Advances in Psychology and Law 2

Brian H. Bornstein
Monica K. Miller *Editors*

Advances in Psychology and Law

Volume 2

Springer

Advances in Psychology and Law

More information about this series at http://www.springer.com/series/11918

Brian H. Bornstein · Monica K. Miller
Editors

Advances in Psychology and Law

Volume 2

Springer

Editors
Brian H. Bornstein
Department of Psychology
University of Nebraska–Lincoln
Lincoln, NE
USA

Monica K. Miller
Criminal Justice Department
University of Nevada
Reno, NV
USA

ISSN 2366-6099 ISSN 2366-6102 (electronic)
Advances in Psychology and Law
ISBN 978-3-319-43082-9 ISBN 978-3-319-43083-6 (eBook)
DOI 10.1007/978-3-319-43083-6

Library of Congress Control Number: 2016940010

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To our students, from whom we have learned so much

Preface

We are very pleased to introduce this second volume of the *Advances in Psychology and Law* book series. As with Volume 1, the present volume helps to fill a long-standing gap in the field of law-psychology—namely, the absence of thorough reviews of existing research that can reveal policy implications and suggest important directions for future research. The chapters are virtually an honor roll of "hot topics" in psychology and law: interrogations and confessions, guilty pleas, how jurors deal with emotional or scientific evidence, gangs, and psychopathy, to mention but a few. Each chapter presents a "state of the science" summary for researchers, while also highlighting important policy implications. The chapters are deliberately interdisciplinary, integrating various psychological subdisciplines (e.g., developmental, social, cognitive), law, and other social science disciplines (e.g., sociology, criminal justice).

The overwhelming majority of convictions in the US are the result of guilty pleas. For a guilty plea to be valid, it must be made knowingly, intelligently, and voluntarily, and by defendants who are factually guilty. The first chapter on "The Validity of Pleading Guilty", by Redlich, examines the validity of pleading guilty, as well as the reliability of the methods used to ensure that guilty pleas are valid. According to her analysis, these methods, and consequently the guilty pleas themselves, often fall short of legal and moral standards.

When criminal defendants do not plead guilty, or when civil litigants fail to reach a settlement, the next step is usually a trial. The next six chapters deal with issues regarding trial procedures. First is a chapter by Yelderman, Peoples, and Miller which reviews the research on death qualification in capital trials. Jury selection in capital trials is unique in the extent to which the court explores jurors' attitudes toward case-relevant issues, namely the death penalty. Although it may seem perfectly reasonable to exclude jurors who would refuse to consider one of the relevant punishment options—for example, jurors who would oppose a death sentence under any circumstances—as Yelderman and colleagues show, death qualification calls the fairness of the entire process into question.

The following three chapters deal with specific kinds of evidence, all of which can lead jurors to make non-normative decisions. First, Nunez, Estrada-Reynolds,

Schweitzer, and Myers discuss the role of emotion in juror decision making. The law often expects jurors to act rationally and to ignore emotion, or to consider emotion only in a circumscribed way, but jurors cannot easily set their emotions aside. As the authors show, emotion in the courtroom implicates a number of general theories of emotions and decision making. The chapter by Devine and Macken addresses a type of evidence that is, in many respects, the polar opposite of emotional evidence, namely, scientific evidence, which is generally presumed to be detached, dispassionate, and objective. Yet just as with emotional evidence, jurors' comprehension of scientific evidence is limited, variable, and often inconsistent with legal guidelines. Similarly, and as explored in the chapter by Costanzo, Blandón-Gitlin, and Davis, jurors do not always appreciate the intricacies of evidence dealing with confessions and interrogation practices. The authors review the research literature on interrogations and confessions with an eye toward whether expert testimony on these issues can help jurors weigh confession evidence more appropriately. A common feature of all three of these chapters is that they offer recommendations for improving jurors' use of these different kinds of evidence.

Many evidentiary rules, which dictate the admissibility of various kinds of evidence, reflect a belief that they would protect against flaws in decision making by excluding some evidence while admitting other evidence that satisfies certain criteria. The next two chapters call some of these assumptions and practices into question. First, Kleynhans and Bornstein explain how some of the assumptions embodied in the Federal Rules of Evidence (FRE) are themselves flawed and can even result in greater disparity to the parties involved. Because evidentiary rules have such a broad scope, the authors also address areas that have not yet been examined in depth by psycholegal researchers, proposing avenues for future research. In the final evidentiary chapter, DeMatteo, Hodges, and Fairfax-Columbo review the scientific literature on the most widely used tool to measure psychopathic characteristics—the Psychopathy Checklist-Revised (PCL-R)—to examine whether it satisfies the test for the admissibility of evidence. They conclude that the PCL-R has probative value in some contexts but not others. Courts therefore need to adopt a nuanced approach to this assessment instrument.

The next chapter, on "A Synthetic Perspective on the Own-Race Bias in Eyewitness Identification", by Wilson, Bernstein, and Hugenberg, addresses the own-race bias (or cross-race effect), which refers to eyewitnesses' tendency to be better at recognizing members of their own race/ethnicity/group than persons of other races/ethnicities/groups. The own-race bias raises a number of theoretical and applied issues, most prominently, what causes it and what can the justice system do about it. Wilson and colleagues discuss the cognitive and social mechanisms underlying the bias, as well as possible interventions for reducing it.

The final two chapters focus on populations of special interest to the criminal justice system. The chapter by Heilbrun, DeMatteo, King, Thornewill, and Phillips casts a relatively broad net. They address the potential of interventions to reduce the risk of subsequent criminal offending by "justice-involved" individuals—that is, anyone who has been arrested on criminal charges. Given the high prevalence of behavioral health disorders in this population, their discussion necessarily speaks to

the debate about whether such interventions should include a behavioral health component. As Heilbrun and colleagues make clear, the best possible interventions would both improve behavioral health and reduce criminal offending risk. The final chapter deals with a smaller population, but one that is nonetheless sizeable and extremely problematic. Wood, Alleyne, and Beresford review strategies for deterring gangs and reducing their impact on communities. A number of individual and group processes can help explain why many criminal justice anti-gang tactics are ineffective; by calling attention to those processes, the chapter by Wood and colleagues contributes both to research on gangs and to public policy efforts to combat them.

Many people helped to make this book a reality: Sharon Panulla and Sylvana Ruggirello, our editors at Springer; the Springer production staff; and the book series editorial board. We are grateful for their commitment to the book series and for being such a pleasure to work with. In addition, we have continually benefited from the efforts of our students—past, present, and future. They force us to clarify our thinking, pique our curiosity, challenge our assumptions, work with us side by side in conducting and writing up research, and enrich our lives in many tangible and intangible ways. We hope that this volume will inspire the research efforts of current students, future students, and established scholars.

