steblay, 2006). After making mistaken identifications, some eyewitnesses were given confirming feedback by the lineup administrator ("Good, you identified the suspect") whereas others were given no feedback about

their identification. This feedback served to distort the eyewitnesses' recollections of the certainty they had in their identifications. Those given confirming feedback recalled having been very certain in their identification compared to those given no confirming feedback. This certainty-inflation effect is greater for eyewitnesses who make mistaken identifications than it is for those who make accurate identifications, resulting in a significant loss in the certainty-accuracy relation (Bradfield et al. 2002). In actual cases, it is common for lineup administrators (usually the detective in the case) to give confirming feedback to eyewitnesses, thereby inflating the certainty of the eyewitness and confounding the certainty-accuracy relation. Even if the lineup administrator refrains from giving the witness confirming feedback, the witness is likely to make confirming inferences from later events (e.g., an indictment of the identified person).

Another real-world factor that can muddle the meaning of eyewitness certainty is repeated testing (see Wells and Olson, 2003). Shaw and his colleagues (Shaw 1996; Shaw and McClure 1996) have shown that repeated questioning of eyewitnesses on a matter about which they were inaccurate serves to inflate their certainty that they were accurate.

System Variables

The critical distinction between estimator and system variables is that errors in system variables can often be reduced and can sometimes be prevented. We can do nothing about poor lighting conditions or the brevity of exposure to the criminal, but police can work to eliminate practices that have been shown to lead to further inaccuracies in reports and in evidence collection.

When the police have a suspect, they usually ask any victim or other eyewitness to identify him or her through the use of a lineup (called an "identity parade" in Great Britain) or a photoarray (also called a photospread). A special mention should be made of the procedure called the "showup"—essentially a lineup composed of only one person. Both psychologists and the courts have assumed that showups are inherently more suggestible than lineups that include 4, 5, or 6 foils (*Stovall v. Denno*, 1967, p. 302), though courts commonly allow eyewitness evidence obtained with showups anyway. In fact, experimental psychologists who study the accuracy of memory are quite strong in their belief that the procedure is prejudicial (Malpass and Devine, 1983; Wells, Leippe, and Ostrom, 1979; Yarmey et al., 1996; Yarmey, 1979; though see Gonzalez et al., 1993; Davis and Gonzalez, 1996).

It is easy to see why lineups and photospreads should be a better procedure for law enforcement to use. Used effectively, a lineup will serve two purposes: to determine whether a suspect is in fact the perpetrator observed by the witness, and to assess the reliability of the witness. Picking someone other than the suspect suggests the latter—unreliable witness memory—and discredits the witness rather than the suspect. The lineup witness who selects a foil may rightly be considered

an unreliable source for subsequent identification evidence. On the other hand, the showup witness has no foil options. A witness who rejects the showup retains police trust as a reliable witness, even in the case where the witness incorrectly says it is not the perpetrator. Therefore, if foil choices are considered useful indications that witnesses are willing to identify innocent people, lineups and photospreads may have an important evidentiary advantage—one that actually transcends the rates of correct identification or errors in the two procedures (see Steblay et al., 2003). It is also worth noting that in today's modern world, one could imagine a time when a photolineup could be generated by a police officer in his or her car in minutes, using a digital camera and photos obtained over the police car's computer, effectively rendering the showup technique obsolete (Wrightsman and Fulero, 2005).

Certainly, the evidence collection procedures used by the police in eyewitness cases can have an effect on witnesses' reactions. The most frequent kinds of errors that are seen in such cases are (see Wrightsman and Fulero, 2005):

1. Implying that the criminal is definitely one of the stimulus persons.
2. Pressuring the witness to make a choice (i.e., creating a demand characteristic).
3. Asking the eyewitness specifically about the suspect while not asking those same questions about the foils (or what Wells and Seelau [1995] call a "confirmation bias").
4. Encouraging a loose recognition threshold in the eyewitness by asking the witness if there is "anyone familiar," or "anyone who looks like the person."
5. Leaking the police officer's hunch, by making it obvious to the eyewitness which is the suspect (Wells and Seelau, 1995, pp. 767–8).
6. After an eyewitness's selection, telling the eyewitness that his or her choice is the "right" one.

Studies have shown that the confidence level of witnesses' reports, as well as their memories of the circumstances of their view of the event, can be manipulated by giving them feedback that their choice is correct, such as by telling them that another witness identified the same person (Luus and Wells, 1994; Luus, 1991; Semmler et al., forthcoming; Wells and Bradfield, 1998; 1999; Bradfield et al., 2002; Wells et al., 2003).

It is clear that the procedures used by some police have the potential of increasing the rate of false identifications (Loftus, 1993). Wells and his colleagues (Wells and Seelau, 1995; Wells et al., 1998) have suggested that the application of four straightforward rules can reduce such errors, rules which have now become part of the material found in the Eyewitness Evidence Guide for Law Enforcement (Technical Working Group on Eyewitness Evidence, 1999; 2003):

Rule 1: "The person who conducts the lineup or photospread should not be aware of which member of the lineup or photospread is the suspect" (Wells et al., 1998, p. 627).

Customarily, the detective who has handled the case administers the lineup. The problem is that this officer, knowing who the suspect is, may communicate this knowledge, even without intending to do so. A variation in eye contact with the witness, a subtle shift in body position or facial expression, or tone of voice, may be enough to communicate feedback to the witness, who often is unsure and hence seeks guidance and confirmation from the detective. And, as we know, some detectives are not reluctant to tell witnesses when their choices identified the suspect.

But if a *double-blind procedure* were to be used, in which the lineup administrator is unaware of the "correct" answer, neither subtle nor overt communication would be made, and a purer estimate of the accuracy of the witness's memory and their confidence level could be determined (Haw and Fisher, 2004).

Rule 2: "Eyewitnesses should be told explicitly that the perpetrator might not be in the lineup or photospread and therefore eyewitnesses should not feel that they must make an identification. They should also be told that the person administering the lineup does not know which person is the suspect in the case" (Wells et al., 1998, p. 629).

Consider the reaction of an eyewitness when he or she is shown a lineup; it probably is something like this: "They wouldn't have gone to this trouble unless they have a suspect. So one of these guys must have done it." If the lineup is seen as a "multiple choice" question without the option "none of the above," the question is an easier one and in fact one could use the "relative judgment" strategy, comparing your memory to the one that "looks" most like the one you remember. Thus, it is essential for the investigator to *emphasize* that the culprit might not be in the photoarray or lineup, by means of an instruction that states clearly that the perpetrator "may or may not be in the set of photos you are about to view." Empirical studies, analyzed by Steblay (1997), find that an explicit warning such as this significantly reduces the rate of incorrect identifications when the offender is *not* in the lineup, without significantly affecting correct identifications.

Rule 3: "The suspect should not stand out in the lineup or photoarray as being different from the distractors based on the eyewitness's previous description of the culprit or based on other factors that would draw extra attention to the suspect" (Wells et al., 1998, p. 630).

In previous lineups the ways that the suspect stood out included:

1. He or she was the only one who fit the verbal description that the eyewitness had given to the police earlier (Lindsay and Wells, 1980);
2. He or she was the only one dressed in the type of clothes worn by the perpetrator (Lindsay et al., 1987); or
3. The suspect's photo was taken from a different angle than the foils' photos (Buckhout and Friere [1975], cited by Wells and Seelau, 1995).

Wells et al. emphasize that distractors should not necessarily be selected to look like the police detectives' prime suspect; instead, they should be chosen to match

the description of the criminal given by the witness. Note that this recommendation goes against the common police procedure in which they choose foils to resemble the suspect, rather than resembling the witness's description of the offender.

Rule 4: "A clear statement should be taken from the eyewitness at the time of the identification and prior to any feedback as to his or her confidence that the identified person is the actual culprit" (Wells et al., 1998, p. 635).

Repeated questioning by authorities (police, investigators, prosecutors) may increase the confidence of the witness's answers (Shaw, 1996; Shaw and McClure, 1996). By the time witnesses reach the witness box at the actual trial, they may act quite differently from their initial response. The initial levels of confidence should be recorded. In response to the above guidelines (and especially Rule 4), suggesting that they do not go far enough, Kassin (1998) has suggested one more rule, that the identification process (especially the lineup and the interaction between the detective and the witness) be videotaped, so that attorneys, the judge, and the jury can later assess for themselves whether the reports of the procedure by police are accurate (see also Judges, 2000). Unfortunately, since this is rarely if ever done, the attorneys, judge, and jury only see the end-product of an identification procedure, rather than the actual collection of the eyewitness evidence.

A fifth and important "rule" not included in Wells et al. (1998), has been strongly advocated by eyewitness researchers and mentioned in the United States Department of Justice guidelines (*Eyewitness Evidence: A Guide for Law Enforcement*, see below): "Scientific research indicates that identification procedures such as lineups and photo arrays produce more reliable evidence when the individual lineup members or photographs are shown to the witness sequentially—one at a time—rather than simultaneously" (Technical Working Group on Eyewitness Evidence, 1999, p. 9). Standard police lineups have traditionally used simultaneous procedures. However, under those conditions, eyewitnesses tend to compare lineup members to each other to determine which one most closely resembles their memory of the perpetrator, a process called "relative judgment" (see above). Lindsay and Wells (1985) devised an alternative lineup presentation technique, sequential presentation, that reduces or eliminates relative judgment by essentially forcing the witness to use an absolute criterion on each picture (i.e., a yes or no decision) before seeing the next one. This sequential presentation technique has been shown to reduce the rate of false alarms with little or no effect on correct identification rates (Lindsay et al., 1991; Steblay et al., 2001). Recent field studies testing the sequential method in the field has yielded success (Hennepin County, Minnesota; see Klobuchar et al., 2006), though results were unfortunately methodologically compromised in another (Cook County, Illinois; see the group of articles on the Cook County study in the February 2008 issue of *Law and Human Behavior*; see also O'Toole, 2006).

Since the publication of these "rules," there has been a clear acceptance of their worth and importance in psychology, law enforcement, and the courts. In October 1999, the United States Department of Justice published a set of guidelines or recommendations for the collection and preservation of eyewitness evidence,

entitled *Eyewitness Evidence: A Guide for Law Enforcement* (Technical Working Group on Eyewitness Evidence, 1999; a training manual for law enforcement was released later; see Technical Working Group on Eyewitness Evidence, 2003). The guide covers interview techniques such as those discussed in this chapter, and recommends procedures for the collection of eyewitness evidence by use of lineups, photospreads, and so on, including double-blind and sequential techniques. In 2001, the Attorney General of New Jersey, John Farmer Jr, ordered the official adoption and implementation of the recommendations of the guide for all lineups and photospreads in that state (see Kolata and Peterson, 2001). Other places (e.g., North Carolina in 2007), have adopted the recommendations of the guide by statute. It is also worth noting that Kebbell (2000) suggests that the law in England and Wales comports reasonably well with the recommendations in Wells et al. (1998) and thus in the guide.

Judges' Decisions on the Admissibility of Expert Testimony

In trials in which the testimony of an eyewitness is potentially pivotal and eyewitness's accuracy is an issue, psychologists have often been denied the opportunity to testify. Buckhout (1983) reported that, in New York by that time, "I have testified before juries in about 10 cases and been kept out too many times to count" (1983, p. 67). Fulero (1988) concluded that by 1988, psychologists have been allowed to testify about eyewitness accuracy for the defense in at least 450 cases in 25 states, but some states still prevent them from doing so (see for example, *Commonwealth of Pennsylvania v. Abdul-Salaam*, 1996).

Why not? Some judges fear that an eyewitness expert's testimony will be so powerful that it will usurp the jury's role as the fact finders in the case. A second reason is that judges may fear a "battle of the experts." Yet a third is that judges may feel that psychology supposedly does not possess information beyond the common knowledge of ordinary persons, and therefore eyewitness expert testimony would not qualify for the usual criterion for expert testimony. These two reasons can be collapsed together because psychology has generated research, the conclusions of which experts generally support, and yet at the same time, controversy exists within the field over the propriety of testifying and the appropriate role. Each of these issues is discussed here.

Kassin et al. (1989) surveyed 63 experts on eyewitness testimony; at least 80 percent of these experts agreed that research results on each of the following topics were consistent enough to present in court: the relationship between accuracy and confidence, the lineup instructions, the impact of exposure time, and unconscious transference (i.e., the misidentification of someone familiar to the witness from a context other than the crime), as well as other topics. More than 70 percent of the experts believed that the tendency to overestimate the duration of the event, the cross-racial identification bias of White witnesses, and lineup fairness generated

consistent research findings. This survey was recently repeated in 2001 (Kassin et al., 2001), with quite similar results.

Such experts have often testified as expert witnesses in criminal and civil cases around the country, and even in other countries (see Buckhout, 1983; Loftus, 1983; Wells, 1986; Penrod et al., 1995; Leippe, 1995). But others have proposed that the research is not sufficiently conclusive or applicable (Konecni and Ebbesen, 1986; McCloskey and Egeth, 1983; McCloskey, Egeth, and McKenna, 1986). Some of these psychologists have testified to that effect (see e.g., *People v. Legrand*, 2002, a decision disallowing eyewitness expert testimony in New York that was recently reversed, *People v. Legrand*, 2007 for example), though judges increasingly appear to be convinced of the scientific merit of such expert testimony (see *United States v. Smithers*, 2000; *United States v. Norwood*, 1996; *State of Ohio v. Echols*, 1998; *People v. Smith*, 2002; *State of Tennessee v. Copeland*, 2007; see also Penrod et al., 1995). Interestingly, while other countries have been at the forefront of psychological research on eyewitness identification, they have tended to lag behind in the use of expert testimony (see in Canada, *R. v. McIntosh*, 1997; but see also *R. v. Miaponoose*, 1996 for an indication of possible change).

Despite that, it seems clear that expert witnesses have a good deal to offer with respect to helping jurors understand how the variables affecting eyewitness reliability work (Leippe, 1995; Penrod et al., 1995). Indeed, by now, there are estimates that psychologists have testified in over a thousand cases in the United States (Penrod et al., 1995; Cutler and Penrod, 1995), and this number is increasing as case law becomes more amenable to eyewitness expert testimony (see e.g., *United States v. Smithers*, 2000; *State of Ohio v. Echols*, 1998).

Expert testimony about the determinants of eyewitness accuracy is an example of what Monahan and Walker (1988) have called "social framework testimony"; that is, it presents "general conclusions from social science research" in order to assist the fact-finder (whether that is judge or jury) "in determining factual issues in a specific case" (Monahan and Walker, 1988, p. 470). A judge's decision to admit or exclude scientific testimony is usually based on a combination of four criteria: the scientific nature of the work, the relevance of the work, the general agreement among experts in the area, and the extent to which the expert might unduly influence the jury (Wells, 1995, p. 729; see Chapter 1 of this volume for a more complete discussion of the admissibility of expert testimony). But in real life, matters are not so straightforward; "From a legal and public policy perspective … there is a problem to the extent that the variation in admissibility decisions is attributable more to ambiguity in the criteria for admissibility, the idiosyncratic views of the trial judge, or the characteristics of the jurisdiction than it is to the specific characteristics or needs of the case" (Wells, 1995, p. 729).

How can psychologists convince trial judges of the importance of the psychological findings? There are two important points that emerge from the research findings: the tendency for fact-finders not to be adequately informed on the topic, and the high level of consistency in the conclusions drawn by experts in this area. Recent United

States Supreme Court case law (*Daubert v. Merrell Dow Pharmaceuticals Inc.*, 1993) reinforces the importance of the expert's helpfulness to the jury by provision of information that is not "within the ken of the average layperson," and of the scientific reliability and validity of the information that is to be provided (see Penrod et al., 1995).

How Accurate is the Knowledge of Jurors?

Until the mid-1970s, expert testimony in such cases was rarely offered or admitted; among reasons given by judges for exclusion were that "jurors already know all this" and that experts would "waste the court's time" (Leippe, 1995, p. 912; see also Penrod et al., 1995). But jurors are often in error in two respects: They overestimate the level of accuracy of eyewitnesses and they do not appreciate the impact of either estimator or system factors on reducing accuracy. Laypersons usually begin with the assumption that the memory of an adult eyewitness is accurate (Leippe, 1995) and hence they expect a far greater percentage of witnesses to be accurate than are found in the field studies that create a mock crime and determine actual levels of eyewitness accuracy (Brigham and Bothwell, 1983; Wells, 1984b; 1984c; Wells and Leippe, 1981; Lindsay et al., 1981).

An assumption that "jurors already know all this" is clearly unwarranted. Four different surveys came to the same conclusion: "that much of what is known about eyewitness memory—that eyewitness experts might talk about in court—is not common sense" (Leippe, 1995, p. 921). Specific findings of these surveys documented this conclusion, as follows:

1. Deffenbacher and Loftus (1982) gave a set of multiple-choice questions on variables associated with eyewitness accuracy to college students and nonstudents with and without jury experience. At least half the respondents chose the wrong answer (i.e., an answer in conflict with the direction of empirical findings) on questions about the confidence-accuracy relationship, cross-racial bias in identification, and weapon focus.
2. Using law students, legal professionals, undergraduate students, and adults as participant subjects, Yarmey and Jones (1983) found that respondents did not recognize the empirically-derived relationships between level of accuracy and such factors as eyewitness's confidence, the presence of a weapon, and the status of the witness (i.e., that police are no better at identification than are other witnesses).
3. Using those 13 empirical findings deemed by experts to be reliable enough to testify about, Kassin and Barndollar (1992) found that significantly fewer students and adults than experts considered the findings as reliable. In four of the 13 reliable findings, the majority of the students and adults disagreed with the experts.

4. Brigham and Wolfskeil (1983) surveyed trial attorneys and found, not surprisingly, that prosecutors were much more likely to believe that eyewitnesses were accurate than were criminal defense attorneys.

How Generally Accepted is the Work in this Area?

Judges have been shown to harbor misconceptions and errors about the factors affecting eyewitness reliability (Wise and Safer, 2003; 2004). More recently, an extensive and careful survey of actual people called for jury duty in Washington, DC, showed that they held the same sorts of misconceptions and errors as had been found earlier (O'Toole, 2005).

A second argument important in order to persuade judges to admit psychological testimony is the consistency of agreement among experts on the phenomenon. A survey by Kassin, Ellsworth, and Smith (1989; repeated by Kassin et al., 2001) of 63 active psychological researchers determined just which specific phenomena, in their opinion, were reliable enough to testify about in court. Box 1 describes those findings that at least 70 percent of this sample felt were reliable, in both 1989 and then 2001.

Box 1: What is Reliable Enough to Testify About?

The following are the findings that at least 70 percent of the researcher-experts surveyed by Kassin et al. (1989) and by Kassin et al. (2001) rated as reliable enough to include in courtroom testimony (1989 and 2001 percentages):

1. Wording of questions. An eyewitness's testimony about an event can be affected by how the questions put to that witness are worded. (97%) (98%)
2. Lineup instructions. Police instructions can affect an eyewitness's willingness to make an identification and/or the likelihood that he or she will identify a particular person. (95%) (98%)
3. Postevent information. Eyewitness testimony about an event often reflects not only what they actually saw but information they obtained later on. (87%) (94%)
4. Accuracy-confidence. An eyewitness's confidence is not a good predictor of his or her identification accuracy. (87%) (87%)
5. Attitudes and expectations. An eyewitness's perception and memory of an event may be affected by his or her attitudes and expectations. (87%) (92%)
6. Exposure time. The less time an eyewitness has to observe an event, the less well he or she will remember it. (85%) (81%)
7. Unconscious transference. Eyewitnesses sometimes identify as a culprit someone they have seen in another situation or context. (85%) (81%)

8. Showups. The use of a one-person showup instead of a full lineup increases the risk of misidentification. (83%) (74%)
9. Forgetting curve. The rate of memory loss for an event is greatest right after the event, and then levels off over time. (83%) (83%)
10. Cross-racial/White. White eyewitnesses are better at identifying other White people than they are at identifying Black people. (79%) (90%)
11. Lineup fairness. The more the members of a lineup resemble the suspect, the higher is the likelihood that identification of the suspect is accurate. (77%) (70%)
12. Time estimation. Eyewitnesses tend to overestimate the duration of events. (75%) (not asked)
13. Stress. Very high levels of stress impair the accuracy of eyewitness testimony. (71%) (60%)
14. Weapons focus. The presence of a weapon impairs an eyewitness's ability to accurately identify the perpetrator's face. (57%) (87%)
15. Hypnotic suggestibility. Hypnosis increases suggestibility to leading and misleading questions. (69%) (91%)
16. Confidence malleability. An eyewitness's confidence can be influenced by factors that are unrelated to identification accuracy. (not asked) (95%)
17. Mugshot-induced bias. Exposure to mugshots of a suspect increases the likelihood that the witness will later choose that suspect in a lineup. (not asked) (95%)
18. Child suggestibility. Young children are more vulnerable than adults to interviewer suggestion, peer pressures, and other social influences. (not asked) (94%)
19. Alcoholic intoxication. Alcoholic intoxication impairs an eyewitness's later ability to recall persons and events. (not asked) (90%)
20. Presentation format. Witnesses are more likely to misidentify someone by making a relative judgment when presented with a simultaneous (as opposed to sequential) lineup. (not asked) (81%)
21. Child accuracy. Young children are less accurate as witnesses than adults. (not asked) (70%)
22. Description-matched foils. The more that members of a lineup resemble a witness's description of the culprit, the more accurate an identification of the suspect is likely to be. (not asked) (71%)

Note: Percentages of experts rating the statement as "reliable enough" are given in parentheses beside each statement. Data from Kassin et al. (1989) and Kassin et al. (2001).

These conclusions are based, for most of the findings, on a multitude of studies using a variety of methods and types of subjects. As Leippe (1995) observed, "in matters of reliability, a number of eyewitness research findings score highly. They are replicable, the opposite findings (as opposed to simply null findings) are seldom reported, the research has high internal validity, and the settings and measures often

have high mundane realism in terms of approximating certain eyewitness situations. A strong argument can be made for reliability and validity" (Leippe, 1995, p. 918).

Of course, as in any other field of endeavor, not all experts agree with the above statement. A few psychologists, including Rogers Elliott (1993), Vladimir Konecni and Ebbe Ebbesen (1986), and Michael McCloskey and Howard Egeth (1983; Egeth, 1993) have been critical for several reasons, including their assertion that the findings have not reached a level of consistency necessary for application in the courts. But these psychologists are clearly very much in the minority, and sometimes the issue of dispute is more a matter of philosophical disagreement about how and when psychological research findings should be presented in court settings, rather than whether or not there is a stable body of research or what conclusions are being drawn from the research studies.

Conclusions

This chapter attempts to demonstrate that the field of psychology has much to offer police and the legal system to help ensure the most reliable use of eyewitness evidence. Some police detectives are clearly uncomfortable with representatives from another discipline "telling them how to run their business," and psychologists always need to remember the day-to-day pressures and constraints on police conducting crime investigations. Still, it is important to remind those in the field of law enforcement that indeed, the goals of all those who work in the system are the same—because if the wrong person is apprehended and convicted, the right one remains free to commit other crimes. Understanding eyewitness error and the factors that contribute to it, and improving eyewitness evidence collection techniques, can help to ensure the best use of that evidence.

References

Behrman, B., and Davey, S. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behavior*, 25, 475–91.

Bothwell, R.K., Deffenbacher, K.A., and Brigham, J.C. (1987). Correlation of eyewitness accuracy and confidence: The optimality hypothesis revisited. *Journal of Applied Psychology*, 72, 691–5.

Bradfield, S.L., Wells, G.L., and Olson, E.A. (2002). The damaging effect of confirming feedback on the relation between eyewitness certainty and identification accuracy. *Journal of Applied Psychology*, 87, 112–20.

Brewer, N., Keast, A., and Rishworth, A. (2002). Improving the confidence-accuracy relation in eyewitness identification: Evidence from correlation and calibration. *Journal of Experimental Psychology: Applied*, 8, 44–56.

Brigham, J.C., and Bothwell, R.K. (1983). The ability of prospective jurors to estimate the accuracy of eyewitness identifications. *Law and Human Behavior*, 7, 19–30.

Brigham, J.C., Maass, A., Snyder, L.D., and Spaulding, K. (1982). Accuracy of eyewitness identifications in a field setting. *Journal of Personality and Social Psychology*, 42, 673–81.

Brigham, J.C., and Wolfskeil, M.P. (1983). Opinions of attorneys and law enforcement personnel on the accuracy of eyewitness identifications. *Law and Human Behavior*, 7, 337–49.

Brimacombe, C.A.E., Jung, S., Garrioch, L., and Allison, M. (2003). Perceptions of older adult eyewitnesses: Will you believe me when I'm 64? *Law and Human Behavior*, 27, 507–22.

Buckhout, R. (1974). Eyewitness testimony. *Scientific American*, 231, 23–31.

Buckhout, R. (1983). Psychologist v. the judge: Expert testimony on identification. *Social Action and the Law*, 9(3), 67–76.

Buckhout, R., and Friere, V. (1975). *Suggestibility in Lineups and Photospreads: A Casebook for Lawyers* (Center for Responsive Psychology Monograph No. CR-5), New York: Brooklyn College.

Christianson, S.A. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin*, 112, 284–309.

Commonwealth v. Vardinski, 438 Mass. 444, 450, 780 N.E.2d 1278 (2003).

Commonwealth of Pennsylvania v. Abdul-Salaam, 544 Pa. 514, 524, 678 A.2d 342, 347 (1996).

Connors, E., Lundregan, T., Miller, N., and McEwan, T. (1996). Convicted by Juries, Exonerated by Science: Case Studies in the Use of DNA Evidence to Establish Innocence after Trial. Alexandria, VA: National Institute of Justice.

Cutler, B.L., and Penrod, S.D. (1995). Mistaken Identification: The Eyewitness, Psychology, and the Law. New York: Cambridge University Press.

Cutler, B.L., Penrod, S.D., and Martens, T.K. (1987). The reliability of eyewitness identification: The role of system and estimator variables. *Law and Human Behavior*, 11, 233–58.

Daubert v. Merrell Dow Pharmaceuticals Inc., 507 U.S. 579, 113 S.Ct. 2786 (1993).

Davis, J., and Gonzalez, R. (1996, February). Relative and absolute judgments of eyewitness identification. Paper presented at the meeting of the American Psychology-Law Society, Hilton Head, S.C.

Deffenbacher, K. (1983). The influence of arousal on reliability of testimony. In S. Lloyd-Bostock and B.R. Clifford (eds). *Evaluating Witness Evidence: Recent Psychological Research and New Perspectives*. Chichester: Wiley. pp. 235–51.

Deffenbacher, K.A., Bornstein, B.H., Penrod, S.D., and McGorty, E.K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior*, 28, 687–706.

Deffenbacher, K.A., and Loftus, E.F. (1982). Do jurors share a common understanding concerning eyewitness behavior? *Law and Human Behavior*, 6, 15–30.

Douglass, A.B., and McQuiston-Surrett, D.M. (2006). Post-identification feedback: Exploring the effects of sequential photospreads and eyewitnesses' awareness of the identification task. *Applied Cognitive Psychology*, 20, 991–1007.

Douglass, A.B., and Steblay, N. (2006). Memory distortion in eyewitnesses: A meta-analysis of the post-identification feedback effect. *Applied Cognitive Psychology*, 20, 859–69.

Easterbrook, J.A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, 66, 183–201.

Egeth, H.E. (1993). What do we *not* know about eyewitness identification? *American Psychologist*, 48, 577–80.

Elliott, R. (1993). Expert testimony about eyewitness identification: A critique. *Law and Human Behavior*, 17, 423–37.

Ellis, H.D., Davies, G.M., and Shepherd, J.W. (1977). Experimental studies of face identification. *Journal of Criminal Defense*, 3, 219–34.

Ellis, H.D., Shepherd, J.W., and Davies, G.M. (1980). The deterioration of verbal descriptions of faces over different delay intervals. *Journal of Police Science and Administration*, 8, 101–6.

Fulero, S.M. (1988, August). Eyewitness expert testimony: An overview and annotated bibliography, 1931–1988. Paper presented at the meetings of the American Psychological Association, Atlanta, GA.

Fulero, S. (1997). Review of *Mistaken Identification*, by Brian Cutler and Steven Penrod. New York: Cambridge University Press, 1995. *Contemporary Psychology*, 42, 395–6.

Fulero, S. (2000, August). The effects of stress on time overestimation: A field study. Paper presented at the American Psychological Association convention, Washington, DC.

Garrett, B. (2007). Judging innocence. *Columbia Law Review*, 100, 101–90.

Gonzalez, R., Ellsworth, P.C., and Pembroke, M. (1993). Response biases in lineups and showups. *Journal of Personality and Social Behavior*, 6, 1–13.

Haw, R. and Fisher, R.P. (2004). Effects of administrator-witness contact on eyewitness identification accuracy. *Journal of Applied Psychology*, 89, 1106–12.

Hockley, W.E., Hemsworth, D.H., and Consoli, A. (1999). Shades of the mirror effect: Recognition of faces with and without sunglasses. *Memory and Cognition*, 27, 128–38.

Huff, C.R., Rattner, A., and Sagarin, E. (1996). Convicted but Innocent: Wrongful Conviction and Public Policy. Thousand Oaks, CA: Sage.

Judges, D.P. (2000). Two cheers for the Department of Justice's eyewitness evidence: A guide for law enforcement. *Arkansas Law Review*, 53, 231–97.

Kassin, S.M. (1998). Eyewitness identification procedures: The fifth rule. *Law and Human Behavior*, 22, 649–53.

Kassin, S.M., and Barndollar, K.A. (1992). The psychology of eyewitness testimony: A comparison of experts and prospective jurors. *Journal of Applied Social Psychology*, 22, 1241–9.

Kassin, S.M., Ellsworth, P.C., and Smith, V.L. (1989). The "general acceptance" of psychological research on eyewitness testimony: A survey of the experts. *American Psychologist*, 44, 1089–98.

Kassin, S., Tubb, V.A., Hosch, H.M., and Memon, A. (2001). On the "general acceptance" of eyewitness testimony research: A new survey of the experts. *American Psychologist*, 56, 405–16.

Kebbell, M.R. (2000). The law concerning the conduct of identification parades in England and Wales: How well does it satisfy the recommendations of the American Psychology-Law Society? *Law and Human Behavior*, 24, 309–15.

Klobuchar, A., Steblay, N., and Caligiuri, H.L. (2006). Symposium: Reforming eyewitness identification: Convicting the guilty, protecting the innocent: Improving eyewitnesses' identifications: Hennepin county's blind sequential lineup pilot project. *Cardozo Public Law, Policy and Ethics Journal*, 4, 381–413.

Kolata, G., and Peterson, I. (2001, July 21). New way to insure eyewitnesses can ID the right bad guy. *New York Times*, p. A-1.

Konecni, V.J., and Ebbesen, E.B. (1986). Courtroom testimony by psychologists on eyewitness identification issues: Critical notes and reflections. *Law and Human Behavior*, 10, 117–26.

Lavrakas, P.J., and Bickman, L. (1975, August). What makes a good witness? Paper presented at the annual meeting of the American Psychological Association, Chicago, IL.

Leippe, M.R. (1995). The case for expert testimony about eyewitness memory. *Psychology, Public Policy, and Law*, 1, 909–59.

Leippe, M.R., Wells, G.L., and Ostrom, T.M. (1978). Crime seriousness as a determinant of accuracy in eyewitness identification. *Journal of Applied Psychology*, 63, 345–51.

Lindsay, R.C.L., Lea, J.A., and Fulford, J.A. (1991). Sequential lineup presentation: Technique matters. *Journal of Applied Psychology*, 76, 741–5.

Lindsay, R.C.L., Pozzulo, J.D., Craig, W., Lee, K., and Corber, S. (1997). Simultaneous lineups, sequential lineups, and showups: Eyewitness identification decisions of adults and children. *Law and Human Behavior*, 21, 391–404.

Lindsay, R.C.L., Wallbridge, H., and Drennan, D. (1987). Do the clothes make the man? An exploration of the effect of lineup attire on eyewitness identification accuracy. *Canadian Journal of Behavioural Science*, 19, 463–78.

Lindsay, R.C.L, and Wells, G.L. (1980). What price justice? Exploring the relationship of lineup fairness to identification accuracy. *Law and Human Behavior*, 4, 303–13.

Lindsay, R.C.L., and Wells, G.L. (1985). Improving eyewitness identification from lineups: simultaneous versus sequential lineup presentations. *Journal of Applied Psychology*, 70, 556–64.

Lindsay, R.C.L., Wells, G.L., and Rumpel, C. (1981). Can people detect eyewitness identification accuracy within and between situations? *Journal of Applied Psychology*, 66, 79–89.

Loftus, E.F. (1983). Silence is not golden. *American Psychologist*, 38, 564–72.

Loftus, E.F. (1993). Psychologists in the eyewitness world. *American Psychologist*, 48, 550–52.

Loftus, E.F., Loftus, G.R., and Messo, J. (1987). Some facts about "weapon focus." *Law and Human Behavior*, 11, 55–62.

Loftus, E., Schooler, J., Boone, S., and Kline, D. (1987). Time went by so slowly: Overestimation of event duration by males and females, *Applied Cognitive Psychology*, 1, 3–13.

Luce, T. (1974). The role of experience in inter-racial recognition. *Personality and Social* Psychology Bulletin, 1, 39–44.