Lincoln, Nebraska, USA Brian H. Bornstein
Reno, Nevada, USA Monica K. Miller

Contents

The Validity of Pleading Guilty

Allison D. Redlich

> Criminal justice today is for the most part a system of pleas, not a system of trials.
> The Honorable Justice Anthony M. Kennedy, *Lafler v. Cooper* (2012).

An examination of the prevalence of guilty pleas in the U.S. criminal justice system reveals that Justice Kennedy's above sentiment is resoundingly true. Specifically, across the U.S., guilty pleas account for 65 % of all outcomes, and about 95 % of all convictions in large, urban counties (Bureau of Justice Statistics 2010). And in some jurisdictions, pleas account for even higher proportions of outcomes. For example, in two New York counties, for every 100 felony arrests, 85 result in a guilty plea, only one proceeds to trial, and the remainder are dismissed, diverted, or devolve into non-criminal violations (Redlich et al. 2013).

Although this high prevalence of pleas makes for a more efficient system (in comparison to having trials for each defendant), it is unlikely to make for a more just system. The main arguments for plea bargaining are that it empowers choice on the defendant and that it saves the courts time and money. Many contend that the U.S. criminal justice system could simply not operate without plea bargains or without the high frequency of plea bargains (e.g., Bibas 2003; Fisher 2003). The main arguments against plea bargaining are that it is not an empowering choice, but rather a coercive one that eschews the standard safeguards built into the trial process. When guilty pleas are in place, the state does not have to prove the charge against the defendant beyond a reasonable doubt, and the defendant waives his rights to the presumption of innocence, to confront persons against him, to present evidence in his defense, and most other constitutionally afforded safeguards.

Because our criminal justice system is over-capacity and overburdened, plea decisions are often made in haste without adequate knowledge or information (e.g., Bibas 2011; Lynch 2003). Some district attorney offices, like New York County's,

A.D. Redlich (✉)
George Mason University, Fairfax, VA, USA
e-mail: aredlich@gmu.edu

B.H. Bornstein and M.K. Miller (eds.), *Advances in Psychology and Law*,
Advances in Psychology and Law 2, DOI 10.1007/978-3-319-43083-6_1

practice a 'best first' plea offer policy (Kutateladze et al. 2014; see also, Rakoff 2014), which means that subsequent offers will be less favorable to defendants, a practice that helps to ensure quicker decisions. However, the United States Supreme Court has long recognized that "a guilty plea is a grave and solemn act to be accepted only with care and discernment" (*Brady v. United States* 1963, p. 748). "The Constitution insists that the defendant enter a guilty plea that is 'voluntary' and make related waivers 'knowing[ly], intelligent[ly], [and] with sufficient awareness of the relevant circumstances and likely consequences'" (*U.S. v. Ruiz*, citing *Boykin v. Alabama*, p. 623). And, for guilty pleas to be considered valid, a factual basis of guilt must be established. Even *Alford* pleas (*North Carolina v. Alford* 1970), which allow defendants to profess innocence while entering a guilty plea, must establish a factual basis of guilt via means other than the defendant's admissions (Shipley 1987).

In this chapter, I examine the validity of pleading guilty. Validity here concerns (1) knowingness and intelligence—having sufficient information to make a decision, having the ability to understand and appreciate the information, and actually understanding, (2) voluntariness—having the choice to enter a guilty plea (as well as the knowledge that one has the choice), and (3) actual guilt—being factually guilty of the charges (see also, Redlich et al. in press). Within these sections, I also identify gaps in the literature and potential ways to move the knowledge base forward with continued research. Although courts are charged with making these determinations regarding validity, to my knowledge, there is little guidance about *how* courts should make them, and as a result, the process can be quite idiosyncratic. Arguably, the methods used today fall short, leading to the question of whether many, or perhaps even the majority, of plea entries are invalid.

To illustrate what happens in plea hearings, I rely, in part, on three research studies, which I describe briefly here to reduce redundancy. First is an observation study conducted by Miller et al. (1978). Miller and his colleagues observed and coded 720 plea hearings across six U.S. cities from 1975 to 1978. Although the data collected are the most antiquated, they are also the most comprehensive (of the three). Some of the findings reported here were obtained directly from the original dataset, which is publicly available. The second study, conducted by Redlich and Bonventre (2015), analyzed the content and comprehensibility of written tender-of-plea forms used with juvenile and adult defendants across all 50 states and the District of Columbia. These forms, while certainly not replacements for the oral colloquy, are intended to supplement and represent efforts by the court to inform defendants about plea-relevant information. In theory, if defendants later attempted to argue that their pleas were invalid (e.g., not knowing, intelligent, or voluntary), their signatures on these forms could be used to contest this allegation. The third study was an interview study of 99 adult offenders post-plea (within 30 days) about their experiences and understanding of plea information (Redlich and Summers 2012).

Overview of Guilty Pleas and the Process

Before addressing the validity of plea decision-making, it is important to understand the nature of guilty pleas and the process. There are several different types of pleas that defendants can enter, including a not guilty plea, an insanity plea, and a guilty but mentally ill plea. There are also different types of guilty pleas, including the traditional guilty plea, the nolo contendere (or no contest) plea, and the *Alford* plea. No contest pleas allow defendants to enter a guilty plea without acknowledging guilt or innocence. In contrast, *Alford* pleas, after the 1970 U.S. Supreme Court case, *North Carolina v. Alford*, allow defendants to claim innocence while simultaneously pleading guilty. Across the U.S., approximately 59 % of all felony pleas are traditional guilty pleas, 11 % are no contest pleas, 7 % are *Alford* pleas, and 23 % are not guilty pleas (Redlich and Ozdogru 2009).

Guilty pleas, regardless of the type, need not be 'bargains' but most are. More specifically, defendants can simply plead guilty to the original charges without expectations of leniency. Most defendants, however, negotiate a deal in which their sentences, charges, and/or collateral consequences (e.g., having to register as a sex offender) are reduced (Redlich et al. in press). Even *Alford* plea-takers receive sentences shorter than those received if convicted at trial (Redlich and Ozdogru 2009). In large part, it is this deal (or the value of the plea) that spurs the controversy around pleas. For example, is the deal too good to turn down, even for innocent defendants?

Over and above the "deal," what other factors do defendants and other legal actors take into account when making plea decisions? Theorists have argued that plea decisions are made "in the shadow of trial" (Nagel and Neef 1979). In this "shadow of trial" model, decisions to offer, accept, or reject pleas derive from the perceived likelihood of a trial outcome (Kalven and Zeisel 1966; Landes 1971; Smith 1986). That is, in theory, prosecutors, defense attorneys, and defendants base their plea decisions on what evidence they anticipate jurors will hear and weigh. However, legal scholars such as Bibas (2004) and Stuntz (2004) have raised important questions about the face validity of the model. In its simplest incarnation, the theory assumes that actors in the criminal justices system act rationally, yet behavioral economists and psychologists now routinely show that people, including professionals in the criminal justice system, do not act in strictly rational ways (Guthrie et al. 2001; Plous 1993; Tversky and Kahneman 1974). More generally, Bibas (2004) questioned whether the complex institutional choices involved in bargaining are congruent with the simple "shadow of the trial" model.

Recently, Bushway and Redlich (2012; Bushway et al. 2014) set out to test the "shadow of the trial" model. First, using data collected by Miller et al. (1978), Bushway and Redlich (2012) found support for the model, at the aggregate but not at the individual level. More specifically, in the aggregate, the average plea sentence for the sample was equivalent to the average sentence at trial discounted by the probability of conviction for the sample (which was 76.7 % of the defendants who proceeded to trial and were convicted), as would be expected by the theory.

However, at the individual level, in many cases, a defendant's actual plea value was not at all similar to the estimate of that defendant's discounted probability of a sentence at trial. Further, individual estimates of the probability of conviction at trial for those who pled guilty were virtually uncorrelated with key pieces of evidence known to increase the probability of conviction at trial, such as confessions. Second, in a separate study, Bushway et al. (2014) demonstrated that whereas prosecutors and defense attorneys bargained in the shadow of the trial, judges did not. Rather, judges were found to offer fixed plea discounts (from sentences that would be received via trial convictions) without much consideration of the probability of conviction.

In regard to defendant decision-making, there is also evidence indicating that defendants are not always rational actors, especially when innocent (see Tor et al. 2010). For example, in line with expected utility theory (as well as the "shadow of the trial" model), defendants should accept plea offers that maximize outcomes and reject those that do not. However, Bordens (1984) questioned whether defendants make plea decisions that are merely satisficing (a hybrid term indicating a compromise between sufficient and satisfying) rather than optimizing. Norris and Redlich (2014), using the Prisoner's Dilemma paradigm, demonstrated that under certain conditions, plea takers do make plea decisions that are merely satisficing. In brief, the Prisoner's Dilemma, first introduced by two RAND scientists and then formalized by Tucker (1950), is a paradigm involving two suspects, separated from each other, and questioned. The dilemma is whether to confess or deny contingent on what the co-suspect is expected to do; this decision is also tied to higher or lower sentences. Norris and Redlich found that when both co-defendants were innocent, 87 % chose an outcome with a sentence that was 5 years longer than an alternative outcome, which could be viewed as satisficing rather than optimizing.

If attorneys, judges, and defendants are not basing their plea offers/advisements/decisions on their expectations of trial outcomes, the question then becomes what are their decisions based on? Although beyond the scope of this chapter, there is some evidence that extra-legal factors come into play. For example, Edkins (2011) found race of the defendant to influence defense attorneys' plea recommendations. Specifically, she demonstrated that defense attorneys, responding to a hypothetical case, recommended plea bargains with longer sentences for African American clients in contrast to Caucasian ones. Edkins stated "Robert Williams, the African American client, was more than three times more likely to be encouraged to accept a plea that included jail time than Robert Williams, the Caucasian American client" (p. 424). Of importance, Edkins did not find that perceived guilt or perceived probability of conviction at trial differed by defendant race, thus indicating that the longer sentence plea offers were not explainable by these factors.

In summary, most criminal convictions in the U.S. are the result of guilty pleas. The predominant, but largely untested (especially for defendants), theory of plea decision-making is the "shadow of the trial" model. Whereas there is some evidence that a portion of plea decision-making may be made in this manner, there is also evidence suggesting that it often is not. Regardless of the rationale behind the plea

decision, the primary emphasis of this chapter is whether the plea decision is valid—that is, knowing, intelligent, voluntary, and with a factual basis of guilt. Understanding why people plead guilty, however, can help inform the validity of pleas.

Determining the Validity of Pleading Guilty

The main mechanisms used to determine knowing, intelligent, and voluntary plea decision-making are via the oral judicial plea colloquy (or inquiry), which must be preserved in the record (*Boykin v. Alabama* 1969), and written tender-of-plea forms. Federal Rule of Criminal Procedure 11 governs the entry of guilty and nolo contendere pleas in federal court, including provisions designed to ensure that defendants' pleas are valid. Most states follow similar procedures (Acker and Redlich 2011). Rule 11, Part b, "Considering and accepting a guilty or nolo contendere plea," in subparts 1–3, utilizes the word 'must' numerous times. Specifically, the "court *must* address the defendant personally in open court," "the court *must* inform the defendant of, and determine that the defendant understands" a list of 15 rights and other aspects for a knowing and intelligent plea, "the court *must* … determine that the plea is voluntary and did not result from force, threats, or promises (other than promises in a plea agreement)," and "the court *must* determine that there is a factual basis for the plea" (emphases added). The American Bar Association (ABA 1999) also delineates standards for guilty pleas and recommends that the court should reject the plea if defendants did not have effective assistance of counsel and/or if defendants lack "complete understanding of the consequences."

In attempting to adhere to Rule 11 and the *Boykin* mandate, judges typically ask defendants a series of questions or read a series of statements. This colloquy is aimed at determining, for example, whether any promises other than the plea offer itself were made (voluntary), whether the defendant understands the legal rights he is ceding and the possible consequences associated with the plea (knowing and intelligent), and whether the defendant is presently incapacitated (e.g., on medication, intoxicated). Some judges will also ask defendants to allocute to the details of the crime, essentially providing in-court confessions.

Typically, judicial colloquies are described as "boilerplate" (Bibas 2011) and "perfunctory" (Turner 2006; p. 212) in that judges may simply read verbatim off of a standard form without deviation. And, presumably, some judges exert more influence than others. For example, the Colorado Supreme Court overturned a case in which the judge was "improperly involved" in the plea discussions and who later refused to allow the defendant to withdraw his guilty plea. In this case, *James M. Crumb v. The People of State of Colorado* (2010), the judge implied that he would be lenient if the defendant agreed to enter a plea and later stated that he was "not going to be a happy judge" if a plea deal had not been negotiated.