Luus, C.A.E. (1991). Eyewitness confidence: Social influence and belief perseverance. Unpublished doctoral dissertation, Iowa State University.

Luus, C.A.E., and Wells, G.L. (1994). The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology*, 79, 714–23.

Malpass, R.S., and Devine, P.G. (1980). Realism and eyewitness identification research. *Law and Human Behavior*, 4, 347–58.

Malpass, R.S., and Devine, P.G. (1983). Measuring the fairness of eyewitness identification lineups. In S.M.A. Lloyd-Bostock and B.R. Clifford (eds), *Evaluating Witness Evidence*. New York: Wiley. pp. 81–102.

McCloskey, M., and Egeth, H.E. (1983). Eyewitness identification: What can a psychologist tell a jury? *American Psychologist*, 38, 550–63.

McCloskey, M., Egeth, H., and McKenna, J. (1986). [Introduction]: The experimental psychologist in court: The ethics of expert testimony. *Law and Human Behavior*, 10, 1–13.

Meissner, C., and Brigham, J.C. (2001). Twenty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychology, Public Policy, and Law*, 7, 3–35.

Monahan, J., and Walker, L. (1988). Social science research in law: A new paradigm. *American Psychologist*, 43, 465–72.

Morgan, C.A., Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., Baronoski, M., and Southwick, S.M. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *International Journal of Law and Psychiatry*, 27, 265–79.

Murray, D.M., and Wells, G.L. (1982). Does knowledge that a crime was staged affect eyewitness performance? *Journal of Applied Social Psychology*, 12, 42–53.

O'Toole, T.P. (2005, April). District of Columbia Public Defender Survey: What do jurors understand about eyewitness reliability? *The Champion*, 28–32.

O'Toole, T.P. (2006, August). What's the matter with Illinois? How an opportunity was squandered to conduct an important study on eyewitness identification procedures. *The Champion*, 18–20.

Penrod, S.D., Fulero, S.M., and Cutler, B.L. (1995). Expert psychological testimony on eyewitness reliability before and after Daubert: The state of the law and the science. *Behavioral Sciences and the Law*, 13, 229–59.

People v. Legrand, 196 N.Y.Misc. 2d 179 (Sup.Ct. N.Y. UYCounty, 2002), reversed NY Slip Op. 02588 (2007).

People v. Smith, 743 N.Y.S.2d 246 (2002).

Pozzulo, J.D., and Lindsay, R.C.L. (1998). Identification accuracy of children versus adults: A meta-analysis. *Law and Human Behavior*, 22, 549–70.

R. v. McIntosh, 117 C.C.C. (3d) 385 (Ont. C.A.) (1997).

R. v. Miaponoose, 110 C.C.C. (3d) 445 (Ont. C.A.) (1996).

Rattner, A. (1988). Convicted but innocent: Wrongful conviction and the criminal justice system. *Law and Human Behavior*, 12, 283–93.

Read, J.D., Vokey, J.R., and Hammersley, R. (1990). Changing photos of faces: Effects of exposure duration and photo similarity on recognition and the accuracy-confidence relationship. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 870–82.

Scheck, B., Neufeld, P., and Dwyer, J. (2000). *Actual Innocence*. New York: Random House.

Semmler, C., Brewer, N., and Wells, G.L. (forthcoming). Effects of postidentification feedback on eyewitness identification and nonidentification confidence. *Journal of Applied Psychology*.

Shapiro, P.N., and Penrod, S.D. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, 100, 139–56.

Shaw, J.S., III. (1996). Increases in eyewitness confidence resulting from postevent questioning. *Journal of Experimental Psychology: Applied*, 2, 126–46.

Shaw, J.S., III, and McClure, K.A. (1996). Repeated postevent questioning can lead to elevated levels of eyewitness confidence. *Law and Human Behavior*, 20, 629–53.

Sporer, S.L., Penrod, S.D., Read, J.D., and Cutler, B.L. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin*, 118, 315–27.

State of Ohio v. Echols, 128 Ohio App.3d 677 (1998).

State of Tennessee v. Copeland, WL 1498396 (2007).

Steblay, N.M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior*, 16, 413–24.

Steblay, N.M. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior*, 21, 283–98.

Steblay, N., Dysart, J., Fulero, S., and Lindsay, R.C.L. (2003). A meta-analytic comparison of showup and lineup identification accuracy. *Law and Human Behavior*, 27, 523–40.

Steblay, N., Lindsay, R., Fulero, S., and Dysart, J. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic review. *Law and Human Behavior*, 25, 459–74.

Stovall v. Denno, 388 U.S. 293 (1967).

Technical Working Group on Eyewitness Evidence (1999). *Eyewitness Evidence: A Guide for Law Enforcement*. Washington, DC: United States Department of Justice, National Institute of Justice. Document No. NCJ 178240.

Technical Working Group on Eyewitness Evidence (2003). *Eyewitness Evidence: A Trainer's Manual for Law Enforcement*. Washington, DC: United States Department of Justice, National Institute of Justice. Document No. NCJ 188678.

United States v. Norwood, 939 F.Supp. 1132 (Dist. Ct. N.J. 1996).

United States v. Smithers, 212 F.3d 306 (2000).

Wagenaar, W.A., and Schrier, J. van der (1994). Face recognition as a function of distance to Illumination: a practical tool for use in the courtroom. Paper presented at the Fourth European Conference on Law and Psychology, Barcelona, Spain.

Wells, G.L. (1978). Applied eyewitness testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology*, 36, 1546–57.

Wells, G.L. (1984a). The psychology of lineup identifications. *Journal of Applied Social Psychology*, 14, 89–103.

Wells, G.L. (1984b). A re-evaluation of the expert testimony issue. In G.L. Wells and E.F. Loftus (eds), *Eyewitness Testimony: Psychological Perspectives*. New York: Cambridge University Press.

Wells, G.L. (1984c). How adequate is human intuition for judging eyewitness testimony? In G.L. Wells and E.F. Loftus (eds), *Eyewitness Testimony: Psychological Perspectives*. New York: Cambridge University Press.

Wells, G.L. (1986). Expert psychological testimony: empirical and conceptual analyses of effects. *Law and Human Behavior*, 10, 83–95.

Wells, G.L. (1993). What do we know about eyewitness identification? *American Psychologist*, 48, 553–71.

Wells, G.L. (1995). Scientific study of witness memory: Implications for public and legal policy. *Psychology, Public Policy, and Law*, 1, 726–31.

Wells, G.L., and Bradfield, A.L. (1998). "Good, you identified the suspect": Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology*. 83, 360–76.

Wells, G.L., and Bradfield, A.L. (1999). Distortions in eyewitnesses' recollections: Can the postidentification feedback effect be moderated? *Psychological Science*, 10, 138–44.

Wells, G., Fisher, R., Lindsay, R., Turtle, J., Malpass, R., and Fulero, S. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist*, 55, 581–98.

Wells, G.L., and Leippe, M.R. (1981). How do triers of fact infer the accuracy of eyewitness identifications? Memory for peripheral detail can be misleading. *Journal of Applied Psychology*, 66, 682–7.

Wells, G.L., Leippe, M.R., and Ostrom, T.M. (1979). Guidelines for empirically assessing the fairness of a lineup. *Law and Human Behavior*, 3, 285–93.

Wells, G.L., Lindsay, R.C.L., and Ferguson, T. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. *Journal of Applied Psychology*, 64, 440–48.

Wells, G.L., Memon, A., and Penrod, S. (2006). Eyewitness evidence: Improving its probative value. *Psychological Science in the Public Interest*, 7, 45–75.

Wells, G.L., and Olson, E.A. (2003). Eyewitness testimony. *Annual Review of Psychology*, 54, 277–95.

Wells, G.L., Olson, E., and Charman, S. (2003). Distorted retrospective eyewitness reports as functions of feedback and delay. *Journal of Experimental Psychology: Applied*, 9, 42–52.

Wells, G.L., and Seelau, E. (1995). Eyewitness identification: Psychological research and legal policy on lineups. *Psychology, Public Policy, and Law*, 1, 765–91.

Wells, G., Small, M., Penrod, S., Malpass, R., Fulero, S., and Brimacombe, C.A.E. (1998). Good practice recommendations for lineups and photospreads. *Law and Human Behavior*, 22, 603–47.

Wise, R.A., and Safer, M.A. (2003). A survey of judges' knowledge and beliefs about eyewitness testimony. *Court Review*, 40(1), 6–16.

Wise, R.A., and Safer, M.A. (2004). What U.S. judges know and believe about eyewitness testimony. *Applied Cognitive Psychology*, 18, 427–43.

Wrightsman, L., and Fulero, S. (2005). *Forensic Psychology* (2nd ed.). Belmont, CA: Wadsworth.

Yarmey, A.D. (1979). *The Psychology of Eyewitness Testimony*. New York: Free Press.

Yarmey, A.D., and Jones, H.P.T. (1983). Is the psychology of eyewitness identification a matter of common sense? In S.M.A. Lloyd-Bostock and B.R. Clifford (eds), *Evaluating Witness Evidence: Recent Psychological Research and New Perspectives*. Chichester: Wiley. pp. 13–40.

Yarmey, A.D., Yarmey, M.J., and Yarmey, A.L. (1996). Accuracy of eyewitness identifications in showups and lineups. *Law and Human Behavior*, 20, 459–77.

Chapter 4

Insanity in the Courtroom: Issues of Criminal Responsibility and Competency to Stand Trial

Patricia A. Zapf, Tina M. Zottoli, and Gianni Pirelli

Synopsis

Competency to stand trial and criminal responsibility (insanity) are two legal issues wherein a defendant's mental state is called into question. Mental state at the time of the alleged offense is at issue in criminal responsibility whereas present mental state is at issue in competency to stand trial. This chapter introduces each of these legal issues and reviews relevant research and commentary. With respect to criminal responsibility, various insanity defense standards are reviewed, data on the use and success of the insanity defense is presented, and the characteristics of insanity acquittees are discussed. In addition, problems with the application of the insanity defense in the courtroom—such as those pertaining to judicial instruction, juror attitudes, and jurors' implicit theories about insanity—are highlighted. With respect to competency to stand trial, the standard for competency is outlined, procedures related to the determination of competency are explained, and the characteristics of defendants referred for competency evaluation are presented. In addition, problems with the application of the competency doctrine within the courtroom—such as the contextual nature of the evaluation and the balancing of the needs of the state with those of the defendant—are highlighted. Recommendations for policy/procedural solutions to some of the problems highlighted are also presented.

Insanity in the Courtroom: Issues of Criminal Responsibility and Competency to Stand Trial

Criminal responsibility and competency to stand trial are two issues of significance for individuals who have been charged with a crime and for

whom mental state may be an issue. Both criminal responsibility and competency are legal issues involving a defendant's mental state and have been referred to as "insanity" by the courts; however, competency involves an examination of mental state immediately prior to or during the trial whereas criminal responsibility requires an examination of the defendant's mental state at the time of the commission of the alleged offense. Here competency and criminal responsibility are discussed within the context of adult criminal courts. This chapter will introduce each of these legal issues and provide a brief review[1] of relevant research and discussion of practical application within the courtroom.

Insanity/Criminal Responsibility

The issue of criminal responsibility, or insanity, has to do with an individual's mental state at the time of the offense. The basic philosophy—which stems back to the earliest recordings of Hebrew law and is still in effect throughout the United States, Canada, New Zealand, Australia, England, Wales, and other countries—is that to convict a person charged with a crime, he or she must be considered responsible for his or her criminal behavior. That is, the criminal behavior must have been a product of free will. If a defendant's behavior was not a product of free will, then he or she should not be held responsible for the crime (see Zapf et al., 2006 for a review of criminal responsibility). Generally, two basic elements of the crime must be proved: the *actus reus*—Latin for "guilty act," which refers to the physical element or physical act of the crime—and *mens rea*—Latin for "guilty mind," which refers to the mental element of the crime, most often considered to be the intention to commit the crime. Insanity is a defense that, generally, contests the *mens rea* component of the crime in that the individual, as a result of a mental disorder, was unable to formulate the requisite intention for the crime and, thus, should not be held responsible.

Insanity Defense Standards

Early (prior to the mid-nineteenth century) tests of insanity included the "good and evil" test, the "wild beast" test, and the "right and wrong" test; however, the case of Daniel M'Naghten in England in 1843 led to the adoption of the first formal definition of insanity. The M'Naghten rule provided that to establish a defense on the grounds of insanity,

1 As the evaluation of competency and criminal responsibility is beyond the scope of this chapter, the interested reader is referred to the following sources for more detail on the evaluation of these issues: Grisso (2003), Melton et al. (2007), Rogers and Shuman (2000), Zapf and Roesch (2006), and Zapf et al. (2006).

> 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)

Although slight variations on the wording of this legal test have occurred over time, the M'Naghten (knowledge/right-wrong) standard remains the legal test for insanity in Canada (Not Criminally Responsible on Account of Mental Disorder), Australia (Mental Incompetence), New Zealand, England and Wales, and about one-half of the jurisdictions in the United States.

Throughout the history of the U.S. legal system, a number of legal tests or standards of insanity have been used. The various insanity standards include: the irresistible impulse test, the *Durham* standard or Product rule, the American Law Institute (ALI) standard, and the Insanity Defense Reform Act (IDRA) standard.

The irresistible impulse test represented a broadening of the M'Naghten standard to include a volitional component. In 1844, Chief Justice Shaw of the Massachusetts Supreme Court held that while the right-wrong (cognitive) test was appropriate, a defendant who acted under the influence of an irresistible impulse (volitional) was not a free agent (*Commonwealth v. Rogers*, 1844). In 1866, this logic was made explicit in Justice Somerville's holding 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 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* by broadening the test even further. The *Durham* product test stated, "an accused is not criminally responsible if his unlawful act was the product of a mental disease or defect" and was intended to remedy problems associated with "the old right-wrong/irresistible impulse rule for insanity [being considered] antiquated [and] no longer reflecting the community's judgment as to who ought to be held criminally liable for socially destructive acts" (p. 976). In essence, the *Durham* test broadened the insanity standard to include any behavior that might be the product of mental disorder.

In 1962, the American Law Institute (ALI), in Section 4.01 of the Model Penal Code, proposed a legal test that encompassed both cognitive and volitional prongs: "A person is not responsible for criminal conduct if at the time of such conduct as a result of mental disease or defect he lacks substantial capacity either to appreciate

the criminality (or wrongfulness) of his conduct or to conform his conduct to the requirements of law." In 1972, the ALI rule was adopted by the D.C. Court of Appeals in *United States v. Brawner*, thus ending the use of the *Durham* product test by the D.C. Court of Appeals although a variant of the product test still continues to be used by New Hampshire state courts.

Since the much-publicized trial of John W. Hinckley, Jr and the ensuing public outrage at his successful use of the insanity defense after attempting to kill President Reagan in 1981, a great deal of court reform and legislative revision with respect to the insanity defense have occurred. In 1984, two years after Hinckley's NGRI acquittal, the Insanity Defense Reform Act (IDRA) was passed. The IDRA dictated that to be found not responsible a defendant must prove that, "as a result of a severe mental disease or defect, he was unable to appreciate the nature and quality or wrongfulness of his act" (p. 201). In effect, the IDRA standard represents a slightly restricted version of the original M'Naghten standard, which includes only a cognitive prong, and eliminates the volitional prong of the ALI standard. Currently, the most common insanity defense standards used throughout the United States involve either restricted versions of the traditional M'Naghten test (such as the IDRA standard) or the American Law Institute's formulation.

Subsequent to the Hinckley trial five states (Montana, Idaho, Utah, Nevada, and Kansas) abolished the insanity defense and others instituted alternatives such as Guilty But Mentally Ill (GBMI) provisions. The GBMI alternative allows for a finding of guilt but a reduction in either the seriousness of the charge or the severity of the punishment if mental disorder was determined to have influenced the criminal act. Approximately 13 states have provisions that allow for a defendant to be found Guilty But Mentally Ill (GBMI; see Arrigo, 1996 for a review; see also Borum and Fulero, 1999, for a discussion of various proposed insanity defense reforms). GBMI legislation was initially introduced in Michigan in 1975 in *People v. McQuillan* (1974), a case in which Michigan's automatic indefinite commitment of NGRI acquittees was found to be unconstitutional. The verdict was also adopted in Indiana in 1979 under similar circumstances. Following Hinckley's assault on President Reagan the stage was set for other states to pass GBMI legislation in response to the perceived abuses of the insanity plea. The GBMI verdict, which allows a criminal defendant to be sent to prison, rather than to a mental health facility, for treatment was not intended to replace the verdict of not guilty by reason of insanity (although it has in Utah and Nevada). While both the NGRI and GBMI verdicts were introduced in order to stem the perceived tide of violence committed by offenders who escape "justice," the GBMI verdict was aimed primarily at jurors with the hope that it would allow them a middle ground between guilty and NGRI.

Opponents of the GBMI verdict argue that it should be abolished on the grounds that it confuses and deceives jurors (Melville and Naimark, 2002; see also Palmer, 2000). In order to cover the punitive and abolitionist motivation behind the GBMI alternative, 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. As Beasley (1983) noted, however, there was no mention of the fact that few, if any, new funds were to be appropriated to the prison system to provide more treatment. In addition, provisions already existed in every state that passed GBMI to transfer laterally a disturbed prisoner into mental hospital settings for treatment if that was necessary. In commenting on this entire enterprise, Professor Richard Bonnie said, bluntly, "[The guilty but mentally ill verdict] should be rejected as nothing more than moral sleight of hand" (1983, p. 194).

Insanity in the Courtroom

While some research (discussed below) has examined public perceptions of insanity as well as rates of use and success of insanity pleas, other research has examined the process by which insanity pleas make it to court and whether insanity pleas are heard by judge or jury. Janofsky and colleagues (1996) examined the outcomes for all insanity pleas raised over a one-year period in Baltimore City's district and circuit courts and found that 95.8 percent of the pleas did not reach court. The vast majority of defendants withdrew their insanity plea immediately after undergoing an insanity evaluation, an additional group of the defendants had their charges dropped before trial, and others were found incompetent to stand trial. Only 4.2 percent of the initial insanity pleas occurred in court and in all cases the plea was not contested by either the state or the defense resulting in the defendants being found not criminally responsible. Further research found that insanity pleas represented less than 1 percent of all pleas (0.31 per 100 indictments) and successful insanity pleas over this one-year period represented far less than 1 percent of all indictments (0.013 successful insanity acquittals per 100 indictments).

Cirincione (1996) examined the processing of insanity pleas across seven states and determined that the vast majority of insanity cases were handled through either plea bargain (42.9%) or bench (judge) trial (42.7%), with relatively few (14.4%) being handled via jury trial. Of the cases handled via plea bargain, the vast majority resulted in conviction, with 87.9 percent of insanity defendants pleading guilty. Of the cases handled through bench trial, 54.6 percent resulted in an acquittal. Of the cases handled via jury trial, three quarters of the defendants were convicted. In interpreting these results Cirincione speculated that jury trials were more commonly used in contentious cases (involving severe crimes and less mental disorder) whereas bench trials were more likely when the prosecution did not contest the insanity claim.

Use of the Insanity Defense

Public perceptions of the insanity defense are that it is used frequently, is successful often, and serves as a "loophole" that allows guilty people to go free (Hans, 1986; Pasewark and Seidenzahl, 1979). However, empirical research reveals that the public

overestimates both the use and success of the insanity defense and underestimates the length of confinement of insanity acquittees (Silver et al., 1994).

Although exact rates of use and success of the insanity defense vary by jurisdiction, the insanity defense is rarely used and even more rarely successful. Silver et al. (1994) compared public perceptions of the insanity defense with empirical data on its actual use and found that the public estimated the use of the insanity defense to be 37 percent (or 37 per 100 felony indictments) whereas the actual use was 0.9 percent (less than 1 per 100 felony indictments), representing a public estimate that was 41 times greater than its actual use. Similarly, with respect to estimates regarding the success of the insanity defense, these authors report that the public estimated the success rate to be 44 percent (or 44 acquittals per 100 insanity pleas) whereas the actual rate of success was 26 percent. Thus, for every 1,000 felony cases, the public estimated 370 insanity pleas, 163 (44%) of which were estimated to be successful when, in actuality, there would only be 9 insanity pleas, 2 (26%) of which would be successful, representing a public estimate of success that was 81 times greater than the actual success rate.

Silver and colleagues (1994) reported public estimates of the proportion of insanity acquittees hospitalized to be about 50 percent, whereas the actual rate of hospitalization was approximately 85 percent. Similarly, these authors reported that the public overestimated the proportion of insanity acquittees that "go free" upon acquittal, with a public estimate of approximately 26 percent going free compared to the actual rate of about 15 percent. These authors note further that if conditional release and outpatient treatment are excluded from the definition of "going free," then the actual rate drops to a mere 1 percent. When these authors compared public estimates of the length of confinement of insanity acquittees to actual lengths of confinement they found, again, that public perceptions represent an underestimate of the actual length of confinement with the public estimating an average length of confinement of 21.8 months versus the actual average length of confinement of 32.5 months.

In Canada the insanity defense (Not Criminally Responsible on Account of Mental Disorder) is rarely successful (Livingston et al., 2003). Research examining the impact of Bill C-30, introduced in 1992 to change the provisions for the evaluation and treatment of defendants raising the issue of insanity or competency in Canada (and thought to make the insanity plea more attractive to defendants because of the emphasis on the community management of these individuals as well as proposed caps on the length of confinement for acquittees) appears to indicate that while an increase in the number of defendants raising the insanity issue has occurred, rates of acquittal have remained relatively constant (Arboleda-Florez et al., 2000). It appears that the insanity defense is used even more rarely outside North America. In Ireland, for example, Gibbons and colleagues reported an average of one insanity acquittal per year since 1910 and indicate that this defense is used only for the most serious of offenses (95 percent of acquittees had been charged with violent offenses; 72 percent

of which were infanticide or homicide) and is successful in only about 2 to 3 percent of homicides (Gibbons et al., 1997).

Characteristics of Insanity Acquittees

Although there are always exceptions to any rule, a relatively consistent picture of the typical insanity acquittee has emerged. Research examining the demographic characteristics of insanity acquittees in the United States (Cirincione et al., 1995) and Canada (Roesch et al., 1997) indicates that the typical insanity acquittee is male, between the ages of 20 and 29, single, unemployed, minimally educated, diagnosed with a major mental illness, has had prior contact with the criminal justice and mental health systems, and is acquitted for a violent offense (see also Lymburner and Roesch, 1999).

In terms of the types of mental illness that insanity acquittees suffer from, research has consistently demonstrated that the majority of insanity acquittees are diagnosed with psychotic disorders. In the United States, Wack (1993) reported that 62 percent of insanity acquittees in New York were diagnosed with a psychotic disorder whereas Bloom and Williams (1994) reported that 60 percent of insanity acquittees in Oregon were diagnosed with schizophrenia (a major mental disorder with prominent psychotic features). The situation is similar in Canada with Roesch and colleagues (1997) reporting that over 50 percent of insanity acquittees in British Columbia and Hodgins (1993) reporting that 63 percent of insanity acquittees in Quebec were diagnosed with schizophrenia.

In a survey of insanity acquittees in Ireland between 1850 and 1995, Gibbons and colleagues (1997) reported a higher proportion of females than has been seen in Canada or the United States. In addition, it appears that the insanity defense is used more often in Ireland for defendants with a primary diagnosis of personality disorder, encompassing 19.1 percent of all acquittees in Ireland as compared to approximately 10 percent of acquittees in the United States (and perhaps reflecting the higher proportion of females using the defense in Ireland). Similar to the situation in the United States and Canada, Gibbons and colleagues reported that the majority of insanity acquittees in Ireland had previous criminal and mental health histories.

Problems of Application within the Courtroom

Due to restrictions involving access to actual juries and jurors, most empirical investigations of jury decision-making rely on the use of mock juries and jurors. Although there are limitations to the ecological validity of such research, a relatively consistent picture of jury decision-making in insanity cases has emerged (see Zapf et al., 2006 for a review). In contrast to historical criticism leveled at the way juries make decisions (Finkel, 1988), the bulk of the research suggests that jurors make decisions conscientiously and rationally (Finkel and Handel, 1989; see also Zapf et al., 2006). However, while the legal system assumes that jurors apply the law

in strict accordance with the evidence presented, research has shown that jurors' attitudes toward the insanity defense, as well as their implicit theories of insanity, have an effect on the inferences they make about defendants' volitional and cognitive impairments.

Judicial instruction and juror attitudes toward insanity The language of insanity defense standards varies across jurisdictions. Jurors are expected to reach a verdict on insanity by carefully applying the relevant jurisdiction's legal standard to the evidence presented in the case. Research has shown, however, that differences in the language of the standard may have little effect on the actual decision-making process (Finkel, 2000; Finkel and Duff, 1989; Finkel and Handel, 1988; Finkel et al., 1985). Several studies have demonstrated that when jury instructions are manipulated, mock juror verdict patterns are indistinguishable across insanity standards. In fact, verdict patterns among jurors given no specific instruction at all were indistinguishable from those of jurors given specific instruction (Finkel, 2000; Finkel and Handel, 1988; Ogloff, 1991). This evidence does not suggest that jurors nullify judicial instruction, but rather that they rely on their own conceptualization of insanity where judicial instructions are vague or unclear regarding the definition of insanity. Therefore, verdicts may depend in large part on jurors' a priori attitudes and conceptions rather than on the legal standard as applied to the evidence. Research by Ellsworth et al. (1984) demonstrated that juror attitudes were more closely related to the verdict than to the facts presented in the case. This is troubling given that a majority of Americans appear to harbor negative misconceptions about the defense (Hans, 1986; Melton et al., 2007).

Public opinion polls suggest that a majority of Americans, while supporting the logic behind the defense of insanity, do not believe that insanity should be allowed as a complete defense (see for example, Roberts et al., 1987). As described earlier in this chapter, many people believe that the defense is used often and is usually successful (Melton et al., 2007). Media portrayals of the insanity defense likely contribute to such myths. Whether due to dramatization of the defense in television and movies, or to the disproportionate news coverage of highly publicized cases in which someone charged with a heinous crime uses the defense, the portrayal of the defense by the media would suggest that it is raised often when crimes are heinous and that it is easily malingered. Such views, when held by jurors, appear to be relatively inflexible. Jeffrey and Pasewark (1984) reported that approximately 50 percent of participants maintained their opinions that the defense was abused and over used even when presented with factual statistics to the contrary. Furthermore, opinion regarding strict liability—the view that mental state has little bearing on one's blameworthiness—has been shown to discriminate between jurors who find a defendant NGRI and jurors who find the same defendant GBMI (a verdict not significantly different from guilty) and to contribute to predicting verdicts after accounting for variation in the facts of the case (Roberts et al., 1987; Roberts and Golding, 1991; Skeem et al., 2004). Despite this evidence suggesting that there is a biasing effect of negative attitudes on

insanity verdicts, *voir dire* procedures—preliminary questioning of potential jurors to select those that will serve on a case—typically do not include inquiry into such biases; furthermore, there are cases in which judges have allowed the empanelment of jurors who have voluntarily expressed such negative views (Perlin, 1994).

Another misconception held by jurors that may impact a verdict is that defendants found insane are somehow "off the hook" upon acquittal. In fact, most states require initial commitment of acquittees and release is usually dependent upon judicial review (Melton et al., 2007). Moreover, most states do not restrict the length of commitment as long as the basis for commitment remains. U.S. courts have been inconsistent as to whether defendants have a right to judicial instruction to inform the jury about the dispositional consequences of an insanity verdict. In *Shannon v. United States* (1994) the Supreme Court ruled that dispositional instructions were not constitutionally necessary unless an error was made during the trial that would give jurors incorrect information regarding disposition. The premise for this decision was that jurors ought to decide guilt or innocence based only on the facts of the case and not on the consequences of their verdict. In a dissenting opinion, however, Justice Stevens noted that there is good reason to make juries aware of dispositional consequences. Research has shown that knowledge of dispositional consequences has a bearing on verdicts. Wheatman and Shaffer (2001) demonstrated that while judicial instruction regarding disposition had little effect on individual jurors (jurors who were not given the opportunity to deliberate), jurors given the opportunity to deliberate demonstrated a post-deliberation shift away from a harsher verdict as compared to uninstructed jurors. Content analysis of jury deliberations indicated that instructed juries understood that the defendant would be retained and treated whereas uninstructed juries worried that the defendant would be freed. Sloat and Frierson (2005) also found that jurors consider dispositions (accurate or inaccurate) in their verdict decisions, even if judicially instructed against doing so (see also Whittemore and Ogloff, 1995). Such findings cast doubt on the commonsense notion of withholding judicial instruction on dispositional consequences (Golding et al., 1999).

Jurors' implicit theories about insanity In addition to juror attitudes and misconceptions about the legal aspects of the defense itself, jurors' implicit theories about what constitutes insanity have been found to be strongly associated with verdicts—sometimes more strongly than the objective elements of the case (e.g., Finkel and Groscup, 1997; Finkel and Handel, 1989; Skeem and Golding, 2001). While the nature of the process by which jurors' implicit theories affect their verdicts remains unclear, several studies have shown that jurors may categorize defendants according to prototypes that match their implicit theories of insanity defendants. Finkel and Groscup (1997) found that undergraduates construe insanity defendants as young, with stressful histories of mental disorder, violence, and abuse, and who commit their crimes after an emotional precipitating incident. Participants associated grandiose delusions with success and motives of revenge with failure of the insanity defense. Likewise, Skeem and Golding (2001) found that juror prototypes are

systematically related to verdict. Only about 25 percent of jurors held prototypes that reflected the defendant as having legally relevant impairment at the time of the offense. Furthermore, Skeem and Golding found that juror prototypes appeared to be linked to attitudes toward the insanity defense, such that jurors with the legally relevant prototypes were more inclined to believe in the logic underpinning the insanity defense than were jurors who held prototypes of chronic, severe, and uncontrollable mental illness, or who conflated symptoms of psychosis with psychopathy.

Jury decision-making, however, is not based solely on subjective factors. If juror bias creates a tendency for individual jurors to decide all cases in a similar manner, then mock jury simulations should result in consistent verdict patterns regardless of case facts. This does not necessarily occur (Finkel et al., 1985; Finkel and Handel, 1989). Objective case characteristics have an impact on juror verdicts even as jurors construe case facts according to their implicit prototypes of insanity. Jurors appear to make inferences regarding the defendant's cognitive and volitional impairments on the basis of their implicit theories; however, juror construal of case evidence appears consistent and rational. Thus, the reasoning underlying verdict decisions is different across verdicts for individual jurors but related to evidence presented in the case (Finkel and Handel, 1989). Furthermore, Roberts et al. (1987) demonstrated that severity of mental disorder and level of premeditation have an impact on verdict beyond the attitudes and implicit theoretical constructs held by jurors.

Recommendations for Policy and Procedural Solutions

With extant research indicating that juror attitudes toward the insanity defense have a more biasing effect on verdict than case facts and judicial instruction, it seems prudent to establish procedural safeguards to ensure that jurors are applying the law appropriately. Addressing juror biases during *voir dire* and instructing jurors as to the dispositional consequences of their verdicts are two mechanisms whereby juror prejudices and misconceptions can be countered. Except in cases where *voir dire* procedures are narrowed by statute, judges are generally afforded substantial discretion in this process. Skeem et al. (2004) argued that such attention to jury bias during *voir dire* is necessary and noted that at least one instrument has been developed and psychometrically validated for such purposes. While it can be argued that judicial instruction regarding dispositional consequences is a legislative issue, the U.S. Supreme Court has neither deemed such instruction constitutionally necessary nor proscribed it. In light of evidence suggesting that juror perceptions of dispositional consequences factor into their decision-making about insanity verdicts, judges should be encouraged to provide such instruction.

Finally, the media plays an important role in shaping the perception of the insanity defense in the mind of the public. While there is little of legal consequence that can be done to move the media away from disproportionate coverage of high-profile cases or dramatization of the criminally insane, psychologists, psychiatrists and other mental health professionals working in the field can help to educate the

lay public and potential jurors. Advancing scientific evidence on the nature of insanity and factual information about the use and success of the defense may assist in dispelling widespread myths about the defense. Principally, as expert witnesses, mental health professionals can inform the courts about two types of issues: research on the damaging effects of bias, and clinical information regarding the nature and types of mental illnesses that can interfere with an individual's ability to understand the nature and quality of his acts.

Competency to Stand Trial

Postponement of criminal proceedings for those defendants considered incompetent to stand trial has long been a part of legal due process that can be traced back to at least the seventeenth century (Winick, 1983). 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 and Tulis, 1977, p. 1053). The rationale underlying competency doctrine includes: 1) protecting the accuracy of the proceedings by ensuring that the defendant is able to give appropriate assistance; 2) protecting the right of the defendant to due process by allowing the defendant the opportunity to choose and assist legal counsel, confront accusers, and testify on their own behalf; and 3) protecting the dignity and integrity of the proceedings. Bonnie (1992) explained that the dignity, reliability, and autonomy of the proceedings are protected by allowing only those that are competent to proceed.

Competency Procedures

The issue of a criminal defendant's competency may be raised at any point in the proceedings before a verdict is rendered. The U.S. Supreme Court, in *Pate v. Robinson* (1966), held that the issue of competency must be raised if evidence presented by the prosecution, defense, or obtained by the court raises a *bona fide* doubt about a defendant's competency. In addition, the Court in *Drope v. Missouri* (1975) further clarified that a defendant's irrational behavior or demeanor at trial are relevant to determining whether further inquiry on the issue of competency is warranted. Thus, the issue of competency can be raised by any party to the proceedings and must be formally considered if good faith doubt exists about a defendant's competency.

Once the issue of a defendant's competency has been raised, a mental health professional (one or more depending upon the procedures of the relevant jurisdiction) is generally called upon to evaluate the defendant. Unlike the case of insanity, where the defendant's mental state at the time of the crime is under consideration, the scope of consideration for competency is limited to the defendant's present mental state. This evaluation may take place at the jail, at an outpatient facility, or in an institutional setting and the factors considered by such an evaluation usually vary by jurisdiction.

Once the competency evaluation has been completed and a written report submitted to the court, a hearing on the issue of competency may take place. In many cases the prosecution and the defense will stipulate to the report and a hearing on the issue becomes unnecessary. The ultimate decision regarding a defendant's competency rests with the court, which is not bound by the opinion of the evaluator(s); however, in most instances the court accepts the recommendation of the evaluator(s) (Cox and Zapf, 2004; Hart and Hare, 1992; Zapf et al., 2004).

Competent defendants proceed whereas those deemed incompetent have their trials postponed until competency has been restored or their charges have been dismissed. Incompetent defendants are often sent to inpatient facilities for competency restoration. Until the case of *Jackson v. Indiana* (1972), virtually all states allowed for the automatic and indefinite commitment of incompetent defendants. In *Jackson*, the U.S. Supreme Court held that a defendant committed solely on the basis of competency "cannot be held more than the reasonable period of time necessary to determine whether there is substantial probability that he will attain that capacity in the foreseeable future" (p. 738).

Psychotropic medication is the most common form of treatment for competency restoration, although some jurisdictions have established treatment programs designed to increase a defendant's understanding of the legal process or that confront problems that hinder a defendant's ability to participate in their defense (see, for example, Bertman et al., 2003; Pendleton, 1980; Siegel and Elwork, 1990; Webster et al., 1985). With respect to the issue of whether an incompetent individual can be forcibly medicated to restore competency, the U.S. Supreme Court, in *Sell v. United States* (2003), held that antipsychotic drugs could be administered against the 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 involuntary treatment with medication could be justified on other grounds, including dangerousness (see *Washington* v. *Harper*, 1990). Writing for the majority, Justice Breyer identified key factors that a court must consider in determining whether a defendant can be forcibly medicated. Specifically, the court must determine that treatment with medication is medically necessary, is substantially likely to restore competence, but will not result in side effects that might affect a defendant's ability to assist counsel, and that alternative and less intrusive methods that would achieve the same result are not available. The *Sell* decision may serve to limit the use of medication as a treatment option for some incompetent defendants who refuse voluntary treatment and may result in greater emphasis on the development of various treatment alternatives.

Standards for Competency to Stand Trial

The modern standard for competency in U.S. law was established in *Dusky v. United States* (1960). Although the exact wording varies, all states use a variant of the

Dusky standard to define competency (Favole, 1983). In *Dusky*, the Supreme Court held that:

> 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)

Legal precedent, research, and commentary have delineated three abilities relevant to competency to stand trial—understanding, appreciation, and reasoning (see, for example, Bonnie, 1993; Poythress et al., 1999; Winick, 1987). Thus, evaluations of a defendant's competency to stand trial focus on the defendant's ability to understand the criminal process, both rationally as well as factually, appreciate his or her role as a defendant in that process, and reason and make decisions about various issues involved in his or her participation in that process. Decision-making regarding the issues involved in entering a guilty plea or waiving the right to the assistance of counsel appears to be subsumed under the issue of a defendant's competency to stand trial since the U.S. Supreme Court's decision in *Godinez v. Moran* (1993), which specified that the same standard for competence applies to pleading guilty, waiving counsel, and standing trial.

In countries such as England, Wales, Australia, New Zealand, and Canada competency doctrine—known as fitness to stand trial, fitness to proceed, or fitness to plead—exists and is defined in much the same way as it is in the United States. In Canada, for example, the Criminal Code (of which there is one that specifies the laws for the entire country as opposed to the case in the United States where each state has its own criminal code) gives the following definition of *unfit to stand trial*:

> Unable on account of mental disorder to conduct a defence at any stage of the proceedings before a verdict is rendered or to instruct counsel to do so, and, in particular, unable on account of mental disorder to (a) understand the nature or object of the proceedings, (b) understand the possible consequences of the proceedings, or (c) communicate with counsel. (Criminal Code of Canada, Section 2, 1992)

Although the legal systems of each of these countries (Canada, United States, Australia, New Zealand, and Wales) are all grounded in English common law, doctrine with regard to competency or fitness has evolved slightly differently for each country. For example, comparisons between Canada and the United States reveal that competency standards in the U.S. require a defendant to have both factual as well as rational understanding whereas in Canada the standard has been interpreted as requiring only factual understanding. Thus, in Canada, the courts have ruled that the test to be used in determining fitness is one of "limited cognitive capacity" (see *R. v.*

Taylor, 1992). In addition, there is no use of the word "appreciate" in the Canadian standard for fitness and, as such, the standard has been interpreted as requiring only an understanding of the nature and object of the proceedings as well as the possible consequences of the proceedings but not an appreciation of the personal relevance of the proceedings as is the case in the United States. Thus, it appears that the Canadian standards for fitness are narrower in scope than U.S. standards, taking into account fewer abilities, and resulting in a lower standard for being found fit to stand trial (Zapf and Roesch, 2001a; 2001b).

Competency in the Courtroom

Competency to stand trial is the most common type of forensic evaluation conducted. It has been estimated that the issue of competency to stand trial is raised for between 2 and 8 percent of all felony defendants in the United States (Bonnie, 1992; Hoge et al., 1992) with approximately 60,000 competency evaluations conducted in the United States annually (Bonnie and Grisso, 2000). Of those who are evaluated with respect to competency, approximately 20 to 30 percent are found incompetent to stand trial; however, the rates of incompetence vary widely across jurisdictions and evaluations settings (from about 7 to 60 percent; Melton, et al., 2007; Nicholson and Kugler, 1991).

No data are available on the rates at which hearings on the issue of competency are held; however, research has indicated that the issue of competency is rarely contested and the courts agree with the evaluator's opinion in upwards of 95 percent of cases (Cox and Zapf, 2004; Cruise and Rogers, 1998; Freckleton, 1996; Hart and Hare, 1992; Reich and Tookey, 1986; Zapf et al., 2004). Therefore, it appears that hearings on this issue are relatively rare, especially when one considers the large number of competency evaluations conducted annually. In addition, it appears that the hearings that do occur are held in front of a judge as opposed to a jury (although some states allow for the issue of competency to be heard by a jury).

Characteristics of Competency Referrals

Defendants referred for competency evaluations within the United States and Canada are most commonly male, single, unemployed, living alone, have a history of contact with both the criminal justice and mental health systems, and are diagnosed with a major mental disorder (see, for example, Nicholson and Kugler, 1991; Roesch and Golding, 1980; Zapf and Roesch, 1998). Nicholson and Kugler (1991) conducted a meta-analysis of research comparing competent and incompetent defendants and found poor performance on psychological tests measuring legally relevant functional capacities, a diagnosis of psychosis, and psychiatric symptoms indicative of severe psychopathology to be the strongest correlates with incompetency. Warren and colleagues (1991) found diagnoses of schizophrenia, mental retardation, mood disorders, and organic brain disorders all to be strong predictors of incompetence.

A direct comparison of competent and incompetent defendants by Hubbard and colleagues (2003) revealed that incompetent defendants were significantly more likely to be single, unemployed, charged with a minor offense, and diagnosed with a psychotic disorder and significantly less likely to be charged with a violent crime and to have substance use disorders than were competent defendants.

Problems of Application within the Courtroom

Although the concept of a competency is seemingly straightforward, in actuality the courts face a number of problems in evaluating and applying psychological information to legal standards of trial competence. The contextual nature of the competency evaluation and the balance between the needs of the state and those of the defendant are two areas in which the courts face dilemma.

Contextual nature of the evaluation Competency evaluations are intended to address the issue of the defendant's relevant abilities, as determined by the standard of competency set out for the particular jurisdiction. As such, evaluations should necessarily include an assessment of the defendant's ability to understand and appreciate his or her criminal charges. In evaluating this, however, there is the possibility that the defendant will relay self-incriminating information during the evaluation, thereby violating his or her Fifth Amendment right to "... be compelled in any criminal cases to be a witness against himself …." This issue was addressed directly in *Estelle v. Smith* (1981), in which the U.S. Supreme Court held that information acquired during a court-ordered competency evaluation cannot be used by the prosecution at either the guilt or sentencing phases of the trial unless the defendant brings his mental state into evidence (such as by pursuing an insanity defense). Given the potential for prosecutorial misconduct, however, evaluators should be cautious that they do not inadvertently include incriminating information in the report to court.

In addition to an assessment of the defendant's understanding and appreciation of his or her charges, the evaluation of competency should be contextual in nature; that is, tailored to the specific defendant and the specific requirements of his or her case. Thus, the threshold for competence for a defendant whose desire is to plead guilty to a relatively minor charge may very well be lower than the threshold for a defendant who has multiple charges and who will be required to take the stand in his or her own defense, listen to multiple prosecution witnesses, and assist his or her attorney in determining inaccuracies or inconsistencies in the various pieces of evidence. Thus, it is important that evaluators be knowledgeable about what is expected of a particular defendant for participation in his or her own defense to make an accurate assessment of the defendant's relevant abilities and deficits. The issue then becomes one of documenting the evaluation procedure and relevant findings for the court so that the decision-maker (i.e., the judge or the jury) may make a determination regarding the defendant's competency to stand trial. A causal connection between

any noted deficits and the defendant's mental illness needs to be established for the defendant to be found incompetent. The process of educating the courts with respect to the linkage between mental illness and competence-related deficits as well as how this may change with a change in context (e.g., for a defendant who is required to testify versus the same defendant who is not required to testify) is the challenge of the competency evaluator.

Research examining the quality of competency evaluation reports indicates that there is some cause for concern regarding the reliability of these evaluations. Skeem and colleagues (1998) demonstrated that although examiner agreement regarding the global issue of competency was high, agreement on specific competence-related deficits averaged only 25 percent across a series of competency domains. Of course, high levels of reliability on the issue of competency do not ensure that valid decisions are being made. Given that individuals who are found incompetent do not proceed with their trials before being restored to competency, there is no way of determining the validity of competency decisions. Roesch and Golding (1980) proposed that incompetent defendants be allowed to proceed with a provisional trial to allow for a determination regarding the validity of competency decisions. In this way, defendants who were able to proceed adequately would have the benefit of timely resolution of their case whereas those who were not able to participate adequately would have their cases set aside until they were treated and restored to competency. Provisional trials, however, do not occur in any state.

Balancing the needs of the state and the best interests of the defendant Another problem that the courts face is that of attempting to balance the needs of the state with those of the defendant. A competency evaluation delays the trial process; therefore, the state's primary interest in adjudication is held at bay until the defendant's competence is evaluated. Similarly, it may be the case that the defendant's right to a speedy trial is also held in abeyance by the defendant being found incompetent to stand trial. This is particularly troublesome in cases where the offense is relatively minor, may not carry any period of incarceration, and the defendant is willing to plead guilty. Given that the issue of competency must be raised by any party if bona fide doubt as to the defendant's abilities exists, this issue of balancing the needs of the state and those of the defendant becomes particularly troublesome when a defendant is prohibited from proceeding with a guilty plea (to a relatively minor charge) when he or she has proper assistance of counsel who has not raised the issue of competency.

Recommendations for Policy and Procedural Solutions

Two possible policy/procedural solutions have been proposed as a remedy to the problem of the uncertain validity of competency opinions and decisions. The first is the previously discussed idea of allowing defendants to proceed with a provisional trial. This would allow those defendants whose deficits did not interfere with their

adequate participation in their proceedings the opportunity for timely resolution of their cases. In addition, other defendants whose deficits did interfere with their adequate participation would then have their cases postponed until they were restored to competency. Thus, postponement of proceedings would only occur for those whose deficits actually interfered with their participation.

The second solution to the competency problem was proposed by Bonnie (1992; 1993) and involves a reformulation of adjudicative competency. Bonnie (1992) argued that competency doctrine evolved at a time when defendants were expected to represent themselves and did not have the assistance of counsel; however, procedures could now be modified to take into consideration the fact that every defendant has the right to assistance of counsel. The crux of Bonnie's (1992) proposal is that all defendants must have certain basic abilities in order to proceed including "(i) capacity to understand the charges, the purpose of the criminal process and the adversary system, especially the role of the defense counsel; (ii) capacity to appreciate one's situation as a defendant in the criminal prosecution; and (iii) ability to recognize and relate pertinent information to counsel concerning the facts of the case" (p. 297). However, the inability of a defendant to make certain decisions regarding his or her case should not necessarily be a bar to adjudication. Rather, the defense counsel within the context of his or her fiduciary relationship with the defendant could make certain decisions on behalf of the defendant so that he or she might proceed without delay.

Although both of these changes in policy and procedure would serve to allow the majority of defendants with competence-related deficits to proceed with their cases in a manner that retains the dignity, reliability, and autonomy of the legal process, neither of these proposals have been tested or implemented in any jurisdiction.

Conclusions

This brief review of some of the issues pertaining to competency and criminal responsibility serves to highlight the need for responsible mental health professionals to conduct high quality evaluations and to educate the court continually with respect to these issues. Continued high quality research on these important issues is necessary as is the education of the public as well as the professionals who will be dealing directly with these issues in a courtroom setting (whether or not they actually appear before a judge).

As with much of the research on legal issues, ecological validity arises as an issue with respect to research relevant to competency and insanity, particularly the research that investigates judicial instruction and juror attitudes regarding insanity. The research that was cited in this chapter was selected because of the high quality of the methods used and the resulting increased potential for generalizability of results.

As is the case with many legal issues, proposals for reform of policy or procedures related to competency and insanity have been slow to be accepted by the courts. It appears prudent to continue with efforts to educate legal decision-makers through evaluation reports and testimony about the relevant research and commentary on issues of competency and insanity. In addition, mental health professionals should increase their attempts to reach this important legal audience by publishing the results of their research on relevant legal issues in periodicals, journals, and newsletters that are accessed by legal professionals on a regular basis rather than to continue to publish in sources targeted towards other mental health professionals.

References

American Law Institute. (1962). *Model Penal Code*. Philadelphia, PA: American Law Institute.

Arboleda-Florez, J., Nussbaum, D., Ohayon, M.M., and Roesch, R. (2000). Research on the impact of the 1992 Criminal Code amendments on remands for fitness to stand trial and criminal responsibility. In D. Eaves, J.R.P. Ogloff, and R. Roesch (eds), *Mental Disorders and the Criminal Code: Legal Background and Contemporary Perspectives*. Burnaby, BC: Mental Health, Law, and Policy Institute. pp. 101–60.

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, 572–92.

Beasley, W.R. (1983). An overview of Michigan's guilty but mentally ill verdict. *Michigan Bar Journal*, 62, 204–5; 215–17.

Bertman, L.J., Thompson, J.W., Jr., Waters, W.F., Estupinan-Kane, L., Martin, J.A., and 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.

Bloom, J.D., and Williams, M.H. (1994). *Management and Treatment of Insanity Acquittees: A Model for the 1990s*. Washington, DC: American Psychiatric Press.

Bonnie, R.J. (1983). The moral basis of the insanity defense. *American Bar Association Journal*, 69, 194–7.

Bonnie, R.J. (1992). The competence of criminal defendants: A theoretical reformulation. *Behavioral Sciences and the Law*, 10, 291–316.

Bonnie, R.J. (1993). The competence of criminal defendants: Beyond *Dusky* and *Drope*. *University of Miami Law Review*, 47, 539–601.

Bonnie, R.J., and Grisso, T. (2000). Adjudicative competence and youthful offenders. In T. Grisso and R.G. Schwartz (eds), *Youth on Trial: A Developmental Perspective on Criminal Justice*. Chicago, IL: University of Chicago Press.

Borum, R., and Fulero, S. (1999). Empirical research on the insanity defense and attempted reforms: Evidence towards informed policy. *Law and Human Behavior*, 23, 375–94.

Cirincione, C. (1996). Revisiting the insanity defense: Contested or consensus? *Bulletin of the American Academy of Psychiatry and the Law*, 24, 165–76.

Cirincione, C., Steadman, H.J., and McGreevy, M.A. (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).

Cox, M.L., and Zapf, P.A. (2004). An investigation of discrepancies between mental health professionals and the courts in decisions about competency. *Law and Psychology Review*, 28, 108–31.

Criminal Code of Canada (1992) R.S.C. c. C-46.

Cruise, K.R., and Rogers, R. (1998). An analysis of competency to stand trial: An integration of case law and clinical knowledge. *Behavioral Sciences and the Law*, 16, 35–50.

Drope v. Missouri, 420 U.S. 162 (1975).

Durham v. United States, 214 F. 2d 862 (D.C. Cir., 1954).

Dusky v. United States, 362 U.S. 402 (1960).

Ellsworth, P., Bukaty, R., Cowan, C., and Thompson, W. (1984). The death-qualified jury and the defense of insanity. *Law and Human Behavior*, 8, 81–93.

Estelle v. Smith, 451 U.S. 454 (1981).

Favole, R.J. (1983). Mental disability in the American criminal process: A four issue survey. In J. Monahan and H.J. Steadman (eds), *Mentally Disordered Offenders: Perspectives from Law and Social Science*. New York: Plenum. pp. 247–95.

Finkel, N.J. (1988). Maligning and misconstruing jurors' insanity verdicts: A rebuttal. *Forensic Reports*, 1, 97–124.

Finkel, N.J. (2000). Commonsense justice and jury instructions: Instructive and reciprocating connections. *Psychology, Public Policy, and Law*, 6, 591–628.

Finkel, N., and Duff, K. (1989). The insanity defense: Giving jurors a third option. *Forensic Reports*, 2, 235–63.

Finkel, N.J., and Groscup, J.L. (1997). Crime prototypes, objective vs. subjective culpability, and a commonsense balance. *Law and Human Behavior*, 21, 209–30.

Finkel, N.J., and Handel, S.F. (1988). Jurors and insanity: Do test instructions instruct? *Forensic Reports*, 1, 65–79.

Finkel, N.J., and Handel, S.F. (1989). How jurors construe "insanity". *Law and Human Behavior*, 13, 41–59.

Finkel, N., Shaw, R., Bercaw, S., and Koch, J. (1985). Insanity defenses: From the jurors' perspective. *Law and Psychology Review*, 9, 77–92.

Freckleton, I. (1996). Rationality and flexibility in assessment of fitness to stand trial. *International Journal of Law and Psychiatry*, 19, 39–59.

Gibbons, P., Mulryan, N., and O'Connor, A. (1997). Guilty but insane: The insanity defense in Ireland, 1850–1995. *British Journal of Psychiatry*, 170, 467–72.

Godinez v. Moran, 509 U.S. 389 (1993).

Golding, S.L., Skeem, J.L., Roesch, R., and Zapf, P.A. (1999). The assessment of criminal responsibility: Current controversies. In I.B. Weiner and A.K. Hess (eds), *Handbook of Forensic Psychology* (2nd ed.). New York: Wiley. pp. 379–408.

Grisso, T. (2003). *Evaluating Competencies: Forensic Assessment and Instruments* (2nd ed.). New York: Kluwer/Plenum.

Hans, V.P. (1986). An analysis of public attitudes toward the insanity defense. *Criminology*, 4, 393–415.

Hart, S.D., and Hare, R.D. (1992). Predicting fitness for trial: The relative power of demographic, criminal and clinical variables. *Forensic Reports*, 5, 53–65.

Hodgins, S. (1993). Mental health treatment services in Quebec for persons accused or convicted of criminal offenses. *International Journal of Law and Psychiatry*, 16, 179–94.

Hoge, S.K., Bonnie, R.J., Poythress, N., and 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, 385–94.

Hubbard, K.L., Zapf, P.A., and Ronan, K.A. (2003). Competency restoration: An examination of the differences between defendants predicted restorable and not restorable to competency. *Law and Human Behavior*, 27, 127–39.

Insanity Defense Reform Act of 1984, Pub. L. No. 98–473, secs. 401, 402, 20 (1984).

Jackson v. Indiana, 406 U.S. 715 (1972).

Janofsky, J.S., Dunn, M.H., Roskes, E.J., Briskin, J.K., and Lunstrum, R.M. (1996). Insanity defense pleas in Baltimore City: An analysis of outcome. *American Journal of Psychiatry*, 153, 1464–8.

Jeffrey, R., and Pasewark, R. (1984). Altering opinions about the insanity plea. *Journal of Psychiatry and Law*, 11, 29–40.

Livingston, J.D., Wilson, D., Tien, G., and Bond, L. (2003). A follow-up study of persons found Not Criminally Responsible on Account of Mental Disorder in Canada. *Canadian Journal of Psychiatry*, 48, 408–45.

Lymburner, J.A., and Roesch, R. (1999). The insanity defense: Five years of research (1993–1997). *International Journal of Law and Psychiatry*, 22, 213–40.

M'Naghten's Case, 8 Eng. Rep. 718 (1843).

Melton, G., Petrila, J., Poythress, N., Slobogin, C., Lyons, P., and Otto, R. (2007). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (3rd ed.). New York: Guilford.

Melville, J.D., and Naimark, D. (2002). Punishing the insane: The verdict of guilty but mentally ill. *Journal of the American Academy of Psychiatry and Law*, 30, 553–5.

Nicholson, R.A., and Kugler, K.E. (1991). Competent and incompetent criminal defendants: A quantitative review of comparative research. *Psychological Bulletin*, 109, 355–70.

Ogloff, J. (1991). A comparison of insanity defense standards on juror decision making. *Law and Human Behavior*, 15, 509–31.

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.A., and Seidenzahl, D. (1979). Opinions concerning the insanity plea and criminality among mental patients. *Bulletin of the American Academy of Psychiatry and the Law*, 7, 199–202.

Pate v. Robinson, 383 U.S. 375 (1966).

Pendleton, L. (1980). Treatment of persons found incompetent to stand trial. *American Journal of Psychiatry*, 137, 1098–100.

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.

Poythress, N., Nicholson, R., Otto, R.K., Edens, J.F., Bonnie, R.J., Monahan, J., and Hoge, S.K. (1999). *The MacArthur Competence Assessment Tool—Criminal Adjudication: Professional Manual*. Odessa, FL: Psychological Assessment Resources.

R. v. Taylor, 77 C.C.C. (3d) 551 (Ont. C.A.) (1992).

Reich, J.H., and Tookey, L. (1986). Disagreements between court and psychiatrist on competency to stand trial. *Clinical Psychiatry*, 47, 29–30.

Roberts, C., and Golding, S. (1991). The social construction of criminal responsibility and insanity. *Law and Human Behavior*, 15, 349–76.

Roberts, C., Golding, G., and Fincham, F. (1987). Implicit theories of criminal responsibility: Decision making and the insanity defense. *Law and Human Behavior*, 11, 207–32.

Roesch, R., and Golding, S.L. (1980). *Competency to Stand Trial*. Urbana, IL: University of Illinois Press.

Roesch, R., Ogloff, J.R.P., Hart, S.D., Dempster, R.J., Zapf, P.A., and Whittemore, K.E. (1997). The impact of Canadian Criminal Code changes on remands and assessments of fitness to stand trial and criminal responsibility in British Columbia. *Canadian Journal of Psychiatry*, 42, 509–14.

Rogers, R., and Shuman, D.W. (2000). *Conducting Insanity Evaluations*. New York: Guilford Press.

Sell v. United States, 282 F. 3d 560 (2003).

Shannon v. United States, 114 S. Ct. 2419 (1994).

Siegel, A.M., and Elwork, A. (1990). Treating incompetence to stand trial. *Law and Human Behavior*, 14, 57–65.

Silten, P.R., and Tullis, R. (1977). Mental competency in criminal proceedings. *Hastings Law Journal*, 28, 1053–74.

Silver, E., Cirincione, C., and Steadman, H.J. (1994). Demythologizing inaccurate perceptions of the insanity defense. *Law and Human Behavior*, 18, 63–70.

Skeem, J.L., and 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., Golding, S.L., Cohn, N., and Berge, G. (1998). Logic and reliability of evaluations of competence to stand trial. *Law and Human Behavior*, 22, 519–47.

Skeem, J.L., Louden, J.E., and Evans, J. (2004). Venirepersons's attitudes toward the insanity defense: Developing, refining, and validating a scale. *Law and Human Behavior*, 28, 623–48.

Sloat, L., and Frierson, R. (2005). Juror knowledge and attitudes regarding mental illness verdicts. *Journal of the American Academy of Psychiatry and the Law*, 33, 208–13.

United States v. Brawner, 471 F. 2d 969 (D.C. Cir., 1972).

Wack, R.C. (1993). Forensic treatment in the United States: A survey of selected forensic hospitals. Treatment services at the Kirby Forensic Psychiatric Center. *International Journal of Law and Psychiatry*, 16, 83–104.

Warren, J.L., Fitch, W.L., Dietz, P.E., and Rosenfeld, B.D. (1991). Criminal offense, psychiatric diagnosis, and psychological opinion: An analysis of 894 pretrial referrals. *Bulletin of the American Academy of Psychiatry and Law*, 20, 63–9.

Washington v. Harper, 494 U.S. 210 (1990).

Webster, C.D., Jenson, F.A.S., Stermac, L., Gardner, K., and Slomen, D. (1985). Psychoeducational programmes for forensic psychiatric patients. *Canadian Psychology*, 26, 50–53.

Wheatman, S.R., and 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, 167–83.

Whittemore, K., and 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.

Winick, B.J. (1983). Incompetency to stand trial: Developments in the law. In J. Monahan and H.J. Steadman (eds), *Mentally Disordered Offenders*. New York: Plenum. pp. 3–38.

Winick, B.J. (1987). Incompetency to stand trial: An assessment of costs and benefits, and a proposal for reform. *Rutgers Law Review*, 39, 243–87.

Zapf, P.A., Golding, S.L., and Roesch, R. (2006). Criminal responsibility and the insanity defense. In I.B. Weiner and A.K. Hess (eds), *Handbook of Forensic Psychology* (3rd ed.). New York: Wiley. pp. 332–63.

Zapf, P.A., Hubbard, K.L., Cooper, V.G., Wheeles, M.C., and 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., and Roesch, R. (1998). Fitness to stand trial: Characteristics of fitness remands since the 1992 Criminal Code amendments. *Canadian Journal of Psychiatry*, 43, 287–93.

Zapf, P.A., and Roesch, R. (2001a). 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., and Roesch, R. (2001b). A comparison of American and Canadian conceptualizations of competence to stand trial. In R. Roesch, R.R. Corrado, and R.J. Dempster (eds), *Psychology in the Courts: International Advances in Knowledge*. London: Routledge. pp. 121–32.

Zapf, P.A., and Roesch, R. (2006). Competency to stand trial: A guide for evaluators. In I.B. Weiner and A.K. Hess (eds), *Handbook of Forensic Psychology* (3rd ed.). New York: Wiley. pp. 305–31.

This page has been left blank intentionally

Chapter 5
Psychological Syndrome Evidence

M. Alexis Kennedy

Synopsis

This chapter will explore the sometimes uneasy intersection of syndrome evidence and legal constructs. Syndromes are a grouping or constellation of symptoms used to identify an underlying undesirable condition. Psychiatric or psychological syndrome evidence can be helpful in legal settings in different ways. For example, expert testimony on psychological syndromes can be used to explain the behavior of offenders or victims. Three areas of syndrome evidence will be considered with specific attention afforded to their scientific status and current legal use. First, the use of Posttraumatic Stress Disorder evidence will be examined. The potential misuse of this relatively well-established psychiatry construct will also be discussed. Second, the controversies surrounding Rape Trauma Syndrome evidence will be considered. Third, the controversy surrounding Parental Alienation Syndrome will be explored. This syndrome is being used in family law despite a lack of scientific evidence to support it. The chapter will conclude with potential limitations of expert psychiatric or psychological witnesses, and will offer suggestions for more appropriate use of syndrome evidence.

Psychological Syndrome Evidence

The term "syndrome evidence" often carries negative connotations in the legal community, bringing to mind battles of expert witnesses, extra long court cases, impatient judges, and additional expenses. This chapter will explore the sometimes uneasy relationship of syndrome evidence and legal constructs. Various roles for syndrome evidence in the courts will be explored. Syndromes that have been successfully admitted by courts will be outlined with particular attention paid to three of the more common though often controversial syndromes seen in court—Posttraumatic Stress Disorder, Rape Trauma Syndrome and civil law's Parental Alienation Syndrome. Psychological or psychiatric research support for

these syndromes will be considered and reviewed. Criticisms of syndrome evidence will also be examined. Finally, the potentially problematic role of expert witnesses and future uses of syndrome evidence will be discussed.

In medicine, a syndrome is simply a group or constellation of symptoms used to identify an underlying undesirable condition (e.g., Sudden Acute Respiratory Syndrome or Acquired Immune Deficiency Syndrome). Psychiatry has also grouped psychological symptoms into syndromes (e.g., Tourette Syndrome). The diagnostic criteria for Tourette's in the DSM-IV TR (APA, 2000) include a range of behaviors such as nonrhythmic motor movements (e.g., facial contractions or twirling when walking) or vocal tics (e.g., clicks or barks). These behaviors have no known physiological origins, and cannot be caused by substance abuse or general medical conditions.

Introducing syndrome evidence into a trial has become unpopular for a number of reasons. First, fitting social science evidence into legal constructs is not always simple. For example, not all psychiatric diagnoses are relevant to the legal definitions; a clinical diagnosis of narcissism would not necessarily be relevant for a legal claim of insanity. Narcissism is defined in the DSM-IV TR as a life-long personality disorder characterized by an exaggerated sense of self-importance, need for excessive admiration, and a lack of empathy for others. Personality disorders have not generally been viewed as sufficiently severe to meet the mental disorder or mental defect requirement of most insanity statutes. While a lack of empathy may make it easier to victimize others, having this diagnosis does not make it relevant to a legal construct like insanity which necessitates a serious mental defect that interferes with cognitive abilities.

Second, cognitive or behavioral limitations established through syndrome evidence are often controversial if they are perceived as "free passes" for criminal offenders or excuses for other parties to litigation. Traditionally, any excusing or absolving of a criminal defendant's responsibility has challenged the public's belief in and support for the criminal justice system. For example, in 1843, the British public was outraged when the delusional Daniel M'Naghten was confined to a mental health hospital rather than sent to prison for shooting the secretary to the Prime Minister. Although M'Naghten was held until his death in 1865, the public outcry forced the House of Lords to create a new set of rules delineating the insanity defense that are still widely used in the U.S. and other countries' legal systems.

Despite the public's suspicion of legal excuses, it is clear that the court system sometimes needs expert explanations for behaviors that fall outside of what would be considered normal behavior. British and North American legal systems have long held that it would be wrong to punish a defendant who is not responsible for their behavior. The courts, when considering excuses like the insanity defense, attempt to assess the defendant's *mens rea* (i.e., his or her guilty mind) to be confident that the defendant was in control or was aware of the impropriety of his or her actions. Courts do not want to criminalize behavior when the accused lacks moral culpability for his or her action.

Courts do not want to criminalize behavior that is justifiable either. A legal justification involves a well-recognized rational reason for intentionally performing what appears to be a criminal act. For example, self-defense is based on a justification defense. The act that the accused performed (i.e., intentionally killing someone) may be justified under certain circumstances. It would be justified in most jurisdictions if: a) the perpetrator reasonably believes that he or she is in immediate danger of serious bodily harm; and b) that the use of such force is necessary and proportional to the danger the perpetrator faces. For instance, a police officer shooting a suspect when that suspect brandishes a loaded gun would meet these criteria. One area of justification evidence that is well-established is expert testimony on the symptoms and behaviors of battered women (for a full review see Schuller and Jenkins, 2007). Dr Lenore Walker (1979) first described "battered woman syndrome" in the context of the cycle of violence seen in domestic abuse situations. She also proposed that learned helplessness may explain why women stay in violent situations (1984). Explaining these dynamics became relevant in assessing fatal violence between intimate partners, in contrast to the archetypal self-defense cases of one-time altercations, typically between men (Schuller and Jenkins, 2007). The utility of explaining the dynamics between intimate partners has been accepted by courts throughout the United States, Canada, Great Britain, Australia, and New Zealand (Raitt and Zeedyk, 2000; Schuller and Jenkins, 2007).

The role of syndrome evidence is most clearly entrenched in criminal liability proceedings where an actor's behavior could be excusable or justified. This is, however, only one role in which psychological evidence can be used to explain behavior. Not all expert evidence will be introduced for the same purpose. This chapter addresses some of the syndrome evidence that may be relevant to offenders and victims in criminal proceedings. Expert psychological evidence is often proffered in civil situation (e.g., family law) as well. This chapter does not explore the admissibility of expert evidence as such a discussion is beyond the scope of the chapter and is covered elsewhere in this book (see Chapter 1 in this volume). This is because admissibility decisions vary based on the legal arena (civil vs. criminal) or when during a trial the evidence is presented (e.g., evidence presented as a defense has a different admissibility standard than evidence presented in rebuttal).