Depending on the judge, there is more or less reliance on competent counsel to explain and provide relevant plea information. However, as argued by Lang (2012),

having "competent counsel does not negate the need for a Fifth Amendment plea colloquy ensuring voluntariness" (p. 952) (and vice versa), indicating that defense attorneys' explanations, no matter how complete, cannot (and should not) stand in for the oral exchange on the record between the judge and the defendant. Nonetheless, some judges appear to replace, or at least greatly reduce, the exchange with defendants in lieu of defense attorney explanations. For example, in the Redlich and Bonventre (2015) study of tender-of-plea forms, more than half the forms made reference to prior explanations from defense attorneys, leading the authors to question whether judges rely on competent attorneys to explain plea information who may, in turn, rely on the forms themselves (which were found to be largely incomprehensible and non-comprehensive).

Finally, it is important to question whether the self-reported answers of defendants and defense counsel alike sufficiently demonstrate validity. If, for example, the judge asks a defendant "Do you understand the charges against you?" and the defendant answers "Yes, sir, I do" the required determination has been met. However, once defendants have negotiated a plea and accepted the offer informally prior to coming before the judge, it is highly probable that some defendants will answer the questions in the plea discussion on the basis of ensuring the plea bargain goes through, rather than on the basis of the truth of the matter. In regard to juvenile defendants, Kaban and Quinlan (2004) made specific note that reliance on such answers to indicate knowing, intelligent, and voluntary decision-making is ill-advised. Further, there is a wealth of empirical research demonstrating that individuals who commit to a position are much more willing to comply with future requests for the same or similar positions, especially when the commitments are made publicly (see Cialdini 2001), such as in open plea hearings.

Thus, what does it mean to make a valid plea decision? "The most common definition of validity is epitomized by the question: Are we measuring what we think we are measuring?" (Kerlinger and Lee 2000; p. 666). Are the judicial colloquy and written tender-of-plea forms meant to test the validity of plea decisions? Is their intent to measure whether plea decisions are voluntary, knowing, and intelligent? These questions are important to answer as an estimated 14,500 defendants every day are faced with the decision to plead guilty or go to trial (Lynch 2003). In theory, every defendant would have the capacity to understand and appreciate the information. Every defendant would be fully informed. Every defendant would make a voluntary choice. And every defendant considering a plea offer would be factually guilty. Every defendant would have competent counsel who had, and then took, the time to explain the information related to and consequences of the plea decision. And every defendant would have the benefit of time to deliberate about the decision. In the sections that follow, the evidence for whether pleas are (1) knowing and intelligent, (2) voluntary, and (3) made by factually guilty individuals is presented.

Knowing and Intelligent Plea Decision-Making

In addressing whether plea decisions are informed, it is important to address three aspects separately. First is whether defendants have the *capacity* for making a knowing and intelligent decision. Second is whether defendants *actually* understand and appreciate the information even if they have the capacity to do so, and third, is whether defendants receive complete and understandable information to make a truly knowing and intelligent decision.

Capacity for Valid Plea Decision-Making

Defendants are presumed competent to stand trial, that is, are presumed to have "sufficient present ability to consult with his lawyer with a reasonable degree of rational understanding and a rational as well as factual understanding of the proceedings against him" (pp. 402–403; *Dusky v. U.S.* 1960). In *Godinez v. Moran* (1993), the U.S. Supreme Court explicitly rejected the notion that competence to stand trial (CST) and competence to plead guilty require different standards. However, in the majority opinion, Justice Thomas distinguished between what he called the 'competence inquiry' and the 'knowing and voluntary' inquiry. He stated "The focus of the competency inquiry is the defendant's mental capacity; the question is whether he has the *ability* to understand the proceedings. The purpose of the 'knowing and voluntary' inquiry, by contrast, is to determine whether the defendant actually *does* understand the significance and consequences of a particular decision" (emphases in original; see Footnote 12). In this section, I focus on the capacity to make valid plea decisions, providing reasons why it is likely that many more defendants than those assessed have deficits in competence.

Assessments of competence are the most common procedure involving the mental health and criminal justice systems (Advokat et al. 2012). Thus, not surprisingly, CST has been the subject of a great deal of research. Recently, Pirelli et al. (2011) conducted a meta-analysis of 40 years of research on CST. In brief, they found that persons diagnosed with psychosis were about eight times more likely to be found incompetent than those without the diagnosis, and that unemployment status and previous hospitalizations affected findings of incompetence. Pirelli et al. also reviewed the multiple instruments developed to assess CST. Although there are differences between them (especially in administration methods), to some degree the instruments attempt to assess defendants' legal abilities and/or links between legal and psychological abilities (Zapf and Viljoen 2003).

All defendants are presumed to be competent to proceed against criminal charges waged against them. It is only when a defendant's competence is questioned that an assessment of ability is conducted. Thus, it is entirely possible that there are defendants who are technically incompetent but allowed to plead guilty because their competence status was never questioned. To wit, in earlier research, my

colleagues and I found that among newly enrolled mental health court clients who pled guilty as a condition to enter the court, 13–27 % demonstrated clinically significant impairments on the understanding portion of a competence measure (Redlich et al. 2010).

Although there are varied reasons for potential incompetence, the two primary reasons to raise questions of competence are mental illness and mental impairment (retardation, developmental disabilities) (Pinals 2005). It has been clearly established that persons with serious mental illness and impairment are (1) overrepresented in the criminal justice system compared to the general population, and (2) have significant deficits in their ability to assist in their defense and understand the proceedings (Redlich and Cusack 2010). It is estimated that at least one million persons with serious mental disorders (i.e., schizophrenia/schizo-affective, psychotic, bipolar, and major depressive disorders) are arrested annually (Redlich and Cusack 2010). And, one comprehensive study demonstrated that 14.5 % of male and 31 % of female jail inmates have a serious mental disorder (Steadman et al. 2009). In contrast, only 1–2 % of the general population has such disorders. Further, up to 10 % of offenders have developmental disabilities, a rate which is about three times more than in the general population (Wall and Christopher 2012).

Competence assessments are estimated at 50–60,000 a year (Mossman et al. 2007; Poythress et al. 1994). Among those whose competence is questioned, about 20–60 %, depending upon the severity and type of mental disability, are found to be incompetent (Everington and Dunn 1995; Mossman et al. 2007). At the conservative end of the spectrum (i.e., 20 %), this means that there are an estimated 12,000 individuals whose competence is questioned per year found to be incompetent. The number of people found incompetent and deemed unrestorable (i.e., unable to become competent) is less clear. Although most individuals are able to be restored to competence, there are characteristics identified to be associated with the inability to be restored to competency: specifically, long-term psychosis with histories of inpatient hospitalization and incompetence stemming from cognitive disorders (such as mental retardation) (Mossman 2007). Indeed, one study found that two-thirds of defendants with mental retardation deemed to be incompetent could not be restored (Anderson and Hewitt 2002).