The utility and relevance of expert evidence in the following three areas—Posttraumatic Stress Disorder, Rape Trauma Syndrome, and Parental Alienation Syndrome—will be considered in more detail. The first area considered, Posttraumatic Stress Disorder, is not always described as a syndrome since it is a well-established psychological disorder. It is, however, used by the legal system much like or as a substitute for syndrome evidence. The second syndrome area considered, Rape Trauma Syndrome, was chosen as it is widely used in courts but its use remains controversial. The third type of syndrome evidence, Parental Alienation Syndrome, will be considered because there appears to be little psychological research to support it use.

Posttraumatic Stress Disorder

Across different legal systems, judges struggle with having to make legal decisions about the admissibility of expert information (Malsch and Freckelton, 2005; Sutherland, 2006). It is clear that courts benefit from the assistance of experts who can provide vital information on specific topics that are beyond the knowledge of the lawyers and judges.

The first area of psychological syndrome evidence that will be considered is that of Posttraumatic Stress Disorder (PTSD). PTSD is presented as a type of syndrome evidence based on a disorder that is widely accepted in the medical and psychological communities.

Scientific Status

PTSD was first recognized by the American Psychiatric Association (APA) as a new diagnostic category in the Diagnostic and Statistical Manual (DSM-III) in 1980. This diagnosis was described as applicable when normal individuals exposed to traumatic experiences "outside the range of usual human experience" (p. 236) displayed long-term symptoms of distress. The original APA description implied that exposure to such traumas was unusual or rare. More recent descriptions of PTSD report a relatively high lifetime prevalence rate of 8 percent for the adult population in the United States (APA, 2000), and an even higher rate of exposure to life threatening trauma among the general populace. This prevalence rate includes a range of traumatic experiences including, among others, witnessing a crime, going to war, or being a victim of a non-sexual crime.

Not all people who experience traumatic incidents will develop PTSD. The current DSM-IV TR definition of PTSD applies to a person exposed to a traumatic event who: a) experienced or witnessed a traumatic event which might lead to actual or threatened injury; and b) responded with intense fear, helplessness, or horror (APA, 2000). According to this definition, ongoing and expected events can be considered traumatic (e.g., being in combat). Also, this definition recognizes the subjective component of a person witnessing trauma having to experience substantial fear or horror (Foa and Riggs, 1995).

PTSD is diagnosed through the occurrence of specific symptoms in three different areas for at least a month. These symptoms are grouped into three clusters: re-experiencing the traumatic event (e.g., intrusive recurring thoughts about the event, nightmares); avoidance behaviors (e.g., avoidance of stimuli reminiscent of the event) or responsiveness numbing (e.g., emotional detachment); and, symptoms of increased arousal (e.g., irritability, exaggerated startle response) (APA, 2000). PTSD can be specified in three ways: "acute" if symptoms have existed for less than three months and persisted for longer than a month; "chronic" if symptoms have existed for more than three months; and, "with delayed onset" if the symptoms appeared more than six months after the traumatic event.

There is considerable research establishing the reliability and validity of the measures used to assess PTSD (Boeschen et al., 1998). A large body of research now considers pretrauma risks to developing PSTD and even sub-types of the disorder (Young and Yehuda, 2006). PTSD's inclusion in the diagnostic manual implies professional community consensus on the existence of a disorder related to experiencing a variety of different traumatic events.

Legal Use of PTSD

Courts have been open to admitting evidence about PTSD in cases involving victims of violence. For example, PTSD evidence has been used to explain the behavior of battered women (Schuller and Jenkins, 2007). It has also been used to explain the behavior of victims of rape and this use of PTSD is discussed below. PTSD has been introduced in legal trials in a number of different countries, for example, it was applied in civil tort cases of "nervous shock" in England and Wales (Adamou and Hale, 2003).

Other uses of PTSD evidence have been more contentious. PTSD has been applied in controversial ways to excuse the violent behavior of war veterans. PTSD was first labeled as "the Vietnam Stress Syndrome" in the 1970s, and current legal theorists argue that it could be applied to new war veterans (Aprilakis, 2005). PTSD suffered by war veterans has been successfully used as a mental illness forming the basis for an insanity defense (Delgado, 1985). For example, in the case of *State v. Cocuzza* (1981), a violent attack on police officers was accepted as a reenactment of trauma when the accused claimed that he had mistaken the officers for combatants in a flashback. (For a more comprehensive discussion of the insanity defense, different standards and severity of symptoms required for these relatively rare successful insanity pleas please see Chapter 4 in this volume.)

Misuse of PTSD-induced insanity has led to serious skepticism about its existence (Aprilakis, 2005). For example, in *State v. Lockett* (1983), where the lower court had accepted the original claim of PTSD-induced insanity, the insanity plea was vacated when it was learned that Lockett had never fought in Vietnam as he had claimed. PTSD, like any type of social science evidence, can develop a negative reputation because of unpopular decisions like these.

A recent Canadian decision saw PTSD-induced insanity expanded to cover the sexual assault of a child (*R. v. Borsch*, 2006). This decision went beyond previous acceptance of PTSD as being an explanation for violent behavior such as attacking police officers. In the *Borsch* case, a veteran's diagnosis of PTSD became the basis for his successful insanity defense. The accused had apparently interrupted a sexual assault of a child while on a peace-keeping mission in Bosnia. He claimed that he had shot that offender and rescued the child. The accused argued that his current sexual assault was a reenactment of the sexual part of that event. The judge concluded that Borsch was insane at the time of his crime, and that he did not understand or appreciate the nature of the sexual assault that he was committing. This is a troubling

extension of PTSD. The notion that PTSD leads to black outs and memory loss for raping a child or any form of intimate violence is an unsupported extension of current research (Swihart et al., 1999). Considering how rarely insanity defenses are successful, there is some doubt that this case will survive its current appeal process.

Rape Trauma Syndrome

Rape Trauma Syndrome (RTS) is the second area of psychological syndrome evidence that will be considered. This psychological construct has sometimes been accepted in court proceedings as a type of evidence useful for explaining the behavior of victims of rape. In contrast to PTSD, this disorder does not enjoy widespread consensus as to its definition nor is it specifically included in the DSM-IV TR.

Scientific Status

RTS was first described in 1974 by Ann Burgess and Lynda Holmstrom when they compared the recovery patterns of 92 adult victims of forcible rape. Courts however, have used the term "rape trauma syndrome" to refer to Burgess and Holmstrom's model of recovery, to any and all post-rape symptomology, and to rape-related PTSD. These three constructs are not identical and have decidedly different applications for those involved in the legal process.

In the initial RTS research by Burgess and Holstrom (1974) women were interviewed immediately after a rape and then again one month later. The researchers identified two phases with different symptoms in each phase. The first phase, an acute crisis phase, contained reactions that lasted for a few days or weeks and tended to be fairly severe. Both physical symptoms (e.g., sleeplessness, loss of appetite, numbness or pain), and psychological symptoms (e.g., extreme fear, persistent nightmares, depression or suicidal thoughts) were described. Long-term reactions were described in the second phase called the "reorganizational" phase characterized by issues such as disturbances in general functioning, development of phobias, sexual problems, and lifestyles changes. The length of time associated with each phase varied among victims but eventually all victims moved into the latter stage.

Since that original conceptualization of RTS in 1974, considerable research has been conducted on survivors of sexual assault. Very little research, however, has attempted to replicate the original study in its entirety (Frazier and Borgida, 1992). The original syndrome conceptualization has been useful for therapeutic purposes but it has not created a psychiatric model supported by research (Frazier, 2005).

The current research on rape victims tends to examine specific symptoms. For example, research has reported on challenges for rape victims such as depression (e.g., Frazier, 1990; Gidycz and Koss, 1989; Hutchings and Dutton, 1997), suicidal behavior (for a full review see Ullman, 2004), long-term physical health problems (e.g., Golding, 1999; Ullman and Brecklin, 2003), and fear and anxiety (e.g., Frazier

and Schauben, 1994; Kilpatrick et al., 1985; Sudderth, 1998). Unfortunately, this research is often presented to the courts as RTS evidence, leading to confusion as this evidence may not be tied to the original RTS theories (Boeschen et al., 1998).

The current research on the reactions of victims of sexual assault, although not formally following the model of RTS, remains an important social science construct relevant to legal decision-making. Evidence introduced by an expert witness during a trial may be able to address seemingly counterintuitive behavior of a victim following an assault (e.g., delays in reporting, appearance of self-doubt in initial reports of the incident). For example, one of the most important areas of research in sexual assault is the issue of self-blame. The attributions a victim makes about a sexual assault also influences the perceptions and reactions of others, therefore, if a victim blames herself, others are likely to adopt this viewpoint, treating her as the person responsible (Frazier and Schauben, 1994). This self-blame may lead to the victim's failure to or delay in reporting the incident because the victim feels partially responsible for the incident occurring.

Overlap with PTSD

PTSD evidence is often introduced and described as RTS evidence in court proceedings. Using the two names interchangeably is an error. Therapists and medical experts now prefer to assess PTSD, rather than RTS, following sexual victimization. PTSD is a cleaner clinical construct, is specifically described in the DSM-IV TR, and possesses a substantial research base. Some people consider RTS to be a subtype of PTSD because of the overlap in symptoms, but it is important to remember that the two constructs developed separately. While rape is specified in the DSM-IV TR as a type of trauma that may induce PSTD, RTS is not in the DSM-IV TR. Further, some of the reactions common to rape victims (e.g., sexual dysfunction, depression, anger) are not criteria specified under PTSD. As RTS focuses on a specific situation (i.e., rape) it is unlikely to be included in future editions of the DSM. The diagnoses included in the DSM-IV tend to describe broad disorders that can be applied to a wide variety of situations.

Current psychological research indicates that the rates of development of PTSD are higher among victims of rape than for people experiencing other traumas (for a full review see Frazier, 2005). Foa and Riggs (1995), for example, found that 94 percent of rape victims met the symptom criteria for PTSD two weeks after an assault (remember, PTSD diagnoses require four weeks of symptoms as a duration requirement). They also reported that in a second sample tracked for a longer period, they found 90 percent of rape victims met the PTSD criteria two weeks following the assault, with 60 percent still meeting the criteria at one month and 51 percent at 12 weeks. These rates were higher at each point than the rates for a comparison group of (non-sexual) assault victims (62% at two weeks, 44% at one month and 21% at 12 weeks). Other psychological research on rape victims indicates that the majority

of such victims meet the criteria for PTSD one year following the assault (Frazier, 2005).

Legal Use of RTS

RTS evidence has been accepted in courts in the United States and to a limited extent in the United Kingdom (Raitt and Zeedyk, 2000). Expert testimony on RTS and PTSD has been allowed in courts considering rape cases for different reasons, including: 1) as information to support a request for damages in civil cases; 2) as a defense for unusual behavior by a rape victim; 3) as evidence of lack of consent; and 4) as an explanation for the behavior of an alleged victim following an assault (Block, 1990). Yet, expert testimony is not likely to be uniformly admissible for all these purposes. This discussion of RTS and PTSD will focus on the fourth purpose—to explain the behavior of the victim.

Due to the legal history of sexual crimes, jurors and judges in sexual assault prosecutions may be vulnerable to rape myth stereotypes or inaccurate assumptions (Frazier and Borgida, 1992; Schnopp-Wyatt, 2000). Throughout history, both legal and cultural representations (e.g., literature) stated that only certain women could be raped and only rape behavior directed toward young, attractive women was plausible (Lalumière et al., 2005). Then and now, women are portrayed as being responsible for protecting their sexual integrity and any claims about failures (e.g., non-consensual sexual activity), therefore, should be assessed with skepticism, particularly if the victims' behavior put them at risk or if the victims are not sympathetic. The common acceptance of these myths is important to address in legal trials as Boeschen and colleagues (1998) state, "These stereotypes and myths have led to a society that typically shifts its critical focus from the rapist to the victim" (p. 414).

Presenting expert witness testimony on RTS, research on consequences of sexual assault or PTSD symptomology can be offered to dispel these myths or to provide an explanation for seemingly counterintuitive behavior of the victim (e.g., a delay in reporting the crime). Evidence, however, cannot be introduced by the prosecution to prove that a specific rape did occur or that a sexual assault necessarily occurred because certain symptoms are present. Psychological trauma may be evidence of being a victim, but it does not necessarily prove who committed the assault or which of possibly many traumas caused specific symptoms.

When faced with admitting expert testimony, judges must balance a number of concerns. On the one hand, jurors with negative stereotypes about rape may benefit from learning how these myths may affect the way they are considering the victim's reactions. On the other hand, this information portraying the victim in a sympathetic light may potentially bias jurors against the defendant (Frazier and Borgida, 1992).

When courts refuse to allow expert evidence about rape trauma, they often are accepting defense lawyers' arguments that such information would be prejudicial to the defendant and potentially confusing. Perhaps the greatest barrier for such evidence is its name—Rape Trauma Syndrome. Courts rejecting RTS evidence have argued

that the combination of the word "rape" with medical terminology "syndrome" may lead the jury to believe that a scientific judgment has been made as to whether an alleged victim has been raped (Schnopp-Wyatt, 2000).

Critics also point out that the term RTS implies that there is a standard set of responses seen in rape victims. Similar to criticisms of the label for "battered woman syndrome", the notion of a "syndrome" oversimplifies the psychological realities into a narrow range of reactions or a single profile of behavior. Experts on battered woman syndrome have recommended that the term itself be dropped from the testimony (Dutton, 1993). Similarly judges are now saying that the RTS term should be avoided.

Expert evidence is most likely to be considered relevant when it is explaining behavior of a victim that appears to be inconsistent with being raped (Faigman et al., 2005). As a recent decision reveals, courts continue to accept the value of expert testimony on RTS as it,

> … may play a particularly useful role by disabusing the jury of some widely held misconceptions about rape and rape victims, so that it may evaluate the evidence free of the constraints of popular myths. (*Chapman v. State*, 2001, §11)

Not surprisingly, evidence to explain unexpected behavior from a victim which is consistent with traumatization is more readily admitted than evidence proffered to show that a victim meets the typical checklist of "normal" victim reactions.

Historically, RTS evidence was presented to bolster the victim's credibility (Costanzo, 2004). If the defendant admitted to sexual activity but claimed that it was consensual, evidence of trauma suffered by the victim increased the credibility of a claim that a rape occurred. Most courts now prohibit RTS evidence which potentially bolsters the credibility of the victim, stating it is the job of the jury to decide if the victim is credible or not (Faigman et al., 2005). Although judges may caution jurors that evidence should only be used to dispel rape myths, it is not possible to determine if jurists are using that information to bolster the victim's credibility. This potential misapplication has led to general criticism of RTS evidence being introduced at trial.

The debate on admissibility of RTS evidence tends to focus more on the purpose of the testimony than on the scientific validity of RTS. Legal evidentiary texts suggest that no U.S. cases have rigorously evaluated the validity of the research underlying RTS (Faigman et al., 2005). For example, Frazier and Borgida's (1992) research reported that courts accepting RTS evidence did not appear to actively scrutinize expert qualifications or the underlying scientific reliability to a significant degree (see Chapter 1 in this volume for a general discussion of the admissibility of psychological expert testimony).

Legal Use of PTSD as RTS Evidence in Rape Cases

Challenges to the admissibility of PTSD expert evidence have been made based upon the notion that PTSD was simply another name for RTS. In *State* v. *Martens*, the Ohio appellate court both acknowledged that RTS evidence is still admissible for certain purposes and went on to discuss the usefulness of the PTSD testimony,

> … we do not find the doctor's testimony to have exceeded permissible limits. The doctor testified that Gaerke suffered from PTSD caused by the incident with appellant. This testimony was beneficial because it eliminated other stressful events as possible causes of Gaerke's PTSD. This testimony also helped to explain why Gaerke waited approximately three months before reporting the rape and why others observed a change in her behavior in the months after the incident. Although the doctor's testimony about Gaerke's emotional demeanor not being feigned came close to commenting on Gaerke's veracity and bolstering her credibility, we find that this threshold was not crossed. (*State v. Martens*, 1993, p. 347)

This quotation may overstate the ability of experts to separate multiple sources of trauma. Multiple negative life experiences may prevent being able to distinguish one specific cause of PTSD symptoms from other possible causes (Boeschen et al., 1998).

Another problem with relying solely on PTSD diagnoses in rape cases is that such expert testimony does not often include the literature on symptoms of rape outside PTSD criteria (e.g., research on depression, suicidal ideation, long-term physical harm) (Boeschen et al., 1998; Faigman et al., 2005). Although research has not been conducted on Rape Trauma Syndrome *per se*, much research has been completed on the long-term negative psychological and physical harm caused by sexual victimization. For the most part, the literature suggests a good deal of scientific agreement on the negative consequences of sexual victimization (Frazier, 2005).

When PSTD evidence is being presented in rape trials, it should include empirical research specific to rape victims. Research on PTSD consistently finds that victims of rape are more likely to develop PTSD than other trauma sufferers and experience greater levels of PTSD symptomology (Frazier, 2005). Unfortunately, even a strong foundation of scientific evidence does not always translate into competent expert evidence. As Boeschen and colleagues (1998) warn:

> Although experts can provide important information when testifying, unsubstantiated, nonscientific testimony on PTSD and especially RTS can harm not only victims and alleged offenders, but also the field of psychology as a whole. (p. 415)

Parental Alienation Syndrome

Parental Alienation Syndrome (PAS) has become prominent in the civil legal arena, specifically in child custody disputes. It has been admitted as evidence in divorce proceedings throughout the United States and Canada. It differs from PTSD and RTS in that PAS has little empirical research to support its use.

Scientific Status

PAS was coined by Dr Richard Gardner in 1985, and is best described as a focused effort to instill negative views of another parent. Gardner defined PAS as a parental campaign of unjustified disparagement against another loving parent (1985; 2003b). The two key characteristics of PAS are: a) active programming or brainwashing of a child against the other parent; and b) that this brainwashing leads to active generation of negative opinions by the child involved in the dispute. PAS only exists when there is no apparent basis for this hostility. In situations where real parental abuse or neglect exists, negative feelings would not be considered symptoms of parental alienation. Although originally the diagnosis was given to the child, this syndrome is more specifically to pathologize the activities of a parent.

There are three levels of PAS: mild, moderate, and severe (Gardner, 2004). Mild PAS is categorized by relatively superficial alienation from a parent where children comply with visitation but complain about or criticize the victimized parent. Moderate PAS is characterized by disruptive and disrespectful behavior by the child to a parent, and is accompanied by a continual negative campaign of disparagement by the offending parent. Finally, in severe PAS, the child becomes so hostile towards the targeted parent that visitation may be unfeasible. Gardner warned that children experiencing severe PAS may act out violently towards the alienated parent.

Gardner proposed that PAS could be identified though evaluating a child's behavior across eight areas (Gardner, 2003b; 2004) including: a) negative acts or statements directed towards the victimized parent; b) criticisms of victimized parent based on absurd rationalizations; c) extreme, polarized emotions towards the parents; d) claims by the child that the rejections of the victimized parent are the child's own ideas; e) reflexive loyalty to the alienating parent; f) no remorse for cruel acts toward the victimized parent; g) imagined or rehearsed scenarios; and h) extension (application) of negative emotions to people associated with the victimized parent.

One unique feature of PAS is that even the leading expert acknowledges that this syndrome exists almost exclusively in cases of divorce (Gardner, 2003b). Other therapists describing this syndrome have argued that it may exist outside of divorce situations (Kopetski, 1998). Legal theorists have described PAS as a form of psychosocial pathology (Williams, 2001).

Nearly all of the research generated on PAS originates in anecdotal legal and therapeutic descriptions of the syndrome (e.g., Cartwright, 1993; Kopetski, 1998; Rand, 1997; Weigel and Donovan, 2006). There is a significant lack of empirical

research (Siegel and Langford, 1998). All of the peer reviewed articles describing Gardner's PAS, are descriptive and not empirical (1985; 1998; 2003b; 2004). There has been no reliability or validity research done on the use of Gardner's syndrome criteria (Rueda, 2004; Warshak, 2001; Williams, 2001).

Research on PAS has not uniformly employed Gardner's categories or criteria. Some research has modified the original symptoms. For example, in the research by Siegel and Langford (1998), an attempt was made to identify alienating parents using the MMPI-2 subscales focused on the parents' behavior not on Gardner's recommended assessment of the children's behavior. Gardner has argued that the PAS could be included in future DSM editions as there are at least 100 articles on the syndrome (2003b). However, inclusion is unlikely for a number of reasons. First, the articles he refers to are not empirical research studies. Second, the majority of publications describing PAS were authored by Gardner, were self-published, and were not peer reviewed (Kelly and Johnston, 2001). Third, it is also unclear how the 2003 death of Dr Gardner will influence the future of this syndrome because his articles are the most frequently cited authorities and he was the most frequent expert witness in the United States and Canada on this syndrome. Fourth, unlike other DSM-IV TR disorders, such as PTSD which applies to a diverse set of circumstances like rape, war, and accidents, PAS only applies to the narrow situation of divorce disputes.

PAS has little in common with research on other established psychiatric diagnoses like depression or anxiety. PAS, for example, does not overlap with PTSD because PAS can only exist in cases where there is no real abuse inflicted by the wrongly accused parent (Gardner, 2004). While Gardner (2003b) introduces PAS by comparing it to medical syndromes, unlike most medical conditions his construct is based on the behavior of two parties or on the family social dynamics not on a single individual's characteristics. In summary, there is little empirical evidence to provide a basis for scientific consensus or general acceptance of a syndrome related to parental alienation.

Current Legal Use of PAS

Despite the lack of scientific support for PAS, it is currently being used in legal cases around the world. Decisions in the United States, Canada, and Europe indicate that judges are making decisions based on expert testimony as to the existence of PAS in families undergoing custody disputes. According to Dr Gardner's website, PAS has been admitted in 22 U.S. states as well as cases in seven Canadian provinces, Australia, Germany, Great Britain, Israel, and Switzerland (2003a).

American courts, however, have not widely accepted PAS evidence. Although acknowledging that it is not a universally accepted syndrome, this evidence has still found its way into some legal decisions (*Pearson v. Pearson*, 2000). In most cases where PAS is introduced, judges often sidestep its admissibility and validity,

claiming that they do not need to come to a decision as to whether PAS exists to come to a decision in their case.

Despite the lack of scientific support for this syndrome, Canadian courts, in contrast, have admitted PAS evidence quite frequently. In a 2002 case in which Dr Gardner's initial testimony was excluded because he had not interviewed all of the parties involved, his expert opinion was allowed after observing the participants testifying during the trial with the following warning,

> His theories are not universally accepted, but they are beginning to realize a wide degree of acceptance particularly in child custody cases, that is, issues in the civil courts as opposed to the criminal courts. (*R. v. C. (K.)*, 2002, §32)

The judge in this case found that Dr. Gardner's testimony met the *Mohan* test for inclusion because the expert testimony was beyond the expertise of the trier of fact (see Chapter 1 in this volume for a discussion of the *Mohan* admissibility test). In another case where PAS was admitted, the judge noted that the parent had such a severe case that he was "not consciously aware of his brainwashing and has no insight into its consequences" (*Rothwell v. Kisko*, 1992, §3). The judge went on to order psychological treatment for "what has been diagnosed as your severe case of parental alienation syndrome" (*Rothwell v. Kisko*, 1992, §14).

The appeal court of Quebec has not yet reversed its decisions admitting expert testimony on PAS (see R. (B.) c. M. (R.), 1994). One party has, however, challenged the lower court's finding of PAS based on the fact that the expert in that case did not use Gardner's original conceptualization of PAS. They further argued that PAS was misdiagnosed and inappropriately applied because the alleged alienating behavior was not seen both in the parent's behavior and in the child's behavior. The appeal court judges clearly stated that both conditions need not exist for PAS to be present. As long as one parent displays ill will, the children do not have to actually denigrate or criticize the other parent for evidence of PAS to be admitted in Canadian courts. This is in direct contrast to Gardner's conceptualization of PAS where the primary assessment is based on the behavior of the child.

In summary, PAS is an example of a syndrome without substantial scientific support or widespread acceptance. Its definition and application in real cases is confused and varies considerably. Although it is proper for courts to want to identify behavior that harms the child, relying on a diagnosis of PAS is unsound.

Problems with Expert Witnesses

Problems with syndrome evidence are not limited to psychological syndrome issues but generalize to all expert testimony on scientific evidence in legal proceedings. Difficulties occur in a number of different areas. The first may be limitations in judges' familiarity with science. This is compounded by a second factor—the

complicated legal standards that govern expert testimony admissibility (see Chapter 1 in this volume for more detail on this issue).

The third problematic area for syndrome evidence is the nature of the experts chosen to present this information. The quality of expert witnesses can vary dramatically. Although the ideal expert witness is always neutral and objective, bias and partisanship may seep into legal reports and testimonies to differing degrees (Malsch and Freckelton, 2005). The most extreme examples of partisan experts are scientists who are purchasable. Malsch and Freckelton describe them as "venal—prepared to tailor their reports and their testimony overtly to assist those who hold the purse strings" (2005, p. 48). In the United States, such experts are often referred to as "hired guns." Lawyers in Canada sometimes refer to them as "jukebox" witnesses—just put a quarter in and they will play any tune that you want. Purchasable expert witnesses exist in both adversarial (e.g., U.K., U.S., Australian) and inquisitorial (e.g., Netherlands) legal systems (Malsch and Freckelton, 2005).

Bias and partisanship can also occur due to lack of training, and some authors argue that legal decision-making has suffered due to a lack of properly trained experts. The growing demand for witnesses forces ill-prepared and unsupervised non-forensic practitioners to testify at trials (Gudjonsson and Haward, 1998). This may contribute to partisanship or bias in two ways. First, experts may be inclined to provide testimony in high-profile legal cases. The excitement of being part of a "dream team" group of professionals working together may lead to experts enthusiastically taking sides. Second, experts may become too invested in their own psychological theories to assess them objectively (Sutherland, 2006). Tips on avoiding partisanship are outlined in the recent work by Saks and Lanyon (2007).

Another pitfall for syndrome evidence may be the undue deference to expert witness testimony afforded by judges and juries. As Sutherland points out, "somewhat paradoxically, it is this very ignorance of science that often results in non-scientists being mesmerized by it" (2006, p. 381). Our general faith in science gives the expert witnesses an inherent credibility. The "truths" in science are often accepted unquestioningly without recognition that science changes far more rapidly than constructs in the law change. When scientists themselves cannot agree on a current truth, it is even more difficult for non-scientists to weigh the testimony of one scientific discipline against another.

Further complicating the assessment of the utility of expert evidence are the debates within a scientific discipline. Psychological and psychiatric experts themselves do not have clear agreement on the validity and diagnostic reliability of syndromes currently being presented in court. This scientific debate clouds the ability of judges to determine what evidence should be admitted. Even the well-established disorders listed in the DSM-IV TR are also not necessarily diagnosed in the same manner by each clinician. There is little research demonstrating whether clinicians are more accurate at diagnosing these disorders (Morse, 1998). Clinicians interpreting the intentions, thoughts, and perceptions of another individual are undertaking a difficult task, unlike the task of other experts who measure the breaking point of a bolt on

a collapsed bridge or the blood alcohol level of a drunk driver. Furthermore, even seemingly measurable physical phenomena are still open to debate and differences in interpretation.

Another limitation of syndrome expert testimony is its failure to recognize the different perspectives of psychology and the law (Canter, 2007). The law attempts to assess each individual as a reason-driven person. Canter argues that the need for a reliable interpretation of one individual's behavior may be pressuring psychologists to inappropriately translate research conducted at the group level to individual decisions.

Recommendations for the Use of Syndrome Evidence

Clarify the "Syndrome" Label

Frances et al. (1995) contrast syndromes to diseases by stating that syndromes often lack a clear pathological nature and a less specified temporal course of development. The syndromes discussed in this chapter are often linked to a specific trauma (e.g., rape, divorce, being battered). However, the characterizations of syndromes like PAS or Battered Women's Syndrome are hampered by their political histories or partisan presentations (Dahir et al., 2005).

Despite the limitations of syndrome classifications, the necessity for introducing information on human behavior remains important in the criminal justice system. In one well documented syndrome area, Battered Women Syndrome, research has begun specifically to consider the use of the word "syndrome" on jurors' decision-making processes (Schuller and Jenkins, 2007). Research has failed to find a significant difference between testimony which is labeled as "syndrome" evidence, and the same testimony presented without using syndromal terms (Schuller and Hastings, 1996). Similar results have emerged with Rape Trauma Syndrome evidence (Schnopp-Wyatt, 2000). It is important to note that the information presented to the triers of fact that addressed specific myths or biases did change the decision-making process, regardless of whether it was labeled syndrome evidence (Schnopp-Wyatt, 2000). Simply avoiding the nomenclature of "syndromes", however, may not fully solve the limitations of the current use of social science evidence in the legal decision-making process.

Ensure that Expert Testimony is Questioned

To admit and use relevant social science evidence properly, lawyers and judges must become more science literate. For judges working in a system where lawyers will introduce contradictory expert witnesses, a familiarity with scientific theory will allow them to move beyond seeing science as simply good or bad. They can focus their scientific examination on how much weight to attribute different pieces of expert

evidence rather than whether it is wholly accurate or inaccurate. Ideally, judges should not rely on theories like Parental Alienation Syndrome while simultaneously acknowledging in their decisions that the theory lacks general acceptance (e.g., *R. v. C. (K.)*, 2002).

Lawyers also need to be more vigilant about limiting expert testimony to the expert's areas of expertise. Blind faith in scientists can lead to disasters like the recent misdiagnoses of Munchausen's Syndrome by Proxy in the United Kingdom. A prominent expert, pediatrician Sir Roy Meadow, has been vilified over his misdiagnoses of Sudden Infant Death cases as criminal instances of Munchausen by Proxy Syndrome (MBPS, Sutherland, 2006; Wilson, 2005). MBPS was a syndrome coined by Dr Meadows in the UK in 1977 (Miller, 2006) whereby parents injure their children to receive medical and emotional attention. MBPS was included in the DSM-III but the name was removed in the DSM-IV when the same diagnosis was described under factitious disorders by proxy instead. Despite its inclusion in the DSM-IV, not all courts accept the existence of this psychological disorder (Miller, 2006).

Lawyers and judges need to develop a simple understanding of different scientific approaches (Wilson, 2005). They will also benefit from abandoning the idea that there is one simple scientific method that will apply to all types of evidence and all scientific disciplines (Edmond and Mercer, 1998; Moreno, 2003). Psychological researchers and experts may have to acknowledge that it is a more subjective area of science than scientific opinions presented on blood alcohol levels or paternity results. Even these "hard sciences" are open to debate and controversy so the softer science (or "art") of psychology should be prepared for criticism and debate. Accepting, understanding, and assessing psychiatric and psychological expert information is more complex than simply distinguishing "good" science from "bad" science. An improved level of scientific sophistication will allow judges to provide guidance on how much weight expert evidence should be given.

In addition, more research is needed exploring how judges are using or misusing expert testimony regarding syndrome evidence (Dahir et al., 2005; Krafka, et al., 2002). Lawyers or courts may benefit from relying more on expert assistance in deciphering the most difficult medical literature. Moreno cites the example of child abuse cases where even lawyers who practice regularly in this area still do not understand the underlying research upon which their expert opinions are based (2003). The scientific measurement of intentions, emotions, and cognitions of people is admittedly a difficult task, and even diagnoses based on highly structured criteria often lead to disagreement among well-trained clinicians. Despite these limitations, as long as legal doctrines continue to explore whether an accused intended to do the criminal act, psychological testimony on theories behind human behavior will be necessary and relevant in criminal cases. While the DSM-IV TR (APA, 2000) specifically cautions against its use to determine legal issues, mental health classifications will nevertheless be relevant in the determination of legal excuses like insanity.

Conclusions

In conclusion, expert testimony on psychological syndromes will continue to be introduced into legal decisions. Widespread acceptance of syndrome evidence is unlikely, however, conscientiously researched areas like Battered Women's Syndrome and Posttraumatic Stress Disorder will continue to be useful in criminal trials. Hopefully, litigation-driven syndromes like Parental Alienation Syndrome will eventually fall from favor in the legal system. Lawyers and judges should become more science literate, and work to ensure that social science evidence is admitted and used properly. While psychological and social science research may not always be an easy fit into the legal system, it will remain an important piece of evidence for legal constructs and fair decisions. Despite the problems outlined in this chapter, psychiatric and psychological evidence has an important role to play in the legal system.

References

Adamou, M.C., and Hale, A.S. (2003). PTSD and the law of psychiatric injury in England and Wales: Finally coming closer? *The Journal of the American Academy of Psychiatry and the Law*, 31, 327–32.

American Psychiatric Association (1980). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders – Text Revision* (4th ed.). Washington, DC: American Psychiatric Association.

Aprilakis, C. (2005). The warrior returns: Struggling to address criminal behavior by veterans with PTSD. *Georgetown Journal of Law and Public Policy*, 3, 541–66.

Block, A.P. (1990). Rape trauma syndrome as scientific evidence. *Archives of Sexual Behavior*, 19, 309–23.

Boeschen, L.E., Sales, B.D., and Koss, M.P. (1998). Rape trauma experts in the courtroom. *Psychology, Public Policy, and Law*, 4, 414–32.

Burgess, L.L., and Holmstrom, A.W. (1974). Rape trauma syndrome. *American Journal of Psychiatry*, 131, 981–6.

Canter, D. (2007). In the kingdom of the blind. In D. Canter and R. Žukausienė (eds), *Psychology and Law: Bridging the Gap*. Aldershot: Ashgate. pp. 1–16.

Cartwright, G.F. (1993). Expanding the parameters of parental alienation syndrome. *The American Journal of Family Therapy*, 21, 205–15.

Chapman v. State, WY 25, 18 P.3d 1164, 1171 (Supreme Court of Wyoming, 2001).

Costanzo, M. (2004). *Psychology Applied to Law*. Belmont, CA: Wadsworth.

Dahir, V.B., Richardson, J.T., Ginsburg, G.P., Gatowski, S.I., Dobbin, S.A., and Merlino, M.L. (2005). Judicial application of *Daubert* to psychological syndrome and profile evidence. *Psychology, Public Policy, and Law*, 11, 62–82.

Delgado, E.J. (1985). Vietnam stress syndrome and the criminal defendant. *Loyola of Los Angeles Law Review*, 19, 473–511.

Dutton, M.A. (1993). Understanding women's responses to domestic violence: A redefinition of battered woman syndrome. *Hofstra Law Review*, 21, 1191–242.

Edmond, G., and Mercer, D. (1998). Trashing "Junk Science." *Stanford Technology Law Review*, 3. Retrieved December 1, 2006 from http://stlr.stanford.edu/STLR/Articles/98_STLR_3.

Faigman, D.L., Kaye, D.H., Saks, M.J., and Sanders, J. (2005). The legal issues of rape trauma syndrome. In D.L. Faigman, D.H. Kaye, M.J. Saks, and J. Sanders (eds), *Modern Scientific Evidence: The Law and Science of Expert Testimony, Volume 2*. Eagan, MN: Thomson/West. pp. 303–17.

Foa, E.B., and Riggs, D.S. (1995). Post-traumatic stress disorder following assault: Theoretical considerations and empirical findings. *Current Directions in Psychological Science*, 4, 61–5.

Frances, A., First, M.B., and Pincus, H.A. (1995). *DSM-IV Guidebook*. Washington, DC: American Psychiatric Press.

Frazier, P.A. (1990). Victim attributions and post-rape trauma. *Journal of Personality and Social Psychology*, 59, 298–304.

Frazier, P.A. (2005). The scientific status of research on rape trauma syndrome. In D.L. Faigman, D.H. Kaye, M.J. Saks, and J. Sanders (eds), *Modern Scientific Evidence: The Law and Science of Expert Testimony, Volume 2*. Eagan, MN: Thomson/West. pp. 317–43.

Frazier, P.A., and Borgida, E. (1992). Rape trauma syndrome: A review of case law and psychological research. *Law and Human Behavior*, 16, 293–311.

Frazier, P., and Schauben, L. (1994). Causal attributions and recovery from rape and other stressful life events. *Journal of Social and Clinical Psychology*, 13, 1–14.

Gardner, R.A. (1985). Recent trends in divorce and custody litigation. *The Academy Forum*, 29, 3–7.

Gardner, R.A. (1998). Recommendations for dealing with parents who induce a Parental Alienation Syndrome in their children. *Journal of Divorce and Remarriage*, 28, 1–23.

Gardner, R.A. (2003a). Court rulings specifically recognizing the parental alienation syndrome in the US and internationally. Retrieved November 15, 2006, from http://www.rgardner.com/refs/pas_legalcites.html.

Gardner, R.A. (2003b). The judiciary's role in the etiology, symptom development, and treatment of the Parental Alienation Syndrome (PAS). *American Journal of Forensic Psychology*, 21, 39–64.

Gardner, R.A. (2004). The relationship between Parental Alienation Syndrome (PAS) and False Memory Syndrome (FMS). *The American Journal of Family Therapy*, 32, 79–99.

Gidycz, C.A., and Koss, M.P. (1989). The impact of adolescent sexual victimization: Standardized measures of anxiety, depression, and behavioral deviancy. *Violence and Victims*, 4, 139–49.

Golding, J.M. (1999). Sexual assault history and long-term physical health problems: Evidence from clinical and population epidemiology. *Current Directions in Psychological Science*, 8, 191–4.

Gudjonsson, G.H., and Haward, L.R.C. (1998). *Forensic Psychology: A Guide to Practice*. London: Routledge.

Hutchings, P.S., and Dutton, M.A. (1997). Symptom severity and diagnoses related to sexual assault history. *Journal of Anxiety Disorders*, 11, 607–18.

Kelly, J.B., and Johnston J.R. (2001). The alienated child: A reformulation of the parental alienation syndrome. *Family Court Review*, 39, 249–66.

Kilpatrick, D.G., Veronen, L.J., and Best, C.L. (1985). Factors predicting psychological distress among rape victims. In C.R. Figley (ed.), *Trauma and Its Wake: The Study and Treatment of Post-traumatic Stress Disorder*. New York: Brunnel/Mazel. pp. 113–41.

Kopetski, L. (1998). Identifying cases of Parental Alienation Syndrome – Part I. *The Colorado Lawyer*, 27(2), 65–8.

Krafka, C., Dunn, M.A., Johnson, M.T., Cecil, J.S., and Miletich, D. (2002). Judge and attorney experiences, practices, and concerns regarding expert testimony in federal civil trials. *Psychology, Public Policy, and Law*, 8, 309–32.

Lalumière, M.L., Harris, G.T., Quinsey, V.L., and Rice, M.L. (2005). *The Causes of Rape: Understanding Individual Differences in Male Propensity for Sexual Aggression*. Washington, DC: American Psychological Association.

Malsch, M., and Freckelton, I. (2005). Expert bias and partisanship: A comparison between Australia and the Netherlands. *Psychology, Public Policy, and Law*, 11, 42–61.

Miller, K.R. (2006). Detecting the undetectable: An examination of the intersection between sudden infant death syndrome and Munchausen by proxy syndrome. *Connecticut Public Interest Law Journal*, 5, 287–315.

M'Naghten's Case 8 Eng. Rep. 718 (House of Lords, 1843).

Moreno, J.A. (2003). Einstein on the bench?: Exposing what judges do not know about science and using child abuse cases to improve how courts evaluate scientific evidence. *Ohio State Law Review*, 64, 531–84.

Morse, S.J. (1998). Excusing and the new excuse defenses: A legal and conceptual review. *Crime and Justice: A Review of Research*, 23, 329–409.

Pearson v. Pearson, 5 P.3d 239, 2000 Alas. LEXIS 69 (Supreme Court of Alaska, 2000).

R. v. Borsch, No. 05-01-25831 (Queen's Bench, Winnipeg Centre, June 22, 2006).

R. v. C. (K.), Carswell Ont 7865 (Ontario Superior Court of Justice, 2002).

R. (B.) c. M. (R.), 66 Q.A.C. 179 (Quebec Court of Appeal, 1994).

Raitt, F.E., and Zeedyk, M.S. (2000). *The Implicit Relation of Psychology and Law: Women and Syndrome Evidence*. London: Routledge.

Rand, D.C. (1997), The spectrum of parental alienation syndrome (part I). *American Journal of Forensic Psychology*. 15, 23–51.

Rothwell v. Kisko, W.D.F.L. 049 (Ontario Court General Division, 1992).

Rueda, C. (2004). Parental alienation syndrome: An inter-rater reliability study. *Dissertation Abstracts International*, 64 (12-B) 6380 (UMI No. AAI3117194).

Saks, M.J., and Lanyon, R.I. (2007). Pitfalls and ethics of expert testimony. In M. Costanzo, D. Krauss, and K. Pezdek (eds), *Expert Psychological Testimony for the Courts*. Mahwah, NJ: Erlbaum. pp. 277–95.

Schnopp-Wyatt, E.N. (2000). Expert testimony in rape trials: Prejudicial or probative? *Dissertation Abstracts International*, 60 (12-B), 6425 (UMI No. 9954909).

Schuller, R.A., and Hastings, P.A. (1996). Trials of battered women who kill: The impact of alternative forms of expert evidence. *Law and Human Behavior*, 20, 167–87.

Schuller, R.A., and Jenkins, G. (2007). Expert evidence pertaining to battered women: Limitations and reconceptualizations. In M. Costanzo, D. Krauss, and K. Pezdek (eds), *Expert Psychological Testimony for the Courts*. Mahwah, NJ: Erlbaum. pp. 203–25.

Siegel, J.C., and Langford, J.S. (1998). MMPI-2 validity scales and suspected parental alienation syndrome. *American Journal of Forensic Psychology*, 16, 5–14.

State v. Cocuzza, No. 1484–79 (New Jersey Superior Court, 1981).

State v. Janes, 121 Wash 2d 220, 850 P.2d 495, 22 A.L.R.5th 921 (Washington Supreme Court, 1993).

State v. Lockett, 121 Misc. 2d 549; 468 N.Y.S.2d 802; 1983 N.Y. Misc. LEXIS 3961 (Supreme Court of New York, Criminal Term, Kings County, 1983).

State v. Martens, 90 Ohio App. 3d 338, 629 N.E.2d 462 (Ohio Court of Appeal, 1993).

Sudderth, L.K. (1998). "It'll come right back at me": The interactional context of discussing rape with others. *Violence Against Women*, 4, 572–84.

Sutherland, E.E. (2006). Undue deference to experts syndrome? *Indiana International and Comparative Law Review*, 16, 375–422.

Swihart, G., Yuille, J., and Porter, S. (1999). The role of state-dependent memory in "Red-outs." *International Journal of Law and Psychiatry*, 22, 199–212.

Ullman, S.E. (2004). Sexual assault victimization and suicidal behavior in women: A review of the literature. *Aggression and Violent Behavior*, 9, 331–51.

Ullman, S.E., and Brecklin, L.R. (2003). Sexual assault history and health-related outcomes in a national sample of women. *Psychology of Women Quarterly*, 27, 46–57.

Walker, L.E. (1979). *The Battered Woman*. New York: Harper and Row.

Walker, L.E. (1984). *The Battered Woman Syndrome*. New York: Springer.

Warshak, R.A. (2001). Current controversies regarding parental alienation syndrome. *Journal of Forensic Psychology*, 19, 29–59.

Weigel, D.J., and Donovan, K.A. (2006). Parental alienation syndrome: Diagnostic and triadic perspectives. *Family Journal: Counseling and Therapy for Couples and Families*, 14, 274–82.

Williams, R.J. (2001). Special issue: Alienated children in divorce: Should judges close the gate on PAS and PA? *Family Court Review*, 39, 267–79.

Wilson, A. (2005). Expert testimony in the dock. *The Journal of Criminal Law*, 69, 330–48.

Young, G., and Yehuda, R. (2006). Understanding PTSD: Implications for court. In G. Young, A.W. Kane, and K. Nicholson (eds) *Psychological Knowledge in Court*. New York, NY: Springer. pp. 55–69.

This page has been left blank intentionally

Chapter 6
Child Sexual Abuse and the Courts

Susan R. Hall

Synopsis

This chapter concerns issues of children in court, particularly alleged sexual abuse victims/survivors, because such cases are more frequently brought to courts in many countries than other cases of abuse. At the same time, the issues discussed are relevant to courts dealing with other types of child maltreatment. The chapter begins with a review of the problem of child maltreatment, and child sexual abuse in particular, and the courts' response to the problem in various countries. Then, after describing psychological research related to child witness competency and needs in criminal child sexual abuse cases, it explains problems with courts applying such research and highlights procedural and policy solutions to meet child witnesses' needs.

Child Sexual Abuse and the Courts

Child abuse is a significant problem that occurs in each society where it has been studied (Schwartz-Kenney et al., 2001). Sociocultural and economic differences among countries affect the definitions of abuse, the emphasis each society places on the problem of child abuse, and the courts' response to it. For example, spanking a child is a criminal act in Sweden (Bottoms and Goodman, 1996), whereas physical or sexual abuse by parents may go overlooked in developing countries like India, where the fact that a child has a home and a caregiver is considered a blessing (Segal, 1996; 2001).

Unlike most countries, the United States gathers fairly reliable prevalence statistics that reveal that child sexual abuse occurs with "alarming frequency" (American Bar Association Criminal Justice Section Task Force on Child Maltreatment, 2002, p. 6). In the past decade, about 3 million reports of child maltreatment (i.e., sexual, physical, emotional, or psychological abuse and/or neglect) have been submitted to U.S. child protective services (CPS) agencies each year (Kuehnle and Sparta, 2006). According to CPS reports in 2004, an

estimated 872,000 children were found to be victims of child maltreatment, 84,584 or 9.7 percent of whom were sexually abused (U.S. Department of Health and Human Services, 2006). Because these numbers only represent substantiated cases involving caretakers and do not take into account underreporting of abuse (which may be partly due to fear, stigma or a "backlash" against reporting abuse), it is generally assumed that they underestimate the actual rates or incidence of maltreatment (Cronch et al., 2006; Faller, 2003; Finkelhor et al., 1990; Kuehnle and Sparta, 2006). Incidence and prevalence rates also vary due to how sexual abuse is defined and how the data are obtained. For example, two meta-analyses produced higher estimates of sexual abuse prevalence, with rates ranging from 12 to 40 percent for females and 3 to 13 percent for males (Bolen and Scannapieco, 1999; Gorey and Leslie, 1997). Thus, Faller (2003) estimated that "as many as one in three or four American females and one in 6 to10 males are sexually abused during their childhoods" (p. 3).

Similar prevalence rates, taking into account similar problems with underreporting, definitional and methodological differences, have been reported in other Western industrialized countries, including Australia (28% of girls and 9% of boys; Hatty and Hatty, 2001), Canada (1 in 8 females and 1 in 25 males; Mian et al., 2001), and Norway (5–16%; Killen, 2001). In countries that do not gather official national statistics about child sexual abuse, the media may either be a frequent source of reports on this problem (e.g., Romania; Muntean and Roth, 2001; Malaysia; Kasim, 2001), or not, due to societal taboos that impede public discussion of sexual abuse (Japan; Kouno and Johnson, 2001).

Although exact incidence and prevalence rates of child sexual abuse are unknown, these generally high numbers mean that professionals whose work brings them into contact with children, including mandated reporters from courts, law enforcement, child welfare, education, health and mental health, will likely encounter cases of child sexual abuse at some point in their careers (Faller, 2003; Wise, 2006). In addition, the crime of child sexual abuse, recognized by all 50 U.S. states, is brought more often to court than other forms of maltreatment (Myers, 1998; Whitcomb, 2003). Criminal prosecution efforts are similarly targeted at sexual abuse in other countries, including Canada (Mian et al., 2001), England (Rogers and Roche, 2001), and Israel (Cohen, 2001). Judges, lawyers, and juries, therefore, need to understand how psychological research can help address some of the problems faced by the courts when hearing cases involving child witnesses who are alleged victims/survivors of sexual abuse and/or other child maltreatment.

The purpose of this chapter is to provide an overview of children in child sexual abuse cases. To accomplish this goal, this chapter will: a) review the problem of child maltreatment, and child sexual abuse in particular, and the courts' response to the problem; b) describe psychological research related to child witness competency and needs in criminal child sexual abuse cases; c) explain problems with courts applying such research; and d) highlight procedural and policy solutions to meet child witnesses' needs. Although the chapter will focus on the United States, relevance of

the research to other contexts (i.e., civil courts; courts in other countries) will also be discussed.

The Problem of, and Legal Responses to, Child Sexual Abuse in the United States

Although child maltreatment has always been a societal problem, it has only been publicly acknowledged in the past several decades. Consider, for example, the United States. Physical abuse was first widely recognized in the United States in the 1960s with Kempe and his colleagues' publication of a seminal paper (Kempe et al., 1962) that coined the term, battered child syndrome (Scheflin, 1998; Wiehe, 1996). It was not until the 1970s, through the research of several sociologists who conducted random telephone and in-person interviews (e.g., Finkelhor, 1984; Russell, 1986), however that child sexual abuse began to be seen as more prevalent than previously suspected (Goodman, 2006).

In response to the increased recognition of the extent of child sexual abuse and problems with the lack of abuse being reported to authorities, three major legal changes occurred in the 1970s (Goodman, 2006). First, mandatory child abuse reporting laws were enacted to facilitate the detection of abuse because victims may be unable or unlikely to report it themselves (Sales et al., 2008). Now in all 50 states, these laws require certain professionals such as physicians and therapists (or "any person" in some states) to report to state officials if they have a reasonable suspicion that a person in a protected class has been abused (Sales et al., 2008). Second, in a considerable shift from seeing children as inaccurate and even dangerous witnesses, children were presumed competent to testify unless shown to be incompetent (Goodman, 2006; Lyon, 2000). This change signaled the recognition that children have the right to testify about events they have experienced, and the ability to provide compelling and accurate accounts of events under appropriate conditions (Bottoms and Goodman, 1996; Westcott et al., 2002). Third, laws that required children's allegations of sexual abuse to be corroborated by physical evidence or another eyewitness were lifted (Goodman, 2006). This change was significant because physical or medical evidence and nonparticipant eyewitnesses are rarely found in such cases (Askowitz and Graham, 1994). Corroborative evidence is more often found for other forms of maltreatment (e.g., physical abuse, neglect) and other crimes (e.g., domestic violence, homicide) (Lyon and Saywitz, 2006). Often the child is the sole witness to the alleged sexual offense (McCauley et al., 2000), and thus is more likely to be called to testify in these cases than other child maltreatment cases (Goodman et al., 1999). These three legal changes have been enacted in some other Western industrialized countries as well (e.g., England, Wales).

The 1980s was marked by an upsurge in public awareness and media attention to child sexual abuse, in part fueled by highly publicized divorce and day care cases (e.g., McMartin Preschool in Manhattan Beach, California; County Walk Day Care

Center in Miami, Florida). By the mid-1980s, the previously taboo subject of child sexual abuse had become "something of a national obsession" (Cohen, 1985, p. 429). In 1990, the issue appeared so severe that the United States government declared child abuse a national emergency (Marks, 1995). In response to public attention and the perceived call to action against such heinous crimes (Askowitz and Graham, 1994), a significant increase in child abuse and neglect incidents were reported to law enforcement and CPS (cases climbed from 850,000 in 1981 to nearly 2.7 million in 1991), resulting in more cases prosecuted and/or pursued through civil litigation (Hall and Sales, 2008). During this time, Ceci and Bruck (1993; 1998) conservatively estimated that over 100,000 children testified in criminal and civil cases each year in the United States, including over 13,000 in sexual abuse cases.

Given concern for the child victims/survivors, mental health professionals began to research the psychological effects of child sexual abuse in the 1980s. Although a wide range of negative effects of sexual abuse were initially identified as indicators of abuse, subsequent research in the past 30 years has confirmed that all sexually abused children do not possess a given set of symptoms, indicators, or syndromes (Kuehnle and Sparta, 2006; Myers and Stern, 2002). Symptoms cannot be used to prove that sexual abuse occurred because non-abused children can experience any behavioral symptom or "indicator" of sexual abuse (Lanning, 2002). However, developmentally unusual/age-inconsistent sexual knowledge and behavior and Posttraumatic Stress Disorder are most consistently reported, having been found to appear more frequently in sexually abused children than in non-abused children (Becker et al., 1995; Kendall-Tackett et al., 1993; Lanning, 2002; McLeer et al., 1998; Wise, 2006). Today, research continues to examine mediating and moderating variables of negative sequelae of sexual abuse, and effective methods to assess and treat negative symptoms as well as highlight strengths (e.g., trauma-focused cognitive behavioral techniques [Cohen et al., 2006; Deblinger and Helfin, 1996]) in a culturally sensitive manner (e.g., Cohen et al., 2001).

In addition, as more children became involved in sexual abuse cases, it became evident that they faced problems in an adult-focused, adversarial legal system. Mental health, criminal justice and legal professionals began to research and write about the needs of allegedly abused children, who often appeared as the sole witness in child sexual abuse cases. As further described below in the section *Children in Court*, research found that many of these children had trouble understanding the criminal justice system (e.g., Saywitz, 1989) and were negatively affected by their involvement in it (e.g., DeFrancis [1969], Eth [1988], Katz and Mazur [1979], Libai [1969], and Parker [1982] as cited in Whitcomb, 2003). Researchers in Australia, Canada and mainland Europe were also studying these common problems and attempting to identify possible solutions (Flin et al., 1996).

In response to the need to accommodate the developmental and psychological needs of children in court and to facilitate the most reliable and complete evidence from these witnesses, by 1988, most states had enacted a number of legal statutory and procedural changes, including allowing children to testify while being hidden

behind screens in the courtroom and videotaped and closed-circuit televised (CCTV) testimony (see also below the section on *Procedural and Policy Solutions*) (Whitcomb, 2003). Similar reforms were enacted in other countries, such as Canada's Bill C-15 and C-126 (Sas et al., 1996), and England and Wales' Criminal Justice Acts of 1988 and 1991 (Wade, 2002). Moreover, the 1989 UN Convention on the Rights of the Child, which is binding international law for some countries including Member States of the Council of Europe, created obligations to allow child witnesses to be heard and to be freed from constraints or emotional distress that might impede their giving evidence. Procedures like CCTV are more likely to be implemented in England, Wales, and some Australian states because of the rebuttable presumption that the child witness in a sexual abuse case can use them (Wade, 2002). In contrast, the United States' Sixth Amendment, which protects the right of the defendant to face the accuser, has presented more challenges to the implementation of CCTV and other courtroom modifications for child witnesses.

In tandem with the enacting of laws designed to benefit child victims in court, there was a burst of research attention paid to the issues of children as witnesses from approximately 1985 to 1995. In the United States, where individual freedoms and defendants' rights are valued, publicity regarding false allegations raised concerns about the validity of child abuse reports. Thus, research in that country focused on the abilities of the child witness, namely the accuracy and suggestibility of children's memories, juror perceptions of their credibility, and questioning techniques that would elicit accurate statements during investigative interviews and courtroom testimony (Bottoms and Goodman, 1996; Whitcomb, 2003). Research in countries that have been able to enact courtroom modifications more easily, such as England, focused their efforts on evaluating innovative procedures in the field. Such research continues to this day.

Psychological Research Related to Child Witnesses in Criminal Child Sexual Abuse Cases

Research regarding child witnesses in sexual abuse cases is very broad and interdisciplinary in nature, spanning the diverse fields of social work, sociology, law, and clinical, community, cognitive, developmental, forensic and social psychology, to name a few (Lyon and Saywitz, 2006; Westcott et al., 2002). Also a "worldwide endeavor," research is being conducted in the United States as well as other countries, including England, Scotland, New Zealand, Australia, Canada, Israel, and Sweden (Saywitz et al., 2002).

Being international in scope, psychological research on child abuse and child witnesses necessarily reflects the sociopolitical values, norms and local issues of the country in which it is generated (Bottoms and Goodman, 1996). Given the historical background of child sexual abuse in the U.S., as just reviewed, it is clear that high-profile cases in our adversarial legal system shaped the burgeoning child witness

research agenda in this country. As a result, it has mainly focused on criminal sexual abuse allegations. Accordingly, this section will focus on current research that informs our knowledge about the competency of children to be witnesses in U.S. criminal courts (i.e., memory, accuracy, suggestibility, truthfulness) as well as their needs in that context (i.e., legal understanding, effects of courtroom participation/testimony).

The research reviewed is also relevant to child witnesses in other kinds of cases (e.g., witness to domestic violence or murder, alleged victim of other maltreatment) and in other kinds of courts (e.g., juvenile, family), but more research is needed to generalize conclusions (Lyon and Saywitz, 2006). Its applicability to cross-cultural and international contexts may also vary due to factors such as the generalizability of laboratory studies to real world practice, the continued inclusion of ethnicity in child maltreatment research, and procedural differences across and within different court systems (i.e., adversarial/common law vs. inquisitorial/civilian). For example, no research has directly compared the needs of child witnesses in the inquisitorial system used in Germany in which alleged child abuse victims are only interviewed by the chairing judge (no juries are used in criminal trials) (Köhnken, 2002), with adversarial systems (e.g., United States) in which defense attorneys cross-examine the child in the presence of a jury.

Child Witness Competency

Although most U.S. state laws presume that all children are competent to testify, if a child is under a certain age, some require a showing of competency or presume young children are incompetent to testify (American Prosecutors Research Institute [APRI], 2004). The judicial test for a child's competency comes from the United States Supreme Court's decision in *Wheeler v. United* States, 159 U.S. 523 (1895). It typically involves a determination of the ability to accurately perceive, recall and share facts, the ability to distinguish truths from lies, and an understanding regarding the obligation to tell the truth (APRI, 2004). Thus, this section will describe the psychological research relevant to elements of this competency determination: children's memory, its accuracy and suggestibility, and their truthfulness.

Memory The study of memory is a complex process that involves acquiring, storing, interpreting, and retrieving information (Faller, 2003). Whether an aspect of an event or experience is encoded into a child's memory depends on the child's behavioral style, whether the child is paying attention to it, is interested and understands it, and the degree of stress the child is experiencing (Baker-Ward and Ornstein, 2002). Also, not understanding an experience, like sexual abuse, may lead to encoding problems (Brainerd and Ornstein, 1991; Pipe et al., 2004). Research findings are mixed and considerable controversy exists regarding the role of stress and arousal in memory. Some researchers conclude that high levels of stress have a debilitating effect on encoding (Ceci and Bruck, 1993; Baker-Ward and Ornstein, 2002), while others

find that children, even three- to four-year olds, are able to retain core rather than peripheral features. Yet each of these findings is affected by each child's individual ability to accurately process stressful events (Goodman, 2006; Saywitz et al., 2002; Schaaf et al., 2002).

Regarding storage, studies have shown that preschool-aged children and adults have similar storage capacity (Brainerd and Ornstein, 1991; Fivush, 1993). Still, how a memory is stored is affected by duration and frequency of exposure to an event, and prior knowledge about an event (Baker-Ward and Ornstein, 2002). Several recent studies have shown that children with more knowledge about an event that they experienced recalled more details about that event than children who had less knowledge about it (Clobb et al. [1993], Greenhoot [2000], and Sutherland, Pipe, Schick, Murray, and Gobbo [2003] as cited in Pipe et al., 2004). Age also plays a role in storage; older children can store information more efficiently than younger children because they have lived longer and gained more experiences and information processing skills (Baker-Ward and Ornstein, 2002).

Also, memories are not stored in a locked box; they are negatively affected by the passage of time, and can be weakened or strengthened by events that happen over time or beliefs that the child has (Baker-Ward and Ornstein, 2002). For example, whether a child experienced an event as involving "shame, perceived responsibility, embarrassment, or guilt, and whether, in turn, it is talked about, reflected on, kept secret, or even negated, may all affect how experiences of abuse or trauma are remembered and recalled over time" (Pipe et al., 2004, p. 444).

Much research attention has focused on how children retrieve, recall, and report non-abusive and allegedly abusive events. Although memory varies at any age due to a number of developmental, personality, and situational factors, children can remember and recall experiences, sometimes over long periods of time (Hewitt, 1999; Pipe et al., 2004). Although memories about experiences that happened to oneself during the first two years of life are not able to be consciously and verbally recalled during childhood or adulthood, Kuehnle and Sparta's (2006) review of research shows that "unique, distinctive, and personally consequential experiences, occurring during the early preschool years, which could include sexual abuse experiences, can be remembered for periods up to 6 years during childhood" (Conway [1996] and Fivush and Schwartzmueller [1998] as cited in Kuehnle and Sparta, 2006, p. 134). Still, young children are generally able to recall less information than older children, and have difficulty remembering the source of their memories. Kuehnle and Sparta also note that when compared to school-aged children (i.e., 7–12 years), preschool children (i.e., 3–6 years) may be less consistent, but not less accurate, when asked to recall specific past events.

In addition to errors of omission (i.e., leaving out information when recalling the memory of an event), people may also make errors of commission (i.e., adding new inaccurate information to a memory of an event) when encoding or recalling an experience. Even though children are more resistant to false abuse-related suggestions than other types of false suggestions (Howe, 2007), some errors of

commission could include false memories or reports of abuse. Anatomically detailed/correct dolls were designed to facilitate children's reports of abuse, but they have been the subject of much controversy and criticism due to concerns that they lead to false reports. Research has shown that using anatomical dolls along with misleading, leading, or suggestive questioning about private parts can lead to false reports of genital or anal touching (Everson and Boat, 2002). Similarly, use of dolls with other props can produce distorted and inaccurate recall of events (Goodman and Melinder, 2007). Everson and Boat noted how additional research is needed to discern the comparative impact of anatomical dolls, interview question types and other props in producing errors of omission and commission in children.

Accuracy How accurate a child is when reporting his or her memory depends on a number of factors, including the strength of the memory (as previously discussed), the way in which the child was asked to report what happened, the type of event that the child experienced, the type of information to be talked about, and influences after the event (Saywitz et al., 2002). One robust finding from psychological research is that children who are asked free recall questions (e.g., "Tell me what happened …") provide the most accurate reports, even though their reports are brief and incomplete (Pipe et al, 2004; Saywitz et al., 2002). They can provide more details when given an open-ended free-recall prompt (e.g., "Tell me more about that …"), more focused recognition questions (e.g., "Did this happen at home or school?"), or cues (e.g., a picture of the child's home or school).

Specific questions are more helpful than free recall questions in eliciting reports of embarrassing experiences or things children have been told to keep secret and such questions can help overcome limits in their vocabulary for sexual topics (Saywitz et al., 2002). Although children's recognition memory is fairly good, the likelihood of errors increases with such questions (Pipe et al., 2004; Saywitz et al., 2002). Focused questions let children know what topics are of interest to the interviewer and may result in the child including erroneous details or producing errors due to feeling pressure to say "yes" or "no" without thinking or to agree with the interviewer given power or status differences (Pipe et al., 2004). Young children often defer to adults (Saywitz et al., 2002). Still, preschoolers in the right conditions can give accurate responses to specific abuse-relevant questions (Rudy and Goodman, 1991). Thus, rather than recommending that yes-no questions be avoided with preschoolers (Peterson and Biggs, 1997), it is preferable to conclude that responses to such questions are less accurate (Lyon, 2002b).

Regarding event type, children can accurately recall experiences that are distinctive, such as is sometimes the case with sexual abuse. Two case studies of children's memory for sexual abuse in which children's recall was compared with corroborating photographic and/or audiotaped evidence found that 50 to 79 percent of details were accurate, but also contained frequent omission errors (Bidrose and Goodman, 2000; Orbach and Lamb, 1999). Comparing abused and non-abused children's memory for a medical exam, both groups of children provided

accurate accounts of the stressful non-abusive events (i.e., venipuncture, anogenital examination) with younger children providing ones that were less complete (Eisen et al., 1998; Eisen et al., 1999). To ensure that an ethical approach is used in medical exam studies, procedures generally include ensuring that a nurse is present when the pediatrician performs the exam and that caregivers can observe or be with the child during the post-medical examination interview with the researcher.

In addition, researchers have also examined children's recall of memories of events that happened once versus those that happened repeatedly. It seems that children's memories for repeated events can be accurate but that younger children especially may confuse episodes and be confused about what details happened at which episode (Pipe et al., 2004). Finally, children's ability to recall temporal information, such as the number, timing, and sequence of events, increases with age (Pipe et al., 2004).

Suggestibility In the U.S., children's suggestibility has been the "primary focus of research, in part reflecting important theoretical issues within academic psychology, but also reflecting U.S. society's concern with false reports and individual freedom" (Bottoms and Goodman, 1996, p. 3). Suggestibility generally refers to errors made due to exposure to false or misleading information or to social or other contextual factors that encourage certain types of responses (Ceci and Bruck, 1993). Many research studies have found that preschool children are more susceptible to suggestion than older children or adults (for reviews, see Ceci and Bruck, 1998; Lyon, 1999; Saywitz et al., 2002), and abused children have shown similar levels of suggestibility when compared to non-abused children (Eisen et al., 1999). Research proliferated on the ability to elicit false reports from children in the 1990s (Goodman, 2006). As a result, some doubt children's abilities to be competent witnesses (e.g., Ceci and Bruck, 1993). Yet, the complexity of suggestibility research merits a more nuanced or balanced approach.

Researchers in this controversial area have therefore begun focusing on the numerous social, cognitive, individual, and family variables that help explain suggestibility (Goodman, 2006; Pipe et al., 2004). Reviews of research (Goodman and Melinder, 2007; Hewitt, 1999; Kuehnle and Sparta, 2006; Pipe et al., 2004; Saywitz et al., 2002) reveal at least eight factors:

- the recency and strength of a child's memory (as previously discussed), as well as if it were imagined rather than experienced (with children asked to recall imagined events being more prone to errors than those asked to report events they actually experienced).
- the need of young children for cues from adult questioners to retrieve their memories; misleading or inaccurate cues increase inaccurate reports, as does reliance on dolls (as noted above) or props, and instructions to think about "pretend" events or to make guesses.