Regardless of the number of people found to be restorable, the pertinent number for the analysis here is the number found to be incompetent, i.e., the conservative estimate, 12,000 of those who were referred for assessments. Restoration is a lengthy process, taking 6 months to a year (or longer) requiring consistent education, treatment, and medication. If defense attorneys, for example, were able to restore defendants to competency, then it is unlikely that the defendants were truly incompetent in the first place. That is, competence assessments take into account the person's *capacity* to partake in their defense, to have a rational as well as factual understanding of the legal process, etc. Defendants undergoing assessments need not initially know the roles of the judge and attorneys, for instance, but they must demonstrate the capacity for learning this information. If the person demonstrates that he can learn (e.g., recite back in his own words what he just learned about the

roles of the judge), then generally this person has met the threshold of competence (see Bonnie 1992; Poythress et al. 2002).

Overall, there appears to be a disconnect between the number of defendants potentially eligible to have their competence assessed and the number who actually do. As mentioned, there are an estimated one million individuals with a serious mental illness arrested per year. If there are 50,000 competence assessments annually, only about 5 % (50,000/1,000,000 arrestees) are assessed.[1] Only 12,000 or 20 % are found to be incompetent. Thus, if 20 % are found to be incompetent and if all 1,000,000 were assessed, then 200,000 would be incompetent. Even assuming that only the 'worst of the worst' get referred for competence assessments, and allowing for a lower rate of 10 % incompetence finding, then 100,000 defendants would be incompetent to proceed. If 65–85 % of all defendants plead guilty (see above), then an estimated 65,000–85,000 (or 130,000–170,000 using the 20 % rate) defendants with serious mental illness *each year* are so severely disabled that they do not have the capacity to make a knowing and intelligent decision, waive their rights and enter guilty pleas, pleas presumably deemed valid by the courts as the pleas went through, and were accepted.

If we extend this reasoning to defendants with mental retardation, the problem is exacerbated. Rates of incompetence with this population are notably higher at 60 %. Approximately, 10 million people are arrested annually, and an estimated 10 % have mental retardation. So, again, using the figure of 1 million, 600,000 would be found incompetent (60 %), and up to 510,000 (85 %) would have pleaded guilty.

These calculations only account for defendants with serious mental illness and mental retardation. They do not account for incompetence due to young age, traumatic brain injuries, other cognitive disorders, and long-term drug use. As in criminal courts, in juvenile court, plea rates dominate, accounting for 99 % of convictions in one court (see Kaban and Quinlan 2004). Seminal research by Grisso et al. (2003) has also made it quite clear that significant minorities of youth do not reach the standards of competence set by the courts.

To summarize, a significant minority of defendants are unlikely to have the capacity to enter into valid plea agreements. Competence is presumed present, and there are reasons to suspect that hundreds of thousands of defendants annually would be found incompetent to proceed if the question of competence was raised. Even among the defendants who meet the low threshold required to be considered

[1]Another issue to consider is why only an estimated 5 % of defendants who have potential competence problems are referred for competence assessments. Reportedly, defense attorneys are sometimes loath to raise the issue of competence for their clients despite indications of deficits. When dealing with clients with mental disabilities, a common dilemma faced by attorneys is to get their client out of jail, often on time served, or to raise the issue of competence which in turn starts a lengthy process of assessment during which the client may or may not remain incarcerated, and may or may not be found restorable to competence (see Pinals 2005). Thus, in efforts to zealously represent their clients, defense attorneys may opt for their clients to be free (non-incarcerated) rather than ensuring due process is met.

legally competent (in a formal sense), they may lack the ability to understand and appreciate plea information provided in oral and written plea inquiries.

Understanding of Plea Information

Adjudicative competence has been said to have a low threshold. Famously, one prosecutor opined that so long as defendants can tell the difference between a grapefruit and the judge, the threshold for competence has been met (Coleman 2002). Thus, even if we assume that most defendants who plead guilty are technically competent to do so, a next-step question would be Does the average defendant have the capacity to understand what the judges and lawyers tell him specifically about pleading guilty? That is, for example, is the language used by legal actors comprehensible to defendants? If defendants are asked more than yes/no questions about their understanding, could they demonstrate adequate comprehension?

When answering the question of whether defendants understand plea information, it is important to consider that most defendants (68 %) did not finish high school, most (66 %) have learning disabilities (Wolf Harlow 2003), and as discussed, significant minorities have mental health problems and cognitive deficits. In addition, legal terms are often a foreign language (i.e., 'legalese') to persons not trained in the law. The little available research on judicial plea colloquies does not delve into the exact language used by judges, but presumably there is some proportion of judges who use technical legal terms unrecognizable to most defendants. To illustrate, the Benchbook for U.S. District Court Judges (2013), in Sect. 2.01, *Taking pleas of guilty or nolo contendere*, includes guidelines and a list of sample questions for judges to use in colloquies. An analysis of these sample questions revealed a required reading grade level of 13.3 (i.e., college level). For example, the first question to ask defendants is "Do you understand that you are now under oath and if you answer any of my questions falsely, your answers may later be used against you in another prosecution for perjury or making a false statement?" Given the educational and intellectual deficiencies present in the majority of defendants, it is probable that many would not truly understand this first question, as their oral comprehension capacity, that is, their ability to listen, understand, and appreciate the meaning of information conveyed verbally, is likely to be low. Of course, this is important because, as mentioned, the main method to determine if defendants' pleas are valid is through the oral plea colloquy by the judge. Although there is limited research on defendants' oral comprehension, one study found that the average listening comprehension ability of recently arrested jail detainees was 9th grade (Blackwood 2009).

More research has been conducted on defendants' reading comprehension abilities. According to the National Adult Literacy Survey (Haigler et al. 1994), most (70 %) adult inmates read at or below the 6th grade level. Juvenile offenders have similarly low (or even lower) reading capabilities (e.g., Hodges et al. 1994).

Redlich and Bonventre (2015) analyzed the reading comprehensibility of tender-of-pleas, the written form of the oral colloquy. They found in their sample of 208 forms that, on average, both juvenile and adult defendant plea forms required an 11th–12th grade reading level to fully understand the content. Only 4.3 % of the forms in their sample of forms could be understood by those who read at a 6th grade level or lower.