- young children's deference to how the adult questioner perceives and interprets an event. Because young children assume that adults know more, have more experience and are accurate, they do not question their truthfulness or correct them if they are wrong. If the adult uses leading and suggestive questions, especially repeatedly, children's memories can become distorted. Different types of leading questions can be posed in such a way as to presume or intonate that something occurred. For example, instead of asking, "What was the man wearing?", a person using a tag question (i.e., a question seeking confirmation, which is used after a statement) would ask, "The man was wearing a football jersey, wasn't he?" Similarly, a person using a suppositional question would ask, "Where did she touch you?" if the child had never mentioned being touched. Research suggests that elementary-aged children are more resistant to leading questions than preschool-aged children.
- interview characteristics. Young children may also be more easily misled by adult questioners who are perceived as coercive and intimidating as well as when interviews are seen to involve accusations that the suspect did a bad thing or was a bad person, or incentives to respond in a certain way.
- young children's difficulty with source monitoring (i.e., the ability to remember if they learned the information from something they saw, heard about from someone else, or inferred from a situation), as well as metacognitive skills (i.e., the ability to analyze one's own thinking). If a young child lacks the ability to compare his thoughts with false suggestions given to him by an adult, then he may not be able to counter the formation of false memories.
- individual differences in a child's receptive and expressive language abilities, temperament, attachment, self-confidence, and global adaptive functioning. For example, a child who is intimidated by being questioned by adults, or who has clinically significant difficulties with speaking, may have less confidence in her memories or her ability to talk about what she remembers; these factors may result in the child being more susceptible to suggestion. On the other hand, a child who adapts well to new situations and is securely attached to his parents may be more comfortable when being questioned about alleged abuse by adults, which may result in retrieving and reporting more correct information.
- parents who give their children incorrect information. Because autobiographical memories are generally constructed through spoken language, how a caregiver discusses an event with a child will influence the way that the child remembers and recalls it. Parent-child conversations help a young child learn how to organize and report past experiences, including which details of the event to include. Thus, parents may unknowingly affect a child's memory of an event through their discussions together, and/or can deliberately include false information. In some cases, a parent may even coach a child on what to say to others about alleged abuse. The influence of family and other motivational factors are in need of further research.

Truthfulness When requiring witnesses to take an oath, courts are assessing witnesses' willingness to tell the truth. With child witnesses, courts want to ensure that the child knows about the importance of telling the truth and of the difference between truth and lies (Myers, 1998). If children show they understand that lying leads to punishment, they are typically found competent (Lyon, 2002a). Much research in developmental psychology concerns whether children lie (they do) and their understanding of the meaning or wrongfulness of lies (e.g., Vrij, 2002). For example, in studies of young children, most four-year olds were able to explain why it was bad to lie, when asked to talk about others than themselves (Lyon, 2002a).

Other courts allow children to promise to tell the truth rather than give the formal oath (Lyon, 2002a), which is an appropriate modification to traditional procedures because many elementary school children do not understand what it means to "swear" to tell the truth (Saywitz et al., 1990). Developmental research supports asking for an affirmation that "I will tell the truth" from young children because "will" becomes a part of children's vocabulary around the age of two and a half years (Astington [1988] as cited in Lyon, 2002a). Given research with maltreated children aged four to seven, asking children to "promise" that they "will" tell the truth and "won't" tell any "lies" was recommended for young children (Lyon, 2002a). Although some evidence exists to support the use of the oath in various forms in encouraging young children to admit information (e.g., Talwar and Lee [2000] as cited in Lyon, 2002a), more research is needed.

Children in Court

Legal understanding Child witnesses have limited understanding of the legal system—its setting, actors, goals, rules, and procedures (Saywitz et al., 2002). A child may find the physical environment of the courtroom to be formal, austere, and intimidating (McGough, 1994). This is not surprising since most criminal courtrooms were designed as if all witnesses were adults. Compared to other familiar people in their lives, people in the courtroom play strange roles and may wear more formal attire than children are used to seeing (Warren-Leubecker et al., 1989). Even when children and adolescents are familiar with court settings and procedures, misconceptions occur. For example, some children erroneously believed that judges would be mean and sarcastic and yell at them, that the courtroom would be packed with noisy spectators, and that they would be on television (Finnegan, 2000). Also, two studies showed how children with actual legal system involvement showed "*less* accurate knowledge and *more* confusion than age mates without legal experience" (Melton et al. [1992] and Saywitz [1989] as cited in Saywitz et al., 2002, p. 359). Knowledge about commonly used legal terms may be limited, especially for those under 10 years of age (Saywitz, 2002; Warren-Leubecker et al., 1989). Children do not obtain a full understanding of the legal system and the various roles people have within a courtroom until their teenage years (Saywitz, 1989).

Developmental differences and misconceptions may lead to false expectations, mistakes, and/or unrealistic or realistic fears before or during courtroom testimony (Saywitz, 1989; Saywitz et al., 1990; Warren-Leubecker et al., 1989). Although courtrooms can be anxiety-provoking settings for people of all ages, for a child witness, "the courtroom is at best a place of confusion, at worst a terrifying world" (McGough, 1994, p. 10). For instance, one five-year-old child was reported to be "so intimidated by the judge in his long black robe that she refused to raise her head and look at him during her testimony" (Tebo, 2003, p. 53). More research is needed to determine if there is a link between children's legal knowledge and their anticipation about or performance during testimony in court (Saywitz et al., 2002).

Several studies in Western industrialized countries have demonstrated that the anticipation of testifying in criminal court is related to increased anxiety and distress in some children. First, Goodman and colleagues (1992) compared 60 U.S. children who testified in criminal court with a matched group of children who did not testify (e.g., matched on age, alleged abuse severity, and pre-court behavioral adjustment), and found that many children expressed pre-trial fears of testifying, and in particular, of facing the defendant. When these researchers followed up with some of these children 12 to 14 years later, they found more psychological difficulties in those children who reported higher anticipatory distress (Quas et al., 2005). Similarly, Berliner and Conte's (1995) study of 82 U.S. children and families found that the anticipation of testifying was related to children's reports of increased distress three and a half years after they had been interviewed. Feelings of anxiety, helplessness, uncertainty, and fears related to intimidation and misperceptions of the legal system were found in field studies with children in England, Wales, and Northern Ireland (Hamlyn et al., 2004; Plotnikoff and Woolfson, 2004; Wade, 2002). Even children in nonadversarial judicial systems have reported anticipatory testimonial distress. For example, about one third of German children suffered from sleep disturbances, fever and diarrhea before their criminal court appearances (Busse et al. [1996] as cited in Köhnken, 2002).

Effects of courtroom participation/testimony Since the 1980s, some ecologically valid U.S. and international studies have examined the effects of children's participation in actual courtroom procedures, revealing both positive and negative effects (Edelstein et al., 2002). Some children may find the experience positive and empowering (Henry, 1997; Melton and Limber, 1989; Walton, 1994), and others may be resilient to stress (Wade, 2002).

The majority of studies indicate that criminal court testimony is related to short-term distress in some, but not all, children (Edelstein et al., 2002). Children may be negatively impacted by different aspects of traditional courtroom testimony, including: its setting, testifying multiple times, a lack of legal knowledge, the length of the legal process, a lack of social support, developmentally inappropriate interviewing techniques, inadequate protections from the stress of cross examination or other harsh courtroom treatment, and confrontation (Edelstein et al. [2002] as

cited in Hall and Sales, 2008; Ghetti et al., 2002; Goodman et al., 1998; Wade, 2002).

For example, Whitcomb and colleagues' (1994) study of 256 U.S. children and adolescents aged 4 to 17 years found that testifying multiple times and lack of social support were factors associated with children's distress. Similarly, Goodman et al. (1992) found that U.S. children in sexual abuse cases who testified exhibited more behavioral problems after seven months than non-testifiers; this was particularly true for those who lacked maternal support, testified multiple times, and whose allegations were not corroborated. In a 10-year follow-up to this study, children who testified multiple times were later found to report higher amounts of sexual problems and defensive avoidance, especially in more severe intrafamilial abuse cases (Quas et al., 2005). Thus, while the long-term effects of testimonial distress are less clear than short-term ones for child witnesses, problems may occur unless buffered by maternal and other forms of support (Edelstein et al., 2002; Saywitz et al., 2002).

Confronting the defendant has been found to be the most prominent fear, concern, and source of stress for child witnesses in the U.S. and internationally (Hall and Sales, 2008; Saywitz et al., 2002). Stress may be intensified when a child testifies against a parent, relative, trusted friend, or teacher (Lusk and Waterman, 1986). Confrontation has been found to affect negatively the quality (e.g., completeness), reliability, and accuracy of children's testimony (Goodman et al., 1998; Marsil et al., 2002). For example, children in one study who appeared most frightened of the defendant were able to answer fewer of the prosecutors' questions (Goodman et al., 1992).

Long and difficult cross examination has also been found to predict older children's distress (Whitcomb et al., 1994). International research shows that cross-examination contains elements that may increase child witnesses' stress, including: complex and specific language, large numbers of questions, and questions that are unconnected and thus difficult to follow (Brennan and Brennan, 1988; Ghetti et al., 2002; Zajac et al., 2003). Cross-examination also may decrease accuracy (Zajac et al., 2003).

Problems Courts Face in Applying Psychological Research on Child Witnesses

"Criminal prosecutions for child sexual abuse are perhaps the quintessential example of the difficulties generated at the border of law and psychology, offering dread aplenty for the fact finder that must determine what really happened" (Miller and Allen, 1998, p. 147). Legal professionals find that sexual abuse cases uniquely and consistently presents complex and unfamiliar social and psychological issues (APRI, 2004). Jurors, judges and attorneys may not be aware of the psychological research on children in sexual abuse cases and may hold biased views of children who make sexual abuse allegations and how they are supposed to behave in court (Berliner, 1998; Massengale, 2001; McAuliff and Kovera, 2002).

To redress problems with a lack of awareness or misinformation, researchers and clinicians can provide consultation and training to court personnel (e.g., during *voir dire*, continuing education courses) and can serve as witnesses in cases where children testify or do not testify. Consultations and training may include, for example, discussing the research on the effects of courtroom testimony on children and modifications that can be taken to meet child witness needs (Hall and Sales, 2008). Lay and expert witnesses may provide the court with substantive evidence to help prove that the abuse occurred, and rehabilitation evidence to shore up a child's credibility if it is called into question by the defense (Myers and Stern, 2002).

Expert testimony, in particular, plays an important role in criminal and civil child abuse cases, although the nature, scope, and legitimacy of some areas of expert social science testimony are controversial (Berliner, 1998). When expert testimony is presented, admissibility standards such as *Frye* or *Daubert* (as described in Chapter 1 of this volume) apply and may limit the admission of clinical and/or research information in court. The ability of experts to provide substantive evidence that a specific child was abused is hotly contested, and courts take different approaches to this form of testimony: allowing some forms, rejecting most forms, and subjecting it all to a *Frye* or *Daubert* analysis (Myers and Stern, 2002). But because the determination of credibility of a witness is left to the jury, U.S. courts do not allow experts to testify as to the credibility of sexually abused children as a group or that an individual child was credible, believable, or told the truth (Myers and Stern, 2002). When a witness's testimony has been attacked, however, expert testimony may be brought in to buttress the witness's credibility. Because its purpose is rehabilitation, such evidence (e.g., to help explain the nature of child abuse generally, behaviors seen in sexually abused children as a group, reporting delays, inconsistencies, recantation, proper child interviewing techniques) is routinely accepted by the courts in most cases and not subject to *Frye* or *Daubert* (Berliner, 1998; Myers and Stern, 2002).

Regardless of the type of psychological evidence that can be presented in court, lay and expert witnesses should remember that their role is to advocate for the facts, not the alleged child victim or abuser, by carefully scrutinizing and objectively analyzing the relevant information in a manner that comports with current professional knowledge and standards (Sparta, 2005). This role includes telling the court about the reliability risks of research findings and the problems with applying research data on groups to the individual child in a case (Ceci et al., 2002). In some countries (e.g., Germany, the Netherlands, Sweden), expert witnesses may analyze the veracity of children's statements using a procedure called criteria-based content analysis (CBCA) (Köhnken, 2002), which is a component of statement validity analysis that originated in Germany and Sweden. Although CBCA has been widely admitted into evidence to differentiate true from false reports of abuse in some countries, Vrij (2005), in the most recent published review of CBCA research, argues that it does not meet *Daubert* criteria given its high and unknown error rates (30 percent in laboratory studies; no error rates available for field studies of actual cases) and lack of methodological acceptance in the scientific community. While

courts may wish that witnesses could provide more certain or definitive knowledge, such as predictions about whether an individual child was abused or gave a false report, this is not often possible (Kuehnle and Sparta, 2006).

Procedural and Policy Solutions

In response to the unique needs of child witnesses in criminal sexual abuse cases, three main types of changes have been made to investigative court procedures in the U.S. and internationally: a) increasing the skills of professionals involved; b) preparing children for court; and c) modifying the court environment and creating new legal procedures for child witnesses (Cashmore, 2002). Much work has been done to increase the skills of mental health, child protection, and criminal justice professionals who interview children who have made allegations of sexual abuse. Based on the research previously reviewed, professional standards and investigative interviewing protocols have been developed to reduce the suggestibility and maximize the reliability and accuracy of children's reports, such as the National Institute of Child Health and Human Development (NICHD) structured interview protocol (Pipe et al., 2004). In some cases, pre-trial interviews conducted by trained forensic interviewers may be shown at trial in addition to, or in lieu of, calling the child to the witness stand. In addition to the benefit of sparing some children the stress of in-court testimony, the use of pre-recorded testimony has been shown not to affect conviction rates in a Western Australia field study (Wade, 2002). Developmentally appropriate questioning techniques have also been recommended for attorneys and judges for use with child witnesses (e.g., APRI, 2004; Massengale, 2001). The coordination of professionals in multidisciplinary teams and at Children's Advocacy Centers to improve investigations and reduce child stress is another procedural and policy advance.

In addition to the support that a clinician may provide to help a child to cope with the stressors involved in the legal process, including courtroom testimony, communities have developed court support and preparation programs (Whitcomb, 2003). Such programs typically involve education and stress reduction techniques. With regard to ecological validity, evaluations of court preparation programs in different countries have demonstrated effectiveness (e.g., Hamlyn et al., 2004 [England and Wales]; Sas [n.d.] [Canada] as cited in Jones et al., 2005; Köhnken, 2002 [Germany]). Victim advocates in criminal proceedings and Court Appointed Special Advocates or guardian *ad litems* in civil court proceedings also may be involved in preparing the child and the court for his/her special needs.

Finally, a number of policy recommendations (e.g., the American Bar Association's 1985 *Guidelines for the Fair Treatment of Child Witnesses in Cases Where Child Abuse Is Alleged*, the U.S. Department of Justice's 1999 *Breaking the Cycle of Violence: Recommendations to Improve the Criminal Justice Response to Child Victims and Witnesses*, and the National Conference of Commissioners on Uniform

State Laws' 2002 *Uniform Child Witness Testimony by Alternative Methods Act*), support a variety of evidentiary and procedural modifications that can be enacted to accommodate the needs of child witnesses in criminal child abuse cases (McAuliff and Kovera, 2002; Hall and Sales, 2008). Similar policy recommendations and legal changes have been implemented in international contexts (e.g., UN Declaration on the Rights of the Child; Israel's Law of Evidence Revision Protection of Young Children [Sternberg et al., 1996]; the South African Law Commission's 1991 *Report on the Protection of Child Witnesses* [Louw and Olivier, 1996]). These include (Hall and Sales, 2008; Whitcomb, 2003, p. 150):

1. abolishing arbitrary age limitations for competency determinations,
2. protecting the child's identity (e.g., encouraging the media not to enable identification of child victims),
3. excluding spectators when children testify,
4. introducing children's out of court statements via reliable hearsay,
5. admitting children's testimony via closed circuit television or videotape,
6. extending the statute of limitations to allow prosecutions of reports that surface many years after the alleged incidents,
7. allowing attorneys to use leading questions, and demonstrative evidence (e.g., anatomically correct dolls, drawings) when questioning children in court,
8. permitting courtroom design changes,
9. expediting the adjudication process,
10. limiting the length of child testimony,
11. coordinating actions in multiple courts,
12. conducting 'child-friendly' interviews,
13. limiting the number of child interviews,
14. employing a team approach to investigations and prosecution, and
15. instituting specialized child victim assistance or advocacy programs.

Research indicates that many of these procedures are used infrequently in U.S. courts, in part due to legal constraints as well as fears that jurors will perceive children using them as less credible (McAuliff and Kovera, 2002; Whitcomb, 2003). Further research is needed on courtroom modifications and guidelines from other countries to test the impact of these modifications on child well-being and case outcomes in diverse contexts, and to assist professionals in deciding which children are in need of which legally permissible modifications.

Conclusions

The problem of child sexual abuse—which has only fairly recently been brought to societal attention—has also brought the unique needs of children to the attention of the legal system. Courts can use psychological research to understand factors that

contribute to the competency of children to serve as witnesses in trials against their alleged abusers, including their abilities to store and retrieve memories accurately, to tell the truth, and to understand what is expected of them in the courtroom. At the same time, judges and attorneys can benefit from psychological research that informs them about the best ways to talk to and treat children so that they can get the best evidence from child witnesses in criminal and civil cases. It is essential that children's voices be taken into account, as international law recognizes the right of children to be heard in judicial proceedings that affect them (UN Convention on the Rights of the Child, Article 12). At the same time, the needs of adversarial legal systems must be taken seriously. To that end, children and the court system can both benefit from more consistent application of developmentally and culturally appropriate legal and procedural changes that not only protect defendants' rights but also improve children's understanding of the legal system, assist them in managing anxiety while preparing to testify, and support them when giving testimony. Future research is needed to develop further and refine such modifications in diverse legal and sociopolitical contexts.

References

American Bar Association Criminal Justice Section Task Force on Child Maltreatment (2002). *The Child Witness in Criminal Cases*. Washington, DC: American Bar Association.

American Prosecutors Research Institute (2004). *Investigation and Prosecution of Child Abuse* (3rd ed.). Thousand Oaks, CA: Sage.

Askowitz, L.R., and Graham, M.H. (1994). The reliability of expert psychological testimony in child sexual abuse prosecutions. *Cardozo Law Review*, 15, 2027–101.

Baker-Ward, L., and Ornstein, P.A. (2002). Cognitive underpinnings of children's testimony. In H.L. Westcott, G.M. Davies, and R.H.C. Bull (eds) (2002). *Children's Testimony: A Handbook of Psychological Research and Forensic Practice*. Chichester: Wiley. pp. 21–35.

Becker, J.V., Alpert, J.L., BigFoot, D.S., Bonner, B.L., Geddie, L.F., Henggeler, S.W., Kaufman K.L., and Walker C.E. (1995). Empirical research on child abuse treatment: Report by the child abuse and neglect treatment working group, American Psychological Association, *Journal of Clinical Child Psychology*, 24 (Suppl.), 23–46.

Berliner, L. (1998). The use of expert testimony in child sexual abuse cases. In S.J. Ceci and H. Hembrooke (eds), *Expert Witnesses in Child Abuse Cases: What Can and Should Be Said in Court*. Washington, DC: American Psychological Association. pp. 11–27.

Berliner, L., and Conte, J.R. (1995). The effects of disclosure and intervention on sexually abused children. *Child Abuse and Neglect*, 19(3), 371–84.

Bidrose, S., and Goodman, G.S. (2000). Testimony and evidence: A scientific case study of memory for child sexual abuse. *Applied Cognitive Psychology*, 14, 197–213.

Bolen, R.M., and Scannapieco, M. (1999). Prevalence of child sexual abuse: A corrective metanalysis. *Social Service Review*, 73, 281–313.

Bottoms, B.L., and Goodman, G.S. (1996). International perspectives on child witnesses: An introduction to the issues. In B.L. Bottoms, and G.S. Goodman (eds), *International Perspectives on Child Abuse and Children's Testimony: Psychological Research and Law*. Thousand Oaks, CA: Sage. pp. 1–11.

Brainerd, C., and Ornstein, P. (1991). Children's memory for witnessed events: The developmental backdrop. In J. Doris (ed.), *The Suggestibility of Children's Recollections*. Washington, DC: American Psychological Association. pp. 10–20.

Brennan, M., and Brennan, R.E. (1988). *Strange Language: Child Victims under Cross Examination* (3rd ed.) Wagga Wagga, NSW: Riverina Murray Institute of Higher Education.

CAPTA (1974). The Child Abuse Prevention and Treatment Act, U.S. Public Law 93-247.

Cashmore, J. (2002). Innovative procedures for child witnesses. In H.L. Westcott, G.M. Davies, and R.H.C. Bull (eds), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice*. Chichester: Wiley. pp. 203–17.

Ceci, S.J., and Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*, 113, 403–39.

Ceci, S.J., and Bruck, M. (1998). Children's testimony: Applied and basic issues. In W. Damon (ed.), *Handbook of Child Psychology: Vol. 4. Child Psychology in Practice* (5th ed.). New York: Wiley. pp. 713–74.

Ceci, S.J., Crossman, A.M., Scullin, M.H., Gilstrap, L., and Huffman, M.L. (2002). Children's suggestibility research: Implications for the courtroom and the forensic interview. In H.L. Westcott, G.M. Davies, and R.H.C. Bull (eds), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice*. Chichester: Wiley. pp. 117–30.

Cohen, A. (1985). Note, the unreliability of expert testimony on the typical characteristics of sexual abuse victims. *Georgetown Law Review*, 74, 429–56.

Cohen, J.A., Deblinger, E., Mannarino, A.P., and de Arellano, M.A. (2001). The importance of culture in treating abused and neglected children: An empirical review. *Child Maltreatment*, 6, 148–57.

Cohen, J.A., Mannarino, A.P., and Deblinger, E. (2006). *Treating Trauma and Traumatic Grief in Children and Adolescents*. New York: Guilford.

Cohen, T. (2001). Israel. In B.M. Schwartz-Kenney, M. McCauley, and M.A. Epstein (eds), *Child Abuse: A Global View*. Westport, CT: Greenwood Press. pp. 85–98.

Cronch, L.E. Viljoen, J.L., and Hansen, D.J. (2006). Forensic interviewing in child sexual abuse cases: Current techniques and future directions. *Aggression and Violent Behavior*, 11, 195–207.

Deblinger, E., and Helfin, A.H. (1996). *Treating Sexually Abused Children and Their Non-offending Parents*. Thousand Oaks, CA: Sage.

Edelstein, R.S., Goodman, G.S., Ghetti, S., Alexander, K.W., Quas, J.A., Redlich, A.D., Schaaf, J.M., and Cordon, I.M. (2002). Child witnesses' experiences post-court: Effects of legal involvement. In H.L. Westcott, G.M. Davies, and R.H.C. Bull (eds), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice*. Chichester: Wiley. pp. 261–77.

Eisen, M.L. Goodman, G.S., Davis, S.L., and Qin, J. (1999). Individual differences in maltreated children's memory and suggestibility. In L.M. Williams and V.L. Banyard (eds), *Trauma and Memory*. Thousand Oaks, CA: Sage. pp. 31–46.

Eisen, M.L., Goodman, G.S., Qin, J., and Davis, S.L. (1998). Memory and suggestibility in maltreated children: New research relevant to evaluating allegations of abuse. In S.J. Lynn and K.M. McConkey (eds), *Truth in Memory*. New York: Guilford. pp. 163–89.

Everson, M.D., and Boat, B.W. (2002). The utility of anatomical dolls and drawings in child forensic interviews. In M.L. Eisen, J.A. Quas, and G.S. Goodman (eds), *Memory and Suggestibility in the Forensic Interview*. Mahwah, NJ: Erlbaum. pp. 383–408.

Faller, K.C. (2003). *Understanding and Assessing Child Sexual Maltreatment* (2nd ed.). Thousand Oaks, CA: Sage.

Finkelhor, D. (1984). *Child Sexual Abuse*. New York: Free Press.

Finkelhor, D., Hotaling, G., Lewis, I.A., and Smith, C. (1990). Sexual abuse in a national study of adult men and women: Prevalence, characteristics, and risk factors. *Child Abuse and Neglect*, 14, 19–28.

Finnegan, M.J. (2000). Creating and administering a kids court program. *APRI Update*, 13(5). Retrieved November 15, 2006, from http://www.ndaa-apri.org/publications/newsletters/update_volume_13_number_5_2000.html.

Fivush, R. (1993). Developmental perspectives on autobiographical recall. In G.S. Goodman and B.L. Bottoms (eds), *Child Victims, Child Witnesses*. New York: Guilford. pp. 1–24.

Flin, R., Kearney, B., and Murray, K. (1996). Children's evidence: Scottish research and law. In B.L. Bottoms, and G.S. Goodman (eds), *International Perspectives on Child Abuse and Children's Testimony: Psychological Research and Law*. Thousand Oaks, CA: Sage. pp. 114–31.

Friedman, W. (1982). *The Developmental Psychology of Time*. New York: Academic Press.

Ghetti, S., Alexander, K.W., and Goodman, G.S. (2002). Legal involvement in child sexual abuse cases: Consequences and interventions. *International Journal of Law and Psychiatry*, 25, 235–51.

Goodman, G.S. (2006). Children's eyewitness memory: A modern history and contemporary commentary. *Journal of Social Issues*, 62i, 811–32.

Goodman, G.S., and Melinder, A. (2007). Child witness research and forensic interviews of young children: A review. *Legal and Criminological Psychology*, 12, 1–19.

Goodman, G.S., Pyle-Taub, E., Jones, D.P.H., England, P., Port, L.K., Rudy, L., and Prado, L. (1992). Testifying in criminal court: Emotional effects on child sexual assault victims. *Monographs of the Society for Development in Child Research*, 229(5).

Goodman, G.S., Quas, J., Bulkley, J., and Shapiro, C. (1999). Innovations for child witnesses: A national survey. *Psychology, Public Policy, and Law*, 5, 255–81.

Goodman, G.S., Tobey, A.E., Batterman-Faunce, J.M., Orcutt, H., Thomas, S., Shapiro, C., and Sachsenmaier, T. (1998). Face-to-face confrontation: Effects of closed-circuit technology on children's eyewitness testimony and jurors' decisions. *Law and Human Behavior*, 22, 165–203.

Gorey, K.M., and Leslie, D.R. (1997). The prevalence of child sexual abuse: Integrative review adjustment for potential response and measurement biases. *Child Abuse and Neglect*, 21, 391–8.

Hall, S.R., and Sales, B.D. (2008). *Courtroom Modifications for Child Witnesses*: *Law and Science in Forensic Evaluations*. Washington, DC: American Psychological Association.

Hamlyn, B., Phelps, A. Turtle, J., and Sattar, G. (2004). *Are Special Measures Working? Evidence from Surveys of Vulnerable and Intimidated Witnesses*. Home Office Research Study, No. 283. London: Home Office Research, Development and Statistics Directorate.

Hatty, S.E., and Hatty, J. (2001). Australia. In B.M. Schwartz-Kenney, M. McCauley, and M.A. Epstein (eds), *Child Abuse: A Global View*. Westport, CT: Greenwood Press. pp. 1–15.

Henry, J. (1997). System intervention trauma to child sexual abuse victims following disclosure. *Journal of Interpersonal Violence*, 12, 499–512.

Hewitt, S.K. (1999). *Assessing Allegations of Sexual Abuse in Preschool Children: Understanding Small Voices*. Thousand Oaks, CA: Sage.

Howe, M.L. (2007). Children's emotional false memories. *Psychological Science*, 18, 856–60.

Jones, L.M., Cross, T.P., Walsh, W.A., and Simone, M. (2005). Criminal investigations of child abuse: The research behind "best practices". *Trauma, Violence, and Abuse*, 6, 254–68.

Kasim, M.S. (2001). Malaysia. In B.M. Schwartz-Kenney, M. McCauley, and M.A. Epstein (eds), *Child Abuse: A Global View*. Westport, CT: Greenwood Press. pp. 131–44.

Kempe, C.H., Silverman, F.N., Steele, B.F., Drogemueller, W., and Silver, H.K. (1962). The battered child syndrome. *Journal of the American Medical Association*, 181, 17–24.

Kendall-Tackett, K., Williams, L., and Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. *Psychological Bulletin*, 113, 164–80.

Killen, K. (2001). Norway. In B.M. Schwartz-Kenney, M. McCauley, and M.A. Epstein (eds), *Child Abuse: A Global View*. Westport, CT: Greenwood Press. pp. 161–73.

Köhnken, G. (2002). A German perspective on children's testimony. In H.L. Westcott, G.M. Davies, and R.H.C. Bull (eds), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice*. Chichester: Wiley. pp. 233–44.

Kouno, A., and Johnson, C.F. (2001). Japan. In B.M. Schwartz-Kenney, M. McCauley, and M.A. Epstein (eds), *Child Abuse: A Global View*. Westport, CT: Greenwood Press. pp. 99–115.

Kuehnle, K., and Sparta, S.N. (2006). Assessing child sexual abuse allegations in a legal context. In S.N. Sparta and G.P. Koocher (eds), *Forensic Mental Health Assessment of Children and Adolescents*. New York: Oxford. pp. 129–48.

Lanning, K.V. (2002). Criminal investigation of sexual victimization of children. In J.E.B. Myers, L. Berliner, J. Briere, C.T. Hendrix, C. Jenny, and T.A. Reid (eds), *The APSAC Handbook on Child Maltreatment* (2nd ed.). Thousand Oaks, CA: Sage. pp. 329–47.

Louw, D., and Olivier, P. (1996). Listening to children in South Africa. In B.L. Bottoms, and G.S. Goodman (eds), *International Perspectives on Child Abuse and Children's Testimony: Psychological Research and Law*. Thousand Oaks, CA: Sage. pp. 168–81.

Lusk, R., and Waterman, J. (1986). Effects of sexual abuse on children. In K. MacFarlane and J. Waterman (eds), *Sexual Abuse of Young Children*. New York: Guilford. pp. 15–29.

Lyon, T.D. (1999). The new wave of suggestibility research: A critique. *Cornell Law Review*, 84, 1004–87.

Lyon, T.D. (2000). Child witnesses and the oath: Empirical evidence. *Southern California Law Review*, 73, 1017–74.

Lyon, T.D. (2002a). Child witnesses and the oath. In H.L. Westcott, G.M. Davies, and R.H.C. Bull (eds), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice*. Chichester: Wiley. pp. 245–60.

Lyon, T.D. (2002b). Expert testimony on the suggestibility of children: Does it fit? In B.L. Bottoms, M.B. Kovera, and B.D. McAuliff (eds), *Children, Social Science, and the Law*. New York: Cambridge University Press. pp. 378–411.

Lyon, T.D., and Saywitz, K.J. (2006). From post-mortem to preventive medicine: Next steps for research on child witnesses. *Journal of Social Issues*, 62(4), 833–61.

Marks, R.G. (1995). Should we believe the people who believe the children: The need for a new sexual abuse tender years hearsay exception statute. *Harvard Journal on Legislation*, 32, 207–54.

Marsil, D.F., Montoya, J., Ross, D., and Graham, L. (2002). Children as victims and witnesses in the criminal trial process: Child witness policy: Law interfacing with social science. *Law and Contemporary Problems*, 65, 209–42.

Massengale, J. (2001). Facilitating children's testimony. *Update*, 14(6). Retrieved February 23, 2009, from http://www.ndaa.org/publications/newsletters/update_volume_14_number_6_2001.html.

McAuliff, B.D., and Kovera, M.B. (2002). The status of evidentiary and procedural innovations in child abuse proceedings. In B.L. Bottoms, M.B. Kovera, and B.D. McAuliff (eds), *Children, Social Science, and the Law*. New York: Cambridge University Press. pp. 412–45.

McCauley, M., Schwartz-Kenny, B.M., Epstein, M.A., and Tucker, E.J. (2000). United States [Electronic version]. In B.M. Schwartz-Kenny (ed.), *Child Abuse: A Global View*. Westport, CT: Greenwood. pp. 241–55.

McGough, L.S. (1994). *Child Witnesses: Fragile Voices in the American Legal System*. New Haven, CT: Yale University Press.

McLeer, S.V., Dixon, J.F., Henry, D., Ruggiero, K., Escovitz, K., Niedda, T., and Scholle, R. (1998). Psychopathology in non-clinically referred sexually abused children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 37, 1326–1333.

Melton, G.B., and Limber, S. (1989). Psychologists' involvement in cases of child maltreatment: Limits of role and expertise. *American Psychologist*, 44, 1225–1233.

Mian, M., Bala, N., and MacMillan, H. (2001). Canada. In B.M. Schwartz-Kenney, M. McCauley, and M.A. Epstein (eds), *Child Abuse: A Global View*. Westport, CT: Greenwood Press. pp. 17–33.

Miller, J.S., and Allen, R.J. (1998). The expert as educator. In S.J. Ceci and H. Hembrooke (eds), *Expert Witnesses in Child Abuse Cases: What Can and Should Be Said in Court*. Washington, DC: American Psychological Association. pp. 137–55.

Muntean, A., and Roth, M. (2001). Romania. In B.M. Schwartz-Kenney, M. McCauley, and M.A. Epstein (eds), *Child Abuse: A Global View*. Westport, CT: Greenwood Press. pp. 176–93.

Myers, J.E.B. (1998). *Legal Issues in Child Abuse and Neglect Practice* (2nd ed.). Thousand Oaks, CA: Sage.

Myers, J.E.B., and Stern, P. (2002). Expert testimony: Current research and practice implications. In J.E.B. Myers, L. Berliner, J. Briere, C.T. Hendrix, C. Jenny, and T.A. Reid (eds), *The APSAC Handbook on Child Maltreatment* (2nd ed.). Thousand Oaks, CA: Sage. pp. 379–401.