In regard to actual comprehension, there is a general dearth of research on defendant understanding and appreciation of the information supplied in plea colloquies or written forms. However, there are four studies that provide some insight. First is a study by Kaban and Quinlan (2004), who questioned Massachusetts juveniles aged 9–17 years awaiting adjudication about their understanding of 36 words used in the plea form (e.g., disposition, restitution). They found that, even with instruction and consideration of prior experience with the legal system, on average, only five words (14 %) were correctly defined. Many times the juveniles claimed to know the meaning of the word but when tested, it was clear that they did not. Second, Redlich and Shteynberg (2016) measured the vocabulary comprehension of words used in a tender-of-plea form among study participants aged 13–24 years involved in a hypothetical plea scenario. Although the young adults accurately defined significantly more words than youth, it was also clear that the adults lacked adequate comprehension. For example, more than 90 % of youth and adults could not accurately define the word 'concession' (as used in the legal context), despite claiming that they (1) understood the plea form, and (2) had no questions.

Third, Bordens and Bassett (1985) interviewed 67 (adult) defendants who had pled guilty, examining what they called 'knowledge of plea information.' They concluded that "In general, the data show that defendants have a knowledge [*sic*] of the information necessary to make an informed decision" (p. 106). However, the knowledge examined was awareness of the sentence to be received if the plea was accepted or rejected, and the likelihood of conviction at trial. Whether participants' awareness of the sentences and likelihood was actually *accurate* is unclear, and other indications of understanding (e.g., that the plea was a voluntary choice, that rights were being ceded) were not measured. Further, not all participants indicated awareness; for example, only 78 % claimed to have been aware of the sentence if the plea had been accepted.

Fourth, Redlich and Summers (2012) interviewed recent plea-takers; when asked directly about legal understanding, 89 % said they understood the plea process, 96 % said they understood the possible penalties associated with their plea, and 87 % said they understood the legal proceedings. Nevertheless, when their comprehension of plea information was assessed, two-thirds of the sample (people who had just entered valid pleas) answered 60 % or more of questions incorrectly.

To summarize, even if defendants can meet the low threshold of competence, the possibility remains that they do not understand nor appreciate the information given to them. As stated by Bibas (2011), "There are two distinct barriers to informed decision-making: first, defendants must have enough information; and second, they must be able to understand, digest, and use that information" (p. 85).

Sufficiency of Information to Make Informed Plea Decisions

In *Boykin v. Alabama* (1969), the court stated "the utmost solicitude of which courts are capable in canvassing the matter with the accused to make sure [the defendant] has a *full* understanding of what the plea connotes and of its consequence" (emphasis added; pp. 243–44). However, the word 'full' in this instance is misleading as there are systematic and non-systematic reasons indicating that complete understanding is not required. For one, 'consequence' concerns only direct, and not collateral, consequences (see Roberts 2009). In large part, direct consequences refer to penal sanctions, including jail/prison sentences, probation, and fines. Collateral consequences are varied and many, and include revocations of, or limitations on future access to, licenses (e.g., driving, professional), access to firearms, voting privileges, housing and welfare benefits, and effects on future charges and sentences. Collateral consequences need not be made known to defendants who are deciding whether to accept plea offers, with one recent exception. In *Padilla v. Kentucky* (2010), the Supreme Court held that defense counsel must inform defendants of their risk of deportation if they plead guilty. However, defense counsels who choose to make collateral consequences known to their clients must provide accurate information. The provision of inaccurate information on collateral consequences could result in assessments of inadequate defense counsel. As argued by Roberts (2009), the fact that not informing defendants about the possible consequences of pleading guilty is constitutional whereas choosing to inform clients but providing misadvice is not, which creates a perverse incentive for defense counsel to simply keep their mouths shut and remain silent.

Although the Supreme Court does not require knowledge of collateral consequences (except deportation, as mentioned, and some subsequent extensions; see Lang 2012) as part of a valid plea decision, there are many who believe that defendants should be aware of such consequences, and that this knowledge may lead to a different plea decisions. For example, the American Bar Association (1999) recommends in Standard 14-1.4(c):

> Before accepting a plea of guilty or nolo contendere, the court should also advise the defendant that by entering the plea, the defendant may face additional consequences including but not limited to the forfeiture of property, the loss of certain civil rights, disqualification from certain governmental benefits, enhanced punishment if the defendant is convicted of another crime in the future, and, if the defendant is not a United States citizen, a change in the defendant's immigration status. The court should advise the defendant to consult with defense counsel if the defendant needs additional information concerning the potential consequences of the plea.

Additionally, the courts have tended toward limiting the type of information that defendants must have to make valid decisions. For example, in the *U.S. v. Ruiz* (2002) decision, a case involving the duty to provide impeachment information (relating to the credibility of informants) prior to plea decision-making, the Court put forth that defendants need not have complete knowledge when deciding to plead guilty, stating "[T]he Constitution does not require the prosecutor to share all

useful information with the defendant" (p. 629). Although the primary issue in *Ruiz* was voluntariness (as opposed to knowing and intelligence), the Court ruled that knowledge of potential impeachment information was relevant to the fairness of a trial but not to the voluntariness of plea decisions. Further, as argued by Taslitz (2008) and others (e.g., Bibas 2004), in general, discovery rules are weak, and most defendants are left in the dark as to the evidence the state has when making decisions about pleading guilty. Put simply, "the myth of the fully informed rational actor" is one of the biggest and most problematic assumptions made in our justice system (Bibas 2011; p. 80).

What information then is given to defendants during plea hearings? Although systematic data are generally lacking on the comprehensiveness and content of judicial colloquies, there has been some research. As a whole, the studies demonstrate that the amount and content of information supplied to defendants is highly variable. For example, in the Redlich and Summers (2012) interview study, 99 individuals who pleaded guilty within the past 30 days (across eight judges) reported on what information they had been asked or told as part of their plea. All reported being asked if they understood the charges against them (100 %), and nearly all reported being asked if it was their choice to plead guilty (96 %), and that they would be giving up certain Constitutional rights (95 %). More nuanced information was reportedly asked/told less frequently, including what evidence the prosecution had against them (54 %), if they were intoxicated on alcohol, drugs, or medications (46 %), and that the judge did not have to accept the sentence recommendation of the prosecutor (86 %). Amount and type of information also varied between judges (except for the three pieces of information near ceiling level), all of whom were in the same jurisdiction. For instance, the proportion of defendants who reported being asked if they were presently intoxicated on drugs, alcohol, or medications ranged from 10 % with one judge to 100 % with another judge. Moreover, when asked if someone had explained to them the conditions of the plea, 30 % claimed they received an explanation on the day of their plea, 3 % claimed receipt after taking their plea, and 7 % claimed never to have received an explanation.

Miller et al. (1978) observed and systematically coded plea colloquies. They found that of 720 plea hearings across six cities, length of plea hearings in total ranged from less than 1 min to about 1 h. However, the average length was 8.04 min (SD = 7.83), the median length was 5 min and the modal length was 3 min. Recently, Boruchowitz et al. (2009) opined that for misdemeanor defendants, the plea and sentencing can take "less time than it takes to get a hamburger from a McDonald's drive-through window" (p. 32).

Miller and colleagues also found that 11 % of the plea hearings had no inquiry and 9 % had the defendant sign a form only. Most of the oral inquiries were also what Miller et al. labeled as 'standard' as opposed to 'individualized.' As shown in Table 1, Miller et al. observed that the judge was primarily the one to ensure the components of a valid plea were met. However, the last column of Table 1 is revealing. In about one-quarter of the hearings, no one reviewed the defendant's Constitutional rights, asked if the defendant understood the rights s/he was waiving, or explained the charges. In 41.4 % of the hearings, whether the defendant

understood the charges was not addressed. Moreover, the overwhelming majority of hearings—94 %—included no mention of collateral consequences. In a similar study (albeit much smaller), Sanborn (1992) observed plea hearings in juvenile courts in different settings. He found that whereas the judicial plea colloquy used by urban judges tended to be more comprehensive (especially for those 14 years and older charged with felonies or weapon–misdemeanor offenses) and lasted 5–10 min, colloquies by suburban and rural judges were much less complete and lasted only 1–2 min. Sanborn also reported that the courts "ignored the voluntary and intelligent nature of the guilty plea," assuming instead that the defense lawyer had fulfilled this role (p. 139).

Finally, in Redlich and Bonventre's (2015) analysis of the content of 208 tender-of-plea forms representing 160 counties and 48 states, they coded aspects related to (1) voluntariness, (2) rights, (3) knowing and intelligence, (4) collateral consequences, and (5) an 'other' catch-all category. Regarding the rights and knowing and intelligence components, a combined total of 81 aspects were systematically coded. On average, the forms included mention of nine rights [e.g., right to (speedy, public, bench) trial, right to presumption of innocence] and seven aspects related to knowing and intelligence (e.g., plea has same effect as being tried and found guilty, possible court costs/fines/restitution, whether defendant understands English). A review of Federal Rule 11, Part 1 reveals that some but not all forms make mention of what must be made known to defendants. For instance, 74 % of forms mentioned the right to a jury trial, but only 35.6 % mentioned the right to plead not guilty. Most forms covered the nature of the charges (75.5 %) and the possible penalties (79.3 %). However, fewer covered the right to counsel (45.7 %), the right to testify and present evidence (41.8 %), and the court's authority to order restitution (35.1 %). For all 81 components, accompanying explanations were quite rare. And, most often, the forms presented the information as statements (82 %) to be read, rather than as questions to be answered.

Regarding collateral consequences (which again is not required with the exception of the recent *Padilla* and subsequent rulings), Redlich and Bonventre (2015) found that about one-quarter of the forms did not mention any collateral consequences. Of the remaining forms that did include mention, most focused on criminal justice consequences rather than social welfare ones. For example, parole/probation violations, effects on future offenses and sentences, and registering as a sex offender appeared in 35–60 forms (of 208). In contrast, consequences around employment, welfare benefits, child custody, student loans, and voting were present in fewer than 10 forms. None of the forms mentioned ineligibility for public housing as a possible consequence of the plea. The average number of collateral consequences mentioned per form was about two.

Overall, the data available on plea hearings demonstrate that most take place in 10 min or less. During these minutes-long exchanges, it appears that the information supplied is highly variable; most defendants, but not all, are likely to be asked the basics, that is, what is minimally required (do they understand the charges, their rights). The specifics of exactly what rights are being waived, for instance, appear to be less likely to be reviewed.

Table 1 Miller et al. (1978) data of observed plea hearings

	Judge[a] (%)	Defense (%)	Prosecutor	Other	Not recited/asked
Who recited constitutional rights	57.7	15.6	0	2.2 %	24.4
Who asked defendant if understood rights	65.1	8.4	<1 %	<1 %	26.2
Who explained charges	57.9	2.8	7.8 %	1.1 %	30.2
Who asked if defendant understood charges	54.2	2.7	1.8 %	<1 %	41.1
Who asked if other promises had been made	24.0	7.8	0	0	68.2
Who asked if threatened or coerced	45.9	8.9	0	0	45.1

[a]Judge percentages combine when the Judge recited/asked alone or with other people (defense counsel or the court administrator)

Needed Research on Knowing and Intelligent Plea Decision-Making

Above, I reviewed three aspects of a knowing and intelligent plea decision and the related decision-making process; one, whether defendants have the capacity to understand and appreciate plea-related information (i.e., competence); two, whether defendants actually do understand and appreciate the information; and three, whether defendants are given sufficient information to make a knowing and intelligent plea decision. In regard to the first, the psychological research on competence can be described as plentiful. Scholars have studied competence for more than forty years (see Pirelli et al. 2011), and the topic remains a robust area of inquiry. Nonetheless, there is still more research that needs to be done, especially research specific to pleas. As indicated above, there are reasons to suspect that thousands of plea-takers every year do not have the requisite abilities to make knowing and intelligent legal decisions. Research should continue to explore this prevalence, as well as determine the precise reasons why legal actors, such as defense attorneys, do not raise questions of competence more often. Further, more in situ research is needed on what actually occurs in plea hearings and the degree to which defendants truly understand what it means to plead guilty; that is, what rights they are ceding and what direct and collateral consequences they are facing. How defendants receive plea information, from whom, and how often are questions that are also unclear and in need of empirical attention. Finally, in regards to the third aspect, how much information is needed to make a fully knowing and intelligent decision should be researched. Would the dichotomous decision change if defendants had more information than typically given? Or would the incentives built into the plea bargaining system be too tempting to alter the plea decision, even if more information was provided?

Voluntary Plea Decision-Making

In legal contexts, it is not always clear what is meant by 'voluntary' and its opposite, 'coercive,' despite the fact that voluntariness is such an integral, defining aspect of our legal system. This lack of definitional clarity becomes quite important when discussing pleas. In 2002, in *United States v. Speed Joyeros, S.A.*, the court stated "There is no single clear definition of 'voluntary' for all legal purposes. Even in the criminal-law-plea context, it is unclear whether 'voluntary' means freedom from any coercion or whether it means freedom only from 'wrongful' or 'undue' coercion. A pristine rule of 'no coercion' would preclude many plea agreements" (p. 14). There are a variety of social influence factors that can serve to limit the perceived voluntariness of plea decision-making, including time-limited offers, obedience to authority, and the value of the bargain itself (see Redlich et al. in press).