Orbach, Y., and Lamb, M.E. (1999). Assessing the accuracy of a child's account of sexual abuse: A case study. *Child Abuse and Neglect*, 23, 91–8.

Peterson, C., and Biggs, M. (1997). Interviewing children about trauma: Problems with "specific" questions. *Journal of Traumatic Stress*, 10, 279–90.

Pipe, M., Lamb, M.E., Orbach, Y., and Esplin, P.W. (2004). Recent research on children's testimony about experienced and witnessed events. *Developmental Review*, 24, 440–68.

Plotnikoff, J., and Woolfson, R. (2004). *Executive Summary. In Their Own Words: The Experiences of 50 Young Witnesses in Criminal Proceedings*. London: NSPCC. Retrieved August 3, 2007, from http://www.nspcc.org.uk/Inform/Publications/Downloads/InTheirOwnWords pdf gf25460.pdf.

Quas, J.A., Goodman, G.S., and Ghetti, S. (2005). Childhood sexual assault victims: Long-term outcomes after testifying in criminal court. *Monographs of the Society for Research in Child Development*, 70, 1–145.

Rogers, W.S., and Roche, J. (2001). England. In B.M. Schwartz-Kenney, M. McCauley, and M.A. Epstein (eds), *Child Abuse: A Global View*. Westport, CT: Greenwood Press. pp. 35–50.

Rudy, L., and Goodman, G.S. (1991). Effects of participation on children's reports: Implications for children's testimony. *Developmental Psychology*, 27, 1–26.

Russell, D.E.H. (1986). *The Secret Trauma: Incest in the Lives of Girls and Women*. New York: Basic Books.

Sales, B.D., DeKraai, M.B., Hall, S.R., and Duvall, J.C. (2008). Child therapy and the law. In R.J. Morris and T.R. Kratochwill (eds), *The Practice of Child Therapy* (4th ed.). New York: Taylor and Francis.

Sas, L.D., Wolfe, D.A., and Gowdey, K. (1996). Children and the courts in Canada. In B.L. Bottoms, and G.S. Goodman (eds), *International Perspectives on Child Abuse and Children's Testimony: Psychological Research and Law*. Thousand Oaks, CA: Sage. pp. 77–95.

Saywitz, K.J. (1989). Children's conceptions of the legal system: "Court is a place to play basketball." In S.J. Ceci, D.F. Ross, and M.P. Toglia (eds), *Perspectives on Children's Testimony*. New York: Springer-Verlag. pp. 131–57.

Saywitz, K.J. (2002). Developmental underpinnings of children's testimony. In H.L. Westcott, G.M. Davies, and R.H.C. Bull (eds), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice*. Chichester: Wiley. pp. 3–35.

Saywitz, K.J., Goodman, G.S., and Lyon, T.D. (2002). Interviewing children in and out of court: Current research and practice implications. In J.E.B. Myers, L. Berliner, J. Briere, C.T. Hendrix, C. Jenny, and T.A. Reid. (eds), *The APSAC Handbook on Child Maltreatment* (2nd ed.). Thousand Oaks, CA: Sage. pp. 349–77.

Saywitz, K., Jaenicke, C., and Camparo, L. (1990). Children's knowledge of legal terminology. *Law and Human Behavior*, 14, 523–35.

Schaaf, J.M., Alexander, K.W., Goodman, G.S., Ghetti, S., Edelstein, R.S., and Castelli, P. (2002). Children's eyewitness memory: True disclosures and false reports. In B.L. Bottoms, M.B. Kovera, and B.D. McAuliff (eds), *Children, Social Science, and the Law*. New York: Cambridge University Press. pp. 242–377.

Scheflin, A.W. (1998). Narrative truth, historical truth, and forensic truth. In L.E. Lifson and R.I. Simon (eds), *The Mental Health Practitioner and the Law: A*

Comprehensive Handbook. Cambridge, MA: Harvard University Press. pp. 299–328.

Schwartz-Kenney, B.M., McCauley, M., and Epstein, M.A. (eds) (2001). *Child Abuse: A Global View*. Westport, CT: Greenwood Press.

Segal, U.A. (1996). Children as witnesses: India is not ready. In B.L. Bottoms and G.S. Goodman (eds), *International Perspectives on Child Abuse and Children's Testimony*. Thousand Oaks, CA: Sage. pp. 266–82.

Segal, U.A. (2001). India. In B.M. Schwartz-Kenney, M. McCauley, and M.A. Epstein (eds), *Child Abuse: A Global View*. Westport, CT: Greenwood Press. pp. 51–65.

Smith, K., and Tilney, S. (2007). *Vulnerable Adult and Child Witnesses*. New York: Oxford University Press.

Sparta, S.N. (2005). Expert mental health testimony in child abuse cases. In P.F. Talley (ed.), *Handbook for the Treatment of Abused and Neglected Children*. New York: Haworth Press. pp. 397–421.

Sternberg, K.J., Lamb, M.E., and Hershkowitz, I. (1996). Child sexual abuse investigations in Israel: Evaluating innovative practices. In B.L. Bottoms, and G.S. Goodman (eds), *International Perspectives on Child Abuse and Children's Testimony: Psychological Research and Law*. Thousand Oaks, CA: Sage. pp. 62–76.

Tebo, M.G. (2003, April). The most vulnerable clients: Attorneys must deal with special issues when kids come into contact with the courts. *ABA Journal*, 89, 46–53, 63.

U.S. Department of Health and Human Services, Administration on Children, Youth and Families. (2006). *Child Maltreatment 2004*. Washington, DC: U.S. Government Printing Office. Retrieved December 18, 2006, from http://www.acf.hhs.gov/programs/cb/pubs/cm04/index.htm.

Vrij, A. (2002). Deception in children: A literature review and implications for children's testimony. In Westcott, H.L., Davies, G.M., and Bull, R.H.C. (eds) (2002). *Children's Testimony: A Handbook of Psychological Research and Forensic Practice*. Chichester: Wiley. pp. 175–94.

Vrij, A. (2005). Criteria-based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law*, 11, 3–41.

Wade, A. (2002). New measures and new challenges: Children's experiences of the court process. In H.L. Westcott, G.M. Davies, and R.H.C. Bull (eds), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice*. Chichester: Wiley. pp. 219–32.

Walton, E. (1994). The confrontation clause and the child victim of sexual abuse. *Child and Adolescent Social Work Journal*, 11, 195–207.

Warren-Leubecker, A., Tate, C., Hinton, I., and Ozbeck, N. (1989). What do children know about the legal system and when do they know it? First steps down a less traveled path in child witness research. In S. Ceci, M. Toglia, and D. Ross (eds), *Perspectives on Children's Testimony*. New York: Springer-Verlag. pp. 158–83.

Westcott, H.L., Davies, G.M., and Bull, R.H.C. (eds) (2002). *Children's Testimony: A Handbook of Psychological Research and Forensic Practice*. Chichester: Wiley.

Wheeler v. United States, 159 U.S. 523 (1895).

Whitcomb, D. (2003). Legal interventions for child victims. *Journal of Traumatic Stress*, 16, 149–57.

Whitcomb, D., Goodman, G.S., Runyan, D.K., and Hoak, S. (1994). *The Emotional Effects of Testifying on Sexually Abused Children*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, National Institute of Justice.

Wiehe, V.R. (1996). *Working with Child Abuse and Neglect: A Primer*. Thousand Oaks, CA: Sage.

Wise, D. (2006). Child abuse assessment. In M. Hersen (ed.), *Clinician's Handbook of Child Behavioral Assessment*. Burlington, MA: Academic Press. pp. 549–68.

Zajac, R., Gross, J., and Hayne, H. (2003). Asked and answered: Questioning children in the courtroom. *Psychiatry, Psychology and Law*, 10, 199–209.

This page has been left blank intentionally

Chapter 7
Sexual Harassment: Antecedents, Consequences, and Juror Decisions

Sarah M. Greathouse, Lora M. Levett, and Margaret Bull Kovera

Synopsis

Although men and women have coexisted within the workforce for some time, only within the past few decades has there been formal recognition that employees may be negatively targeted and unfairly treated because of their gender. This recent recognition of sexual harassment in the workplace has spurred interest in scholars from the legal arena to the social sciences (Wiener and Gutek, 1999). Thus far, social scientists have mainly focused on estimating how often sexual harassment occurs, the causes of sexual harassment, the consequences of sexual harassment, and how laypeople and jurors perceive incidents of sexual harassment (Goodman-Delahunty, 1999; Wiener and Gutek, 1999). In this chapter, we examine the findings in these burgeoning research areas. We first review the current state of sexual harassment case law to provide a legal background for the subsequent research. We then turn to the social science research on the antecedents and consequences of sexual harassment. Last, we examine the factors that influence juror decisions about liability and damage awards in sexual harassment cases.

Review of Sexual Harassment Law

As recognition of sexual harassment in the workforce increased in the 1960s, the Equal Opportunity Employment Commission (EEOC) addressed the issue by providing a formal definition of sexual harassment and establishing that the victim of sexual harassment has the right to seek redress for the discriminatory behavior (Title VII of the Civil Rights Act of 1964). According the EEOC, sexual harassment was defined as "unwelcome advances, requests for sexual favors, and other verbal or physical conduct of a sexual nature." For the behavior in question to be considered harassing, the victim had to perceive one's responses

to the advances as conditions of his or her employment and these advances had to have a negative impact on the victim's work performance.

As there is often little physical evidence present in sexual harassment cases, determinations of sexual harassment are often based on inferences of the victim's perceptions of the situation and the victim's reactions to the alleged behavior (Kovera and Cass, 2002). Such subjective determinations can create difficulties in a court of law where determinations are supposed to be based on fact. In the development of sexual harassment as a legal concept, the courts have attempted to decrease the role of subjective judgment by establishing guidelines for the trier of fact to use in determining whether the alleged behavior constitutes sexual harassment and if so, how the victim should be compensated.

Although sexual harassment was first thought to constitute a situation in which an employer directly threatens a subordinate's employment status unless the employee complies with the employer's sexual requests, referred to as "*quid pro quo*" sexual harassment, the Supreme Court established in *Meritor Savings Bank v. Vinson* (1986) that a direct threat was not necessary for a situation to constitute sexual harassment. Rather, the Supreme Court ruled that if intimidating, offensive behavior is pervasive in the victim's working environment and the harassing behavior was not welcomed by the victim, a hostile working environment due to sexual harassment was present. In these cases, the trier of fact is instructed to evaluate whether the plaintiff's conduct indicated that the alleged sexual advances were unwelcome. In addition to the plaintiff's actions, her dress and demeanor can be considered when determining the welcomeness of the alleged behavior. In sum, a wide variety of behaviors may be considered to be sexual harassment depending upon the circumstances. Even asking for a date could be construed as sexual harassment if, for example, complying with the request was made a condition of employment (*quid pro quo* harassment), or the requests were unwelcome and pervasive enough to alter the conditions of the workplace (hostile work environment).

In most jurisdictions, jurors are instructed to rely on the reasonable person standard when deciding liability in a sexual harassment case (e.g., deciding whether the sexual behavior was severe, pervasive, and unwelcome). The reasonable person standard for determining whether behavior is considered sexually harassing asks jurors to consider whether the behavior in question was unwelcome and was so severe and pervasive as to create an "environment that a reasonable person would find hostile or abusive" (*Harris vs. Forklift Systems, Inc.*, 1993, p. 21). Several courts have adopted such a reasonable person standard (*Rabidue v. Osceola Refining Co.*, 1986); however, other courts have adopted a reasonable woman standard (*Ellison v. Brady*, 1991). Specifically, the Ninth Circuit Court of Appeals in the *Ellison v. Brady* case ruled that a finding of sexual harassment would be required in cases in which the plaintiff "alleges conduct which a reasonable woman would consider sufficiently severe or pervasive to alter the conditions of employment and create an abusive working environment" (*Ellison v. Brady*, 1991, p. 879). The assumption behind the reasonable woman standard is that men and women differ in their perceptions of

sexual harassment (Rotundo et al., 2001). Therefore, when making a decision about a sexual harassment case, jurors should be required to take the perspective of the person alleging the sexual harassment claim.

The *Ellison v. Brady* decision did not cite empirical research to support the use of a reasonable woman standard; rather, the ruling was based on a commonsense assumption that instructing jurors to adopt the perspective of a sexual harassment victim will cause jurors who might otherwise rely on unfounded stereotypes of sexual harassment in their decision-making to become instead sensitive to the unique perspective of a victim in that situation (Gutek and O'Connor, 1995). Since the *Ellison* decision, social scientists have begun to conduct research testing the *Ellison* court's assumptions. A review of the effects of the different legal standards on jurors' decisions will be presented later in this chapter.

If the jury finds the defendant liable for gender discrimination due to a hostile work environment (using either the reasonable woman or reasonable person standard), jurors are instructed to award damages if they believe that damages are warranted. Specifically, Title VII of the Civil Rights act of 1991 stipulates that plaintiffs in a sexual harassment case may seek compensation for mental anguish, pain, and suffering. In addition, plaintiffs may choose to append state tort law and antidiscrimination claims to their Title VII claim to seek additional damages in sexual harassment cases (Sharkey, 2006). The jury may assign both compensatory and punitive damage awards to the plaintiff. Compensatory damage awards are designed to reimburse plaintiffs for the damage caused by the harassment so that they are returned to the same condition they experienced before the harassment. For example, compensatory damage awards can be assigned to reimburse a plaintiff for lost wages, promotions lost due to missed work, or any other economic loss suffered by the plaintiff. Compensatory damage awards may also be used to reimburse the plaintiff for non-economic losses such as the plaintiff's pain and suffering resulting from the defendant's harassing behavior. The jury may also assign punitive damage awards. Punitive damage awards may be assigned to punish the defendant for their unlawful behavior and to deter them from repeating such behavior. Juries may also assess punitive damages in an attempt to deter others from participating in similar behavior. What evidence is necessary for the plaintiff to recover damages, however, has caused debate (see Chapter 8, in Volume I for a more detailed discussion of civil trial issues).

The Supreme Court addressed this issue in *Harris v. Forklift Systems* (1993) and *Turic v. Holland Hospitality, Inc.* (1996). *Harris v. Forklift Systems* involved a woman who quit her job due to offensive comments made by the president of the company. The court originally ruled that even though the comments may have been considered offensive, they were not so degrading that they were psychologically damaging. The Supreme Court reversed the original decision, stating that Title VII could be violated regardless of whether psychological harm ensued. Thus, the Court determined that damages could be sought regardless of whether the plaintiff experienced psychological harm.

In *Turic v. Holland Hospitality, Inc*., the Supreme Court addressed whether expert evidence from a psychologist is needed to prove that the sexual harassment induced emotional distress, allowing the plaintiff to recover damages to compensate for the distress. This case involved a young mother who asserted she was wrongfully terminated from her employment because she was considering aborting her pregnancy. The original finding ruled for the plaintiff and awarded her compensatory and punitive damage awards. The defense appealed the jury verdict favoring the plaintiff, as well as the awarded compensatory and punitive damages claiming that the defense could not assess whether the plaintiff's claim of emotional distress was valid if a psychologist was unable to examine the plaintiff. The appellate court upheld the trial court ruling stating that the plaintiff's claim of emotional distress and the damages awarded to compensate the plaintiff for that distress could be evaluated through the plaintiff's testimony and the circumstances of the case.

These two cases helped establish the necessary conditions for the recovery of damages in sexual harassment cases. Although plaintiffs need to establish that their place of employment was hostile and caused distress for sexual harassment to be present, they need not prove that they were psychologically damaged. Furthermore, when a plaintiff claimed general emotional distress, pain, and suffering, it was not necessary for a psychologist to affirm these claims.

International Definitions and Legislation Related to Sexual Harassment

Most industrialized nations have not agreed upon one standard definition or policy regarding sexual harassment. Critics have suggested that a more global standard of what constitutes sexual harassment may be necessary in our ever growing global economy, especially as our workforce becomes more international in nature (Efron, 1999; Orihuela and Montjoy, 2000). Although it is beyond the scope of this chapter to cover international sexual harassment law in its entirety, it is useful to note some of the differences and similarities in the laws that countries have adopted to combat sexual harassment in the workforce. Most industrialized nations have either enacted legislation to address sexual harassment in the workplace or established paths for addressing sexual harassment claims in existing legislation. Like the United States, countries that have adopted sexual harassment law through either the courts or legislation have done so relatively recently. For example, Japan decided the first hostile work environment sexual harassment case in 1992 and recently amended their Japanese Equal Employment Opportunity Law in 1997 to include a regulation prohibiting sexual harassment (Yamakawa, 1999). In Canada, sexual harassment cases have been actionable since the late 1980s and are considered a form of sex discrimination regulated under the Canadian Human Rights Act (Kelly and Watt, 1996). Other countries (most commonly found in Central Europe) have not yet adopted a law regulating sexual harassment cases specifically, and instead have chosen to encompass sexual harassment cases under the broader areas of equal

employment opportunity, labor or tort law (e.g., Greece; Magliveras, 2005; Orihuela and Montjoy, 2000).

The variation in definitions of sexual harassment is almost as great as the number of countries with laws regulating sexual harassment. For example, unlike U.S. courts, Canadian Courts have not distinguished between *quid pro quo* and hostile work environment sexual harassment and instead have adopted a broader standard that defines sexual harassment as conduct that is unwelcome, sexual in nature, and negatively affects the work environment or has negative consequences for complainants (Kelly and Watt, 1996). France is one of the only countries that considers all cases of sexual harassment under criminal rather than civil law (Hannelais, 2005). This law prohibiting sexual harassment in the workplace was passed by the French Parliament in 1992, and is relatively restrictive when compared with the definition of sexual harassment used in the United States (Hannelais, 2005).

Similar to U.S. law, some countries have included legislation that holds employers accountable for what goes on in their place of work. French law requires employers to take steps to prevent sexual harassment in the workplace (Hannelais, 2005), and Canadian employers may be held liable for failing to prevent further harassment after an initial complaint. In addition, the successful plaintiff in a Canadian sexual harassment suit is entitled to damages. However, the amount of damages awarded to plaintiffs in successful Canadian sexual harassment cases is substantially lower compared to the amount of damages awarded to plaintiffs in successful U.S. sexual harassment cases (Chotalia, 2005; Kelly and Watt, 1996).

In some countries that have enacted legislation for litigating sexual harassment claims, such legislation is rarely used, either because of a narrow definition of what constitutes sexual harassment (e.g., France), or because of required mediation or conciliation prior to litigation (e.g., Hong Kong, Japan; Petersen, 2005). In Hong Kong, for example, a Sex Discrimination Ordinance exists to govern the litigation of sexual harassment claims; however, most complaints never reach the courtroom level (Petersen, 2005). Instead, the procedure for litigating sexual harassment claims requires that the parties must attempt conciliation through the Equal Opportunities Commission before the complaint is allowed to be litigated. Initially, this procedure was conceived with the idea that such a conciliation would keep the plaintiff's complaint confidential. Critics have argued, however, that obligating complainants to negotiate with respondents may have some unintended negative effects in that doing so may cause undue stress to the complainant, exaggerate the already imbalanced power structure between complainant and respondent, and overall, may make the complainant feel less satisfied by the negotiation process as opposed to a courtroom procedure (Peterson, 2005).

Social Science Research Addressing Sexual Harassment

As the law governing the litigation of sexual harassment claims has evolved, so has the state of social science research on the antecedents and sequelae of sexual harassment. Not every worker is equally likely to experience sexual harassment during their working years. Not every person who experiences sexual harassment experiences the same consequences or exhibits the same coping strategies. Over the past two decades, social science research has provided insights into who is likely to experience sexual harassment, under what conditions, and the factors that influence physical, psychological, and behavioral responses to sexual harassment.

Case Characteristics/Antecedents of Sexual Harassment

Research on worker reports of sexual harassment incidents indicates that females are more likely than men to experience sexual harassment in the workforce (Gutek, 1985; Martindale, 1990). Bastin et al. (1996) reported in a recent large-scale survey that 64 percent of females in the armed forces reported being subjected to sexual harassment in the past year compared to 17 percent of men. In a review of numerous self-report studies, Gutek (1993) arrived at similar numbers. Specifically, an estimated 53 percent of women and 9 percent of men have been sexually harassed over the course of their careers (Gutek, 1993). This reflects the high prevalence of sexual harassment directed at female victims across occupations. For example, Slade (1994) reported that 43 percent of female attorneys surveyed reported sexual harassment by a partner in their firm. In another survey, Schneider (1987) reported that 60 percent of female graduate students claimed harassment by male professors in their program and 22 percent of female graduate students reported receiving requests for dates from male faculty members, which could be considered sexual harassment if there was an implied threat of negative repercussions if the advance was spurned. Research by Schneider et al. (1997) also indicates that sexual harassment is often not a one-time occurrence, but is rather repeated over a period of time, usually lasting one week to six months (U.S. Merit Systems Protection Board [USMSPB], 1987). Repeated offenses were reported for situations that would constitute both *quid pro quo* harassment and hostile working environments (USMSPB, 1987). Furthermore, many victims are required to have continued contact with their harassers (Schneider et al., 1997).

Several theories in the sexual harassment literature have explored why sexual harassment is disproportionately directed at female employees. The social and organizational climate of some organizations may affect the likelihood that women are targets of sexual harassment. Fiske and Glick (1995) have argued that the historical gender dominance of a job may influence the types of harassment that are likely to occur. In jobs that are considered to be traditional female occupations, men may make sexual advances toward women in pursuit of intimate relationships. Conversely, in traditionally male dominated jobs, women may encounter more hostile

forms of harassment in an effort to demean a woman who has attempted to enter a male dominated climate. Fitzgerald et al. (1997) have proposed that the likelihood that sexual harassment will occur depends on both what the researchers term the job gender context (i.e., the gender ratio and job duties typical of that occupation) and the organizational climate (i.e., the degree to which the organization tolerates sexual harassment; Naylor et al., 1980). Research results indicate that both factors have a significant influence on the likelihood of sexual harassment (Fitzgerald et al., 1997).

Some researchers posit that some men possess attitudes and characteristics that increase the likelihood that they will engage in sexually harassing behavior. Pryor et al. (1995) developed a Likelihood to Sexually Harass (LSH) Scale, which measures men's propensity to sexually harass women by assessing their reactions to a series of 10 scenarios depicting interactions between men and women. Specifically, men would rate the likelihood that they would engage in a variety of behaviors toward the woman depicted in the scenario, with one of the options describing sexually exploitative behavior (e.g., would you offer the woman a job in exchange for sexual favors). Men who scored high on the LSH scale were more likely to endorse myths about rape (e.g., women only make claims of rape when they have been rejected by a man), prescribe to stereotypical masculine traits, and endorse stereotypical sex roles for men and women. Whether these types of men will actually engage in sexually harassing behavior, however, seems to be influenced by the social climate of the workplace (Pryor et al., 1993; Pryor and Stoller, 1994). In a study conducted by Pryor and his colleagues (1993), when male participants were placed in positions of power over a female confederate and the male researcher running the experiment exhibited sexually harassing behavior (e.g., engaging in verbal flirtation, openly staring at the female confederate's body), participants who scored high on the LSH displayed more sexually harassing behaviors toward the female confederate compared to low scoring LSH participants or participants who were not exposed to a researcher who modeled sexually harassing behaviors. In another study conducted by Pryor and Stoller (1994), when high scoring LSH men were placed in groups together and provided with opportunities to establish norms accepting of sexually harassing behavior, in a later task they were much more likely than men not given such an opportunity to engage in sexual harassment of a female confederate. These studies seem to indicate that although some men have a propensity to engage in sexual harassment, the social norms deemed acceptable in their workplace will influence whether they will act on this propensity.

Although researchers generally have focused on male harassment of female employees, some researchers have examined same-sex harassment in the workforce. The majority of same-sex harassment occurs among men. When Waldo et al. (1998) studied the experience of men who claim to have been sexually harassed, the men reported that the harasser was male about 50 percent of the time, whereas women who are victims of sexual harassment are rarely victimized by a woman (less than 2 percent). The type of harassment that males receive from other male co-workers

includes subjecting the victim to lewd gestures, jokes, and comments or by degrading males who are perceived as lacking masculine characteristics (Goodman-Delahunty and Foote, 1999). Some researchers posit that these forms of sexual harassment are not as psychologically damaging to males (Goodman-Delahunty and Foote, 1999), and therefore, the consequences of sexual harassment for men in these situations are not as extreme as the consequences reported by women (Cochran et al., 1997; Gutek, 1985). Other research by Magley et al. (1999) seems to indicate that this effect is moderated by the frequency of the harassment. When men were frequently harassed, there were no significant differences between men and women in reported consequences. However, when men reported experiencing low levels of sexual harassment, they reported less negative consequences than women of the same category (see also Waldo and Magley, 1996).

Consequences of Sexual Harassment

Most women who are sexually harassed will be affected in some way by that sexual harassment. However, women experience a wide variety of consequences as a result of sexual harassment. One survey conducted by Crull (1982) revealed that 63 percent of victims of sexual harassment reported adverse physical symptoms and 94 percent of the victims reported emotional distress as a result of sexual harassment that they experienced. Specific consequences of sexual harassment may be job related, including a decrease in reported satisfaction with one's job (Gruber, 1992; Morrow et al., 1994) and relationships with co-workers (Gutek, 1985; Loy and Stewart, 1984; Morrow et al, 1994), an increase in absenteeism from work (U.S. Merit Systems Protection Board, 1981; 1987), sometimes ultimately resulting in losing one's job (Coles, 1986; Crull, 1982; Loy and Stewart, 1987). In addition, women who report having experienced sexual harassment are also more likely to be currently experiencing symptoms of depression and post-traumatic stress disorder (Dansky and Kilpatrick, 1997).

It is important to note that severe incidents of sexual harassment are not necessary for the victim to experience adverse consequences. Although research by Schneider and her colleagues (1997) indicates that more severe, ongoing reports of sexual harassment are associated with more severe consequences, even women who report relatively mild, low frequency forms of sexual harassment still experience adverse consequences for their job performance and psychological well-being compared to women who did not report exposure to sexual harassment. These adverse effects were significant even when the researchers controlled for other factors that could affect well-being such as general job stress and negative affect.

Whether a victim of sexual harassment files a complaint with the EEOC or brings forth a suit in court, they are likely to receive a favorable decision about one-third of the time (Terpstra and Baker, 1988; 1992). In both types of complaints, plaintiffs were more likely to be successful if the harassment was severe, there were witnesses to the harassment, and the victim had notified management that such behavior

was occurring. Research by Terpstra and Baker (1992) found that two additional factors were predictive of favorable findings in trial proceedings: whether there was documentation in support of the allegations and if the company did not previously take any action.

Coping Strategies of Sexual Harassment Victims

To categorize the various responses of actual victims of sexual harassment, Fitzgerald (1990) coded the survey responses of women who had previously reported an incident of sexual harassment and grouped their reported responses to the harassment into 10 different categories. Half of the coping strategies were more broadly categorized as internally focused coping strategies and the other half were described as externally focused coping strategies. Victims employing internally-focused coping strategies focus on managing their own cognitive and emotional responses to the harassment. These strategies include enduring the harassment and doing nothing and simply enduring the behavior, denying that the behavior would constitute sexual harassment, detaching from the situation, blaming themselves for the situation, and reattribution or thinking about the situation in such a way that it could not be categorized as sexual harassment. On the other hand, those victims who cope using externally focused strategies focused on solving the problem. These included actively avoiding contact with the harasser, appeasement by using non-confrontational strategies such as humor and excuses, enlisting social support from family and friends, directly confronting the situation, and employing institutional/organizational support.

Research by Phillips et al. (1989) on the frequency of the various coping strategies indicates that a sizable percentage of victims do nothing in response to the harassing behavior, either though endurance or denial. The most commonly employed externally focused strategies include avoidance (Gutek, 1985) and seeking social support — 68 percent of the surveyed respondents in one study sought support from a co-worker and 60 percent turned to family and friends for support (Phillips et al., 1989). Victims of sexual harassment are least likely to seek industrial/organizational relief. When such measures are taken, complainants avail themselves of the least aggressive reporting options such as reporting the incident to a supervisor rather than filing a formal complaint with the EEOC (Fitzgerald, Swan and Fischer, 1995). In one survey conducted by Schneider and her colleagues (1997), 17 to 33 percent of women had spoken with a supervisor about the behavior, while only 6 to 14 percent had taken the more extreme action of filing a formal complaint.

Several factors moderate the type of coping strategy that victims use. For example, the severity of the harassment is correlated with the victim's response strategy (USMSPB, 1981; Livingston, 1982). When the behavior of the harasser is particularly severe, the victim is more likely to use externally focused strategies such as confronting the aggressor or filing a formal complaint. When the harassment is more subtle and less directly threatening, the victims are more likely to respond with internal, cognitive coping strategies such as denial or appeasement through humor.

Factors Affecting Juror Decision-Making in Sexual Harassment Cases

The primary focus of research on laypeople's decisions about sexual harassment has been on the effects of various factors on how people decide whether particular behaviors constitute sexual harassment both within and outside of a trial context (Wiener and Gutek, 1999). Only recently have investigations begun to assess how laypeople might award damages in these situations. In the subsequent sections, we will review the research addressing how these various factors affect whether people categorize behaviors as sexual harassment, whether jurors find a defendant liable for discrimination due to sexual harassment, and the amount of compensatory and punitive damages that jurors award.

Gender Effects in Juror Decision-Making about Sexual Harassment

Research investigating perceptions of sexual harassment consistently finds gender differences in the way that men and women perceive potentially harassing behavior. Across several different levels of ecological validity, a variety of stimuli, and participant types, women are more likely than men to perceive various types of socio-sexual behavior in the workplace as harassing (Blumenthal, 1998; Kovera et al., 1999; Rotundo et al., 2001; Wiener et al., 1995). Wiener and his colleagues (1995), in one of the earliest studies investigating this gender difference, had participants read one of two written fact patterns based on actual hostile work environment cases, including jury instructions for making a decision in the case. They found that female participants were more likely than male participants to determine that the plaintiff had been subjected to a hostile work environment and to rate the conduct in question as unwelcome, severe, and pervasive. More recently, Kovera and her colleagues (1999) have more closely mimicked the conditions of an actual trial, and presented college student participants with a videotaped simulation of a sexual harassment trial and required them to render a verdict as if they were actual jurors. Again, this study demonstrated a gender difference in juror perceptions of sexual harassment; women were more likely than men to render a verdict in favor of the plaintiff. Meta-analyses by Blumenthal (1998) and Rotundo and colleagues (2001) have confirmed this overall significant effect of gender on perceptions of sexual harassment.

Recently, researchers have begun to investigate different factors that may affect the gender difference in perceptions of sexual harassment. In their meta-analysis examining the effect of participant gender on perceptions of sexual harassment, Rotundo and her colleagues investigated whether the magnitude of the difference in male versus female perceptions of sexual harassment was moderated by the type of socio-sexual behavior perceived. They found that the differences in perception were greater for hostile work environment sexual harassment, derogatory attitudes toward women, dating pressure, and physical sexual contact compared to sexual propositions or sexually coerced acts (Rotundo et al., 2001). Thus, when the sexually

harassing behavior is clear-cut, gender differences are smaller than when the behavior in question is ambiguous.

Moreover, the magnitude of the gender differences in perceptions of sexual harassment may be larger in community samples versus student samples (O'Connor et al., 2004). Specifically, in this study, gender differences in community members' perceptions of sexual harassment were greater than gender differences in college students' perceptions. The researchers hypothesized that this difference in magnitude could be attributed to the difference in work experiences of the two samples. That is, the women in the community member sample would have more experience in a working environment compared to the women in the college student sample. Therefore, the female community members may have been more able to conceptualize the environment experienced by the plaintiff compared to women in the college student sample. However, this moderation of gender differences as a function of community member versus college student sample disappeared in a second study using the same fact pattern but using a videotaped stimulus as opposed to a written scenario (O'Connor et al., 2004).

Models of Juror Decision-Making in Sexual Harassment Cases

The gender of the rater as a determinate of decision-making may not provide a complete picture of what is influencing a juror in the decision-making process. Specifically, researchers have sought to identify factors that may be mediating the relationship between juror gender and verdict. In Wiener and Hurt's (2000) investigation of participants' evaluations of co-worker behavior, self-referencing, or imagining what one would have done in the same situation, mediated the relationship between gender and verdict in ambiguous cases of sexual harassment. Women endorsed higher levels of self-referencing, which increased their findings for the plaintiff. In a trial simulation study, O'Connor et al. (2004) examined a variety of possible mediating mechanisms by exposing mock jurors to one of several trial stimuli and gathering measures of several possible mediating mechanisms, including ratings of benevolent and hostile sexist attitudes (Fiske and Glick, 1995), ratings of the plaintiff's credibility, and reports of juror self-referencing.

In the best fitting model, hostile (but not benevolent) sexist attitudes mediated juror gender and self-referencing, and complainant credibility partially mediated self-referencing and determinations of sexual harassment. The O'Connor et al. (2004) model explained the self-referencing effect by proposing that juror self-referencing affected jurors' evaluations of the plaintiff's credibility. The authors hypothesized that women were better able to relate to the experiences of the victim (i.e., they were able to empathize with the plaintiff, thereby affecting the plaintiff's credibility) and used more self-referencing in their decision-making. Participants who were better able to understand the plaintiff's reaction to the harassing behavior judged the plaintiff to be more credible and reasonable and subsequently judged the

behavior in question to be less welcome than participants who were less likely to self-reference.

Greathouse and Kovera (2005) further tested the mediating roles of self-referencing and complainant credibility in research examining the role of expert testimony in sexual harassment cases. Community member mock-jurors read a sexual harassment trial transcript that contained varied types of expert testimony. Although some participants were not exposed to any expert testimony, the remaining participants read a trial containing one of three types of expert testimony. Some participants read testimony proffered by the plaintiff's expert that addressed the psychological damages suffered by the plaintiff. Other participants read testimony from the plaintiff's expert and an opposing expert who argued that the plaintiff did not suffer psychological damage. The remaining participants read both the plaintiff and defense expert testimony and additional testimony from the opposing expert stating that the plaintiff had previously suffered from childhood sexual abuse and that this abuse caused the plaintiff both to solicit the sexual advances (going to liability) and to overreact to any sexual behavior that was directed toward her (going to damages).

In addition to replicating the mediating role of sexism, self-referencing, and plaintiff credibility in the relationship between juror gender and verdict, this trial simulation demonstrated that the type of expert testimony moderated the mediating effect of juror self-referencing. Specifically, when compared to no expert testimony conditions, juror self-referencing was a significant mediator for two conditions: expert testimony solely for the plaintiff and traditional expert testimony for both the plaintiff and the defense. However, when the opposing expert for the defense also included testimony concerning the prior childhood sexual abuse of the plaintiff, the path through juror self-referencing was not significant. It may be that participants were able to relate most to the plaintiff when there were no claims of specific psychological disorders. It may also be that when both experts testified with opposing opinions, participants negated both of the experts' testimony and used self-referencing, or put themselves in the plaintiff's shoes, to come to a decision. On the other hand, when previous sexual abuse history was associated with the plaintiffs, participants were not able to self-reference. Thus, the fact that self-referencing and plaintiff credibility mediated the effects of variables other than gender speaks to the generalizability of these two mediating variables in sexual harassment decision-making.

In addition to modeling juror decision-making in sexual harassment cases through examining the various mediators between the exogenous variables (gender, type of case, type of expert testimony) and liability decisions, researchers have relied upon the Story Model of juror decision-making (Pennington and Hastie, 1988; 1990; 1992; 1994) to explain how jurors make decisions in sexual harassment cases. Briefly, the Story Model posits that jurors construct a plausible story or stories based on their preexisting attitudes, general world knowledge, and the evidence and facts presented during trial to assist them in determining the events that occurred for the behavior in question. According to the Story Model, after jurors learn the

verdict options available, the rendered verdict will be the verdict that best matches the juror's constructed story of the event in question (see Chapter 3, in Volume I for a more elaborate discussion of the Story Model).

The majority of research on the Story Model has been conducted within the context of a criminal trial. According to Finkel (1995; 1997), jurors possess prototypes of typical crimes and will use these prototypes to provide context to the event. These prototypes may be used in story construction. In addition, jurors may also use contextual factors in the decision-making process that may not be legally relevant, such as situational and motivational factors (Finkel, 1997). The prototypes that jurors possess about a certain crime and the context they provide may influence jurors' interpretation of the evidence presented to create a story which in turn influences their ultimate decision (Casper et al., 1989; Olsen-Fulero and Fulero, 1997).

To test whether jurors hold prototypes that influence decision-making in sexual harassment cases, Huntley and Costanzo (2003) asked mock jurors who attended focus groups for a national trial consulting firm to listen to attorney arguments in one of four sexual harassment cases and to provide feedback concerning the strongest and weakest arguments. Jurors indicated their agreement with the various themes presented as well as their verdict choice. After categorizing the open-ended responses provided by the jurors, the researchers identified two divergent stories across the four different cases, one whose themes seemed supportive of the plaintiff (e.g., the plaintiff was a good employee who was fearful of losing her job) and one supportive of the defendant (e.g., the plaintiff was oversensitive). Jurors' constructed themes correlated with their verdict preference.

In the second phase of the research, Huntley and Costanzo (2003) used the categories derived in the first phase to construct a post-trial questionnaire asking jurors to endorse pro-plaintiff or pro-defendant themes. The questionnaire was scored so as to create a story score for each participant that reflected the extent to which the jurors endorsed either of the stories found in the first phase of this research. When the story scores were entered into a model of juror decision-making, they were significant mediators of the relationship between juror gender and verdict preference. Specifically, women were more likely to endorse a pro-plaintiff explanation of the events which, in turn, was predictive of verdict. The model containing the story scores as a mediator explained a significantly higher percentage of the variance than a model only containing the predictor variables and outcome variables.

The Reasonable Person vs. the Reasonable Woman: The Effects of the Legal Standard on Juror Decision-Making in Sexual Harassment Cases

As previously discussed, in most jurisdictions, jurors determining liability are instructed to find for the plaintiff if they determine that the behavior was unwelcome and that a reasonable person would have found the actions of the harasser to be so pervasive and severe that the plaintiff would have felt he or she worked in a hostile

environment (*Rabidue v. Osceola Refining Co.*, 1986). In *Ellison v. Brady* (1991), the 9th Circuit Court of Appeals altered jurors' instructions so that instead jurors were to adopt the perspective of a reasonable woman. The court reasoned that men and women differ in their perceptions of sexual harassment and providing jurors with the reasonable woman instruction would provide male jurors with the proper framework to evaluate the claim. Although the courts did not cite any empirical research to support their decision, social science research has begun to provide answers to both of the assumptions behind the *Ellison* decision. We have already established that men and women perceive sexual harassment differently. So, does using a reasonable woman standard diminish this gender effect? Generally, the answer is no.

At the juror level, using a reasonable woman standard generally produces similar verdicts and gender effects as using a reasonable person standard (e.g., Gutek et al., 1999; Wiener et al., 1995; Wiener and Hurt, 2000). For example, in one series of five studies, Gutek and her colleagues (1999) investigated the effect of legal standard on juror decisions in sexual harassment cases. In each of the subsequent studies, the stimulus used became more ecologically valid, ranging from a brief written summary to a videotaped trial. Each of the studies manipulated the legal standard to be used by the jurors and measured the sex of the juror. The researchers predicted that using the reasonable woman standard as opposed to the reasonable person standard would result in one of two outcomes: either a) the legal standard would have a direct effect in that all jurors would be more likely to render a verdict in favor of the plaintiff in cases with a reasonable woman standard vs. a reasonable person standard, or b) the legal standard would interact with gender of the juror to affect verdict, in that it would only sensitize jurors who might not normally consider the behavior in question from the perspective of a woman (e.g., jurors with either sexist attitudes or who were male), thus eliminating the gender difference in juror decision-making about sexual harassment cases. Overall, there was very little impact of the legal standard on juror decisions, accounting for less than 2 percent of the variance overall. Thus, it appears that the reasonable woman standard does not have the intended effect of remedying the gender gap in jurors' decisions about sexual harassment judgments.

Even though research suggests that juror decisions are unaffected by changes in legal standard, some research by Perry et al. (2004) suggests that using the reasonable woman as opposed to the reasonable person standard may change decision-making at the judicial level (Perry et al., 2004). In this study, researchers analyzed judges' opinions in 124 sexual harassment cases. The researchers coded for legal standard in two ways: 1) whether the judge explicitly mentioned which standard he or she used in decision-making and 2) the legal precedent for which standard the judge should have been using in the case. After controlling for case characteristics, judge gender, and the year in which the case was decided, researchers reported a weak relationship between legal standard (in the form of precedent used) and outcome. In the average hostile work environment case, a plaintiff was 3.17 times more likely to prevail in cases in which the precedent in the court was the reasonable woman standard than if the precedent in the court was the reasonable person standard (Perry

et al., 2004). Thus, even though the effect of the legal standard appears to have little effect for juror decision-making, using the reasonable woman standard opposed to the reasonable person standard does seem to have a small direct effect on judicial decision-making.

The Effects of Expert Testimony on Juror Decision-Making in Sexual Harassment Cases

It is possible that the reason that using the reasonable woman standard does not lessen or eliminate the gender difference in jurors' decisions about sexual harassment is that jurors do not understand the difference between how a "reasonable person" versus a "reasonable woman" would perceive the situation in question (Gutek et al., 1999). Goodman-Delahunty (1999) suggests that expert testimony explaining the standard may help delineate these differences for jurors, and thus help jurors use the standard properly in decision-making. Some research suggests that this might be a successful strategy. Kovera and her colleagues (1999) investigated the effects of social framework testimony on juror decision-making in a hostile work environment sexual harassment case. In the videotaped trial stimulus, researchers manipulated whether jurors were presented with expert testimony that described the influence of sexualized material on men's behavior toward women. The presence of expert testimony did not affect women's liability judgments; however, men were more likely to find for the plaintiff in those conditions in which an expert was present compared to those conditions in which an expert was absent. The gender difference in decision-making about expert testimony was not eliminated completely by the expert testimony, but it was attenuated. If expert testimony on factors that affect the likelihood of sexual harassment attenuates the gender difference in sexual harassment decision-making, it is possible that it may also have the potential to help jurors apply the reasonable woman standard. More research is needed to better understand whether expert testimony can help jurors use the reasonable woman standard properly.

In addition to assisting jurors in understanding the standard used to make sexual harassment decisions, some researchers have suggested that expert testimony should be used to examine plaintiffs and determine whether the plaintiff had been sexually abused as a child (Feldman-Schorrig, 1995). This abuse defense (Fitzgerald et al., 1999), based on Feldman-Schorrig's narrative review of the literature on correlates of childhood sexual abuse, argues that those who are sexually abused as children are more likely to elicit sexually harassing behavior later on in life. Thus, whether someone was sexually abused as a child would affect the welcomeness of the sexual harassment, in that if one was sexually abused and therefore elicited the sexually harassing behavior, the behavior would not be considered unwelcome (Feldman-Schorrig, 1995). In addition, Feldman-Schorrig (1995) argued that a woman who had been sexually assaulted in the past would be hyper-sensitive to any sexual behavior, and therefore would not be considered a reasonable person or woman. Several

researchers have argued based on the empirical literature that these assertions are based on faulty logic and are fundamentally untrue according to the empirical literature in the area (Fitzgerald et al., 1999; Stockdale et al., 2002). Further, the proposition that a woman who had experienced prior abuse would be hyper-sensitive to sexual behavior and would therefore not be considered a reasonable woman erroneously relates the objective standard to the subjective standard in sexual harassment cases (Kovera and Cass, 2002). Whether the woman was hyper-sensitive to abuse relates to the subjective standard of whether the woman had experienced adverse effects from the behavior in question, not whether a reasonable person or woman would find the behavior in question harassing. However, even if we overlook this flaw in logic, research suggests that those women who have experienced victimization in the past may not be any more sensitive than those women who have not (Stockdale et al., 2002). In this study, researchers reviewed the literature investigating the effects of prior sexual abuse on judgments of sexual harassment, and conducted a study examining the effects of experiencing prior sexual harassment on decision-making in sexual harassment cases (Stockdale et al., 2002). In their review of the literature, they found no evidence of a consistent relationship between those who had been sexually abused and subsequent judgments or perceptions in sexual harassment scenarios. In addition, in their empirical work, prior experience of sexual harassment had no effect on subsequent decisions in a sexual harassment case.

Even though empirical research has demonstrated that the abuse defense rests on flawed logic and is considered by most in the field to be wholly inaccurate, it is still possible that defense attorneys may attempt to call experts that will present this defense to jurors (Kovera and Cass, 2002). Thus, it is important to assess whether jurors are able to correctly assess, weigh, and discount this erroneous testimony. To date, only one study has investigated the effects of the "abuse excuse" defense on juror decision-making (Greathouse and Kovera, 2005). Greathouse and Kovera found when jurors heard testimony from a defense expert evoking the "abuse defense," they were less likely to find for the defendant compared to jurors who heard testimony from either traditional opposing experts or solely an expert testifying for the plaintiff, suggesting that not only is the abuse defense illogical but also does not have the desired effect on juror judgments.

Juror Common Understanding of the Consequences of Sexual Harassment

Some experts have suggested that plaintiffs should claim ordinary or garden variety damages (e.g., embarrassment, sadness, etc.), as opposed to injuries (e.g., Posttraumatic Stress Disorder, Depression) to avoid placing their mental health in a state of controversy, which could trigger a defense motion to compel the plaintiff to submit to a mental health examination by a defense expert (Fitzgerald et al., 1999). For plaintiffs to be successful with such a strategy, it is important that jurors understand the consequences of sexual harassment in the workplace for plaintiffs (Kovera and Cass, 2002). Research thus far has demonstrated that laypersons see

distress caused by gender-based harassment as similar to distress caused by everyday stressors (Lees-Haley et al., 1994). However, low severity gender-based harassment that is frequent can cause severe harm (Piotrkowski, 1998; Schneider et al., 1997). This suggests that there may be a discrepancy between layperson's beliefs about the harm caused by sexual harassment and actual harm experienced by the target of sexual harassment.

To test this possible discrepancy, researchers surveyed experts and individuals reporting for jury duty to examine perceptions of harm experienced by targets in several different workplace scenarios (e.g., gender discrimination, *quid pro quo* sexual harassment, hostile work environment sexual harassment, physical injury, and a work stress situation; Levett et al., 2005). They then compared jurors' beliefs about harm in the stressful workplace situations to experts' opinions about harm in the stressful workplace situations to assess if jurors accurately perceived the harms experienced by the target in those situations. Compared to experts, jurors did not underestimate the harm experienced by a target in sexual harassment scenarios more than they underestimated harm in other scenarios. However, jurors consistently underestimated the amount of harm experienced by targets across situation type. Specifically, compared to experts, jurors rated the targets in the stressful workplace situations as less likely to experience emotional damages (e.g., anxiety, feelings of withdrawing), physical damages (e.g., disrupted sleep, sexual problems), and work related damages (e.g., absences, decreased productivity). Thus, it appears that the litigation strategy of claiming garden variety damages and not calling an expert to testify about damages may be more harmful than helpful, as jurors do not generally understand the negative consequences of sexual harassment (Levett et al., 2005).

Juror Decisions about Damage Awards in Sexual Harassment Cases

In those cases in which the defendant is found liable, jurors are also charged with the task of assigning the proper compensatory damages (designed to restore plaintiffs to their state prior to the injury; Cather et al., 1996; Goodman-Delahunty and Foote, 1995), and if applicable, punitive damages (designed to punish the defendant for especially reprehensible behavior and to act as a deterrent for others; Greene and Loftus, 1998; Title VII of the Civil Rights Act, 1991). To date, most of the research concerning juror decision-making in sexual harassment cases has examined the liability decision-making phase, and how jurors assign damages has been somewhat neglected (Kovera and Cass, 2002). Therefore, we know little about how juries award damages in sexual harassment cases (Sunstein and Shih, 2004). In civil cases generally, jurors generally award damages appropriately (Greene and Bornstein, 2003). However, in most studies investigating juror decision-making about damage awards, the injury is a physical injury, and therefore the damages experienced because of such an injury may be more readily apparent to jurors compared to a psychological injury like those experienced by plaintiffs in sexual harassment cases (Cass et al., under review). One study that investigated damage awards in sexual

harassment cases by examining archives of sexual harassment cases reported that awards in sexual harassment cases were seemingly random (Sunstein and Shih, 2004). In a replication and extension of that study, Sharkey reported a positive relationship between punitive and compensatory damages and that cases governed by the Civil Rights Act of 1991 were awarded statistically significantly more damages compared to cases not governed by the Act (Sharkey, 2006). However, given the archival analyses in these studies, it is difficult to determine from these studies what influences jurors' allocations of damage awards in sexual harassment cases.

In another study, researchers presented individuals reporting for jury duty with the summary of a sexual harassment case that was calibrated so that the majority of jurors would believe the defendant was liable so they could more directly test juror decision-making about damage awards in sexual harassment cases (Cass et al., under review). In this study, researchers manipulated the severity of the harassment experienced by the plaintiff, the organization's behavior (i.e., the response of the company to the plaintiff's complaint and the description of the company's sexual harassment policy), and measured the gender of the juror. Jurors were provided the option to award compensatory and/or punitive damages to the complainant. Results demonstrated that participants' perceptions of whether the work environment was hostile did not influence damage awards in accordance with the law (perceptions of the hostile work environment should only affect liability decisions, not damage awards). However, the manipulation of harassment severity influenced compensatory damage awards, in that participants awarded higher pain and suffering awards in the high severity condition compared to the low severity condition. Thus, jurors may not be awarding compensatory damages in sexual harassment cases correctly, as the pain and suffering demonstrated by the plaintiff remained constant across conditions. Thus, harassment severity should have only influenced liability decisions, not damage awards. According to the law, only the degree of pain and suffering should influence compensatory damage awards, not the severity of the harassment. In addition, manipulations of organizational behavior correctly influenced punitive damages. Specifically, jurors in the conditions in which the company had no response to the plaintiff's complaint and no official sexual harassment policy awarded higher punitive damages to the plaintiff compared to jurors in the conditions with an enforced sexual harassment policy. Finally, juror gender had no effect on damage award decisions (Cass et al., under review). This is particularly interesting in light of the consistent gender effect found in juror liability decisions about sexual harassment. More research is needed to ascertain further what affects juror decisions in these cases and how those decisions might be improved so that jurors make consistent and legally appropriate decisions about damages in sexual harassment cases.

Conclusions

Since sexual harassment in the workforce was first recognized by the legal system as a legitimate issue faced by employees and mechanisms were put into place so that victims could seek redress, the psychological community has produced a significant body of research addressing the antecedents and consequences of harassment. We now know that sexual harassment occurs with alarming frequency, that certain workplace situations combined with specific personality characteristics seem to promote harassing behavior more than others, and that it can have long-lasting consequences for the victims. Researchers have also begun to explore the factors that influence jurors when making determinations in sexual harassment trials. Although much of this early research on evaluations of sexually harassing behavior examined laypersons' reactions outside a trial context (e.g., Wiener et al., 1995; 1997), more recent studies have increased the ecological validity of their methods to include videotaped trials (Kovera et al., 1999; O'Connor et al., 2004) and community member participants (Cass et al., under review; Greathouse and Kovera, 2005; O'Connor et al., 2004).

The majority of this research has focused on understanding the origins of and remedies for gender differences in categorizing socio-sexual behaviors as harassment; more recent research has begun to address additional topics, including how jurors allocate damages for psychological harm. Indeed little is known about how jurors evaluate damages that are not physical or economic, and it is possible that the variables that affect damage awards meant to compensate for physical or economic losses may be different than those that affect damage awards for more intangible psychological harm. Continued research on juror decision-making in sexual harassment cases is important as it may provide insights into how awarding damages in these difficult cases differs from decision-making in cases in which the harms are more easily quantified.

References

Bastin, L., Lancaster, A., and Reyst, H. (1996). *Department of Defense 1995 Sexual Harassment Survey*. Arlington, VA: Defense Manpower Data Center.

Blumenthal, J.A. (1998). The reasonable woman standard: A meta-analytic review of gender differences in perceptions of sexual harassment. *Law and Human Behavior*, 22, 33–57.

Casper, J.D., Benedict, K., and Perty, J.L. (1989). Juror decision making, attitudes, and the hindsight bias. *Law and Human Behavior*, 13, 291–310.

Cass, S., Levett, L.M., and Kovera, M.B. (under review). *The Effects of Harassment Severity and Organizational Behavior on Damage Awards*.

Cather, C., Greene, E., and Durham, R. (1996). Plaintiff injury and defendant reprehensibility: Implications for compensatory and punitive damage awards. *Law and Human Behavior*, 20, 189–205.

Chotalia, S.P. (2005). Sexual harassment laws in Canada. *International Journal of Discrimination and the Law*, 7, 199–227.

Civil Rights Act of 1964. Pub. L. No. 88-352, 78 Stat. 241.

Civil Rights Act of 1991. Pub. L. No. 102-106, U.S.C. Cert. 1981 nt.

Cochran, C.C., Frazier, P.A., and Olson, A.M. (1997). Predictors of responses to unwanted sexual attention. *Psychology of Women Quarterly*, 21, 207–26.

Coles, F.S. (1986). Forced to quit: Sexual harassment complaints and agency response. *Sex Roles*, 14, 81–95.

Costanzo, M., and Costanzo, S. (1992). Jury decision making in the capital penalty phase: Legal assumptions, empirical findings, and a research agenda. *Law and Human Behavior*, 16, 185–201.

Crull, P. (1982). Stress effects of sexual harassment on the job: Implications for counseling. *American Journal of Orthopsychiatry*, 52, 539–44.

Dansky, B.S., and Kilpatrick, D.G. (1997). Effects of sexual harassment. In W. O'Donohue (ed.) *Sexual Harassment: Theory, Research, and Treatment*. Needham Heights, MA: Allyn and Bacon.

Efron, J.M. (1999). Comment: The transnational application of sexual harassment laws: A cultural barrier in Japan. *University of Pennsylvania Journal of International Economic Law*, 20, 133–77.

Ellison v. Brady, 924 F.2d 872 (9th Cir. 1991).

Feldman-Schorrig, S. (1995). Need for expansion of forensic psychiatrists' role in sexual harassment cases. *Bulletin of the American Academy of Psychiatry and the Law*, 23, 512–22.

Finkel, N.J. (1995). *Commonsense Justice: Jurors' Notions of the Law*. Cambridge, MA: Harvard University Press.

Finkel, N.J. (1997). Commonsense justice, psychology, and the law: Prototypes that are common, senseful, and not. *Psychology, Public Policy, and Law*, 3, 461–89.

Fiske, S.T., and Glick, P. (1995). Ambivalence and stereotypes cause sexual harassment: A theory with implications for organizational change. *Journal of Social Issues*, 51, 151–66.

Fitzgerald, L.F. (1990). Sexual harassment: The definition and measurement of a construct. In M.A. Paludi (ed.) *Ivory Power: Sexual Harassment on Campus*. Albany, NY: State University of New York Press.

Fitzgerald, L.F., Buchanan, N.T., Collinsworth, L.L. Magley, V.J., and Ramos, A.M. (1999). Junk logic: The abuse defense in sexual harassment litigation. *Psychology, Public Policy, and Law*, 5, 730–59.

Fitzgerald, L.F., Drasgow, F., Hulin, C.L., Gelfand, M.J., and Magley, V.J. (1997). Antecedents and consequences of sexual harassments in organizations: A test of an integrated model. *Journal of Applied Psychology*, 82, 578–89.

Fitzgerald, L.F., Swan, S., and Fischer, K. (1995). Why didn't she just report him? The psychological and legal implications of women's responses to sexual harassment. *Journal of Social Issues*, 51, 117–38.

ForsterLee, L., Horowitz, I.A., and Bourgeois, M.J. (1993) Juror competence in civil trials: Effects of preinstruction and evidence technicality. *Journal of Applied Psychology*, 78, 13–21.

Goodman-Delahunty, J. (1999). Civil law: Employment and discrimination. In R. Roesch, S.D. Hart, and J.P. Ogloff (eds), *Psychology and Law: The State of the Discipline*. New York: Kluwer Academic/Plenum. pp. 277–337.

Goodman-Delahunty, J., and Foote, W.E. (1995). Compensation for pain, suffering, and other psychological injuries: The impact of *Daubert* on employment discrimination claims. *Behavioral Sciences and the Law*, 13, 183–206.

Greathouse, S.M., and Kovera, M.B. (March, 2005). Does self-referencing mediate the effects of expert testimony on judgments in a sexual harassment case? Paper presented at the annual meeting of the American Psychology-Law Society, La Jolla, CA.

Greene, E., and Bornstein, B.H. (2003). *Determining Damages: The Psychology of Jury Awards*. Washington, DC: American Psychological Association.

Greene, E., and Loftus, E.F. (1998). Psycholegal research on jury damage awards. *Current Directions in Psychological Science*, 7, 50–54.

Gruber, J.E. (1992). A typology of personal and environmental sexual harassment: Research and policy implications for the 1990s. *Sex Roles*, 26, 447–64.

Gruber, J.E., and Bjorn, L. (1986). Women's responses to sexual harassment: An analysis of sociocultural, organizational, and personal resource models. *Social Science Quarterly*, 67, 814–26.

Gutek, B.A. (1985). *Sex and the Workplace: The Impact of Sexual Behavior and Harassment on Women, Men, and Organizations*. San Francisco, CA: Jossey-Bass.

Gutek, B.A. (1993). Responses to sexual harassment . In M. Costanzo and S. Oskamp (eds). *Gender Issues in Contemporary Society*. Newbury Park, CA: Sage.

Gutek, B.A., and O'Connor, M. (1995). The empirical basis for the reasonable woman standard. *Journal of Social Issues*, 51, 151–66.

Gutek, B.A., O'Connor, M.A., Melancon, R., Stockdale, M.S., Geer, T.M., and Done, R.S. (1999). The utility of the reasonable woman legal standard in hostile work environment sexual harassment cases: A multimethod, multistudy examination. *Psychology, Public Policy, and Law*, 5, 596–629.

Hannelais, J. (2005). Sexual harassment law in France. *International Journal of Discrimination and the Law*, 7, 187–98.

Harris v. Forklift Systems, Inc., 510 U.S. 17; 114 S.Ct. 367 (1993).

Huntley, J.E., and Costanzo, M. (2003). Sexual harassment stories: Testing a story-mediated model of juror decision-making in civil litigation. *Law and Human Behavior*, 27, 29–51.

Kelly, J.M., and Watt, B. (1996). Damages in sexual harassment cases: A comparative study of American, Canadian, and British law. *New York Law School Journal of International and Comparative Law*, 16, 79–134.

Kovera, M.B., and Cass, S.A. (2002). Compelled mental health examinations, liability decisions, and damage awards in sexual harassment cases: Issues for jury research. *Psychology, Public Policy, and Law*, 8, 96–114.

Kovera, M.B., McAulliff B.D., and Hebert, K.S. (1999). Reasoning about scientific evidence: Effects of juror gender and evidence quality on juror decision is a hostile work environment case. *Journal of Applied Psychology*, 84, 362–75.

Lees-Haley, P.R., Lees-Haley, C.P., and Price, J.R. (1994). A sexual harassment-emotional distress rating scale. *American Journal of Forensic Psychology*, 12, 39–54.

Levett, L.M., Kovera, M.B., and Goodman-Delahunty, J. (2005, March). Juror common understanding of workplace harm. Paper presented at the meetings of the American Psychology-Law Society, La Jolla, CA.

Livingston, J.A. (1982). Responses to sexual harassment on the job: Legal, organizational and individual actions, *Journal of Social Issues*, 38, 5–22.

Loy, P.H., and Stewart, L.P. (1984). The extent and effects of sexual harassment of working women. *Sociological Focus*, 17, 31–43.

Magley, V.J., Waldo, C.R., Drasgow, F., and Fitzgerald, L.F. (1999). The impact of sexual harassment on military personnel: Is it the same for men and women? *Military Psychology*, 11, 283–302.

Magliveras, K.D. (2005). The regulation of workplace sexual harassment in Greece: Legislation and case law analysis. *International Journal of Discrimination and the Law*, 7, 169–86.

Martindale, M. (1990). *Sexual Harassment in the Military: 1988*. Arlington, VA: Defense Manpower Data Center.

Meritor Savings Bank v. Vinson, 477 U.S. 57 (1986).

Morrow, P.C., McElroy, J.C., and Phillips, C.M. (1994). Sexual harassment behaviors and work related perceptions and attitudes. *Journal of Vocational Behavior*, 45, 295–309.

Naylor, J.C., Pritchard, R.D., and Ilgen, D.R. (1980). *A Theory of Behavior in Organizations*. New York: Academic Press.

O'Connor, M., Gutek, B.A., Stockdale, M., Geer, T.M., and Melancon, R. (2004). Explaining sexual harassment judgments: Looking beyond gender of the rater. *Law and Human Behavior*, 28, 69–95.

Olsen-Fulero, L., and Fulero, S.M. (1997). Commonsense rape judgments: An empathy-complexity theory of rape juror story making. *Psychology, Public Policy, and Law*, 3, 402–27.

Orihuela, S., and Montjoy, A. (2000). The evolution of Latin America's sexual harassment law: A look at mini-skirts and multinationals in Peru. *California Western International Law Review*, 30, 323–44.

Pennington, N., and Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14, 521–33.

Pennington, N., and Hastie, R. (1990). Practical implications of psychological research on juror and jury decision making. *Personality and Social Psychology Bulletin*, 16, 90–105.

Pennington, N., and Hastie, R. (1992). Explaining the evidence: Tests of the story model for juror decision making. *Journal of Personality and Social Psychology*, 62, 189–206.

Pennington, N., and Hastie, R. (1994). The story model for juror decision making. In R. Hastie (ed.), *Inside the Juror*. New York: Cambridge University Press. pp 192–224.

Perry, E.L., Kulik, C.T., and Bourhis, A.C. (2004). The reasonable woman standard: Effects on sexual harassment court decisions. *Law and Human Behavior*, 28, 9–26.

Petersen, C.J. (2005). Negotiating respect: Sexual harassment and the law in Hong Kong. *International Journal of Discrimination and the Law*, 7, 127–68.

Piotrkowski, C.S. (1998). Gender harassment, job satisfaction, and distress among employed white and minority women. *Journal of Occupational Health Psychology*, 3, 33–43.

Phillips, C.M., Stockdale, J.B., and Joeman, L.M. (1989). *The Risks of Going to Work*. London: The Suzy Lamplugh Trust.

Pryor, J.B., Giedd, J.L., and Williams, K.B. (1995). A social psychological model for predicting sexual harassment. *Journal of Social Issues*, 51, 69–84.

Pryor, J.B., LaVite, C., and Stoller, L. (1993). A social psychological analysis of sexual harassment: The person/situation interaction. *Journal of Vocational Behavior*, 42, 68–83.

Pryor, J.B., and Stoller, L. (1994). Sexual cognition processes in men who are high in the likelihood to sexually harass. *Personality and Social Psychology Bulletin*, 20, 163–9.

Rabidue v. Osceola Refining Co., 805 F.2d. 611 (6th Cir. 1986).

Rotundo, M., Nguyen, D.H., and Sackett, P.R. (2001). A meta-analytic review of gender differences in perceptions of sexual harassment. *Journal of Applied Psychology*, 86, 914–22.

Schneider, B.E. (1987). Graduate women, sexual harassment, and university policy. *Journal of Higher Education*, 58, 46–65.

Schneider, K.T., Swan, S., and Fitzgerald, L.F. (1997). Job-related and psychological effects of sexual harassment in the workplace: Empirical evidence from two organizations. *Journal of Applied Psychology*, 82, 401–15.

Sharkey, C.M. (2006). Dissecting damages: An empirical exploration of sexual harassment awards. *Journal of Empirical Legal Studies*, 3, 1–45.

Slade, M. (1994, February 26). Law firms begin reining in sexually harassing partners. *The New York Times*, p. B12.

Stockdale, M.S., O'Connor, M., Gutek, B.A., and Geer, T. (2002). The relationship between prior sexual abuse and reactions to sexual harassment. *Psychology, Public Policy, and Law*, 8, 64–95.

Sunstein, C.R., and Shih, J. (2004). Damages in sexual harassment cases. In C.A. MacKinnon and R.B. Siegel (eds), *Directions in Sexual Harassment Law*. New Haven, CT: Yale University Press.

Terpstra, D.E., and Baker, D.D. (1988). Outcomes of sexual harassment charges. *Academy of Management Journal*, 31, 185–94.

Terpstra, D.E., and Baker, D.D. (1992). Outcomes of federal court decisions on sexual harassment. *Academy of Management Journal*, 35, 181–90.

Turic v. Holland Hospitality, Inc., 201 F.R.D. 85 F.3d 1211 (6th Cir. 1996).

U.S. Merit Systems Protection Board (1981). *Sexual Harassment in the Federal Workplace: Is It a Problem?* Washington, DC: U.S. Government Printing Office.

U.S. Merit Systems Protection Board (1987). *Sexual Harassment in the Federal Workplace: An Update*. Washington, DC: U.S. Government Printing Office.

Waldo, C.R., Berdahl, J.L., and Fitzgerald, L.F. (1998). Are men sexually harassed? If so, by whom? *Law and Human Behavior*, 22, 59–79.

Waldo, C.R., and Magley, V.J. (1996, May). Psychological and job-related outcomes of the sexual harassment of men. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Wiener, R.L., and Gutek, B.A. (1999). Advances in sexual harassment research, theory, and policy. *Psychology, Public Policy, and Law*, 5(3), 507–18.

Wiener, R.L., and Hurt, L.E. (2000). How do people evaluate social sexual conduct at work? A psycholegal model. *Journal of Applied Psychology*, 85, 75–85.

Wiener, R.L., Hurt, L., Russell, B., and Mannen, K. (1997). Perceptions of sexual harassment: The effects of gender, legal standard, and ambivalent sexism. *Law and Human Behavior*, 21, 71–93.

Wiener, R.L., Watts, B.A., Goldkamp, K.H., and Gasper, C. (1995). Social analytic investigation of hostile work environments: A test of the reasonable woman standard. *Law and Human Behavior*, 19, 263–81.

Yamakawa, R. (1999). We've only just begun: The law of sexual harassment in Japan. *Hastings International and Comparative Law Review*, 22, 523–60.

Index