When addressing the voluntariness of plea decisions, the U.S. Supreme Court case of *Bordenkircher v. Hayes* (1978) often arises (e.g., Lynch 2003; Taslitz 2008). In this case, the prosecutor attempted to pressure the defendant into pleading guilty by offering him a deal of five years (for the crime of passing a forged check for $88.93) to "save the court the inconvenience and necessity of a trial," or he would charge the defendant as a habitual offender in which he would serve a mandatory life sentence. Hayes opted for his Constitutional right to a jury trial but was convicted. The Supreme Court approved the prosecutor's actions and upheld Hayes' life sentence, arguing that the defendant was ultimately free to accept or reject the plea offer.

Unlike the Justices in *Hayes,* many view plea deals as inherently coercive, a Hobson's choice (Langbein 1992; Lynch 2003). The term "Hobson's choice" reportedly derives from a sixteenth century English stable owner who would offer potential buyers a choice between the horse closest to the door or no horse at all. In many ways, a plea offer offers this choice of no choice at all. That is, the choice between accepting guilt and the concomitant conviction (and all of its consequences), or risking one's fate at trial. Only about a quarter of people who proceed to trial are acquitted (Bureau of Justice Statistics 2010). The sentences associated with trial convictions are also much harsher than plea convictions, leading to phrases such as 'trial penalty' and 'trial tax' (e.g., Ulmer et al. 2010). On the other side of the debate, there are many who view plea bargains as a positive benefit for defendants, even innocent defendants (Bowers 2008). Indeed, defendants who are guilty and offered a plea bargain may be quite pleased with the outcome in that they often avoid jail via time served or probation. Perceptions of procedural justice and perceived coercion have been found to be inversely correlated among plea-takers (Redlich and Summers 2012), such that those who perceive they were treated fairly and had a voice are more likely to see their plea decision as more voluntary (and vice versa).

Arguably, voluntariness is a more subjective concept than the knowing and intelligence aspects of plea information. Whereas, in theory, knowing and

intelligence can be tested via extended questioning, voluntariness is more akin to a personal feeling of having one's will overborne. According to Federal Rule 11 part 2, *Ensuring That a Plea Is Voluntary*, judges must "address the defendant personally in open court and determine that the plea is voluntary and did not result from force, threats, or promises." In the Redlich and Summers (2012) study, 96 % reported being asked if it was their choice to plead guilty, and 93 % claimed that it was indeed their voluntary choice. When asked if the option of taking a plea versus going to trial was presented to them, a small minority said they were never presented this option (6 %) and 21 % claimed this option was presented for the first time on the day of entering their plea. Nearly one-third of the sample, however, erroneously believed that someone other than themselves (e.g., the judge or their attorney) made the final plea decision *after* the court had accepted their plea. And when asked directly if guilty pleas have to be voluntary, 44 % incorrectly said 'no.' Thus, even if individuals perceive they made the choice voluntarily, they are not necessarily aware that the choice was indeed theirs to make.

In their content analysis of tender-of-plea forms, Redlich and Bonventre (2015) found that 56.7 % of the tender-of-plea forms referred to voluntariness, though very few included accompanying explanations of the term. Additionally, 64.4 % provided that the plea did not result from force or threats, and 74.1 % that the plea did not result from promises. Surprisingly, 17 % of the forms did not mention any aspects of voluntariness. Given that voluntariness is an integral part of guilty pleas, it is was an unexpected finding that nearly one in five forms did not address it. Similarly, in the Miller et al. (1978) data of observed plea hearings, in 68 % of the cases, no one had asked defendants if any promises other than the plea itself had been made, and in 45 % of the cases, no one had asked defendants about threats or coercion (see Table 1).

Self-reports of defendants also bear out that they sometimes feel pressured into pleading guilty. For example, Bordens and Bassett (1985) found that in addition to internal pressures experienced by the defendant, external pressure from the prosecutor was a predominant reason for pleading guilty. Pressure from the prosecutor included threats of charges for many crimes and severe punishment, and reprisals for not accepting the plea offer. In a more recent study, focused on youthful offenders, Malloy et al. (2014) found that 10–14 % of reported reasons for pleading guilty related to duress (see also Redlich et al. 2010).

Overall, one of the primary arguments opponents of plea bargaining have is that it is coercive by its very nature. Social science researchers adhere to a much more stringent standard of coercion in terms of human subject ethics. For example, when conducting research with prisoners, a prisoner representative must review and approve research proposals as a member of the institutional review board who oversees the researcher's study. Often times, offering prisoners the choice between participating in the research with payment and not participating is viewed as unduly coercive. The voluntariness of guilty plea decisions has a much lower standard. The review and approval that takes place typically takes the form of the judge asking the defendant if it was his choice to plead guilty, defendants who may or may not know it is their choice to make.

Needed Research on Voluntary Plea Decision-Making

The U.S. criminal justice system has allowed for plea bargains for more than 150 years. Nonetheless, questions still persist about the true voluntariness of the choice. Psychologists should continue to investigate perceptions of voluntariness associated with plea offers. As found by Bordens and Bassett (1985), the pressure may be internal to the defendant or stem from a variety of external sources (e.g., prosecutor, defense attorney, family members). It would also be worthwhile for researchers to attempt to examine perceptions of voluntariness from defendants, prosecutors, and defense attorneys, and see if they align with each other. Do defenders believe that defendants' choices are voluntary when defendants themselves feel they are not, for example?

A difficult task going forward will be to distinguish between defendants' satisfaction with the leniency they enjoyed as part of the plea deal and the actual voluntariness of the decision. That is, for instance, if defendants were pressured into accepting a plea deal, they may experience cognitive dissonance about how 'good' a deal they received, and inflate the level of voluntariness and/or worth of the deal to harmonize the discordance they may (unknowingly) feel. Finally, comparing perceptions of voluntariness between true and false guilty plea-takers would be an interesting endeavor. Predictably, innocent people would be likely to see their plea decision as less voluntary. According to Tor et al. (2010), innocent persons experience feelings of substantive unfairness when faced with two negative choices: plead guilty or go to trial.

Factual Guilt and Plea Decision-Making

In this 'Age of Innocence' it is unquestionable that innocent persons have accepted guilt and pleaded guilty. Since the inception of the plea bargaining process in the mid-1800s, the possibility of innocents pleading guilty has been recognized (*State v. Kaufman* 1879). And, as noted above, the *Alford* plea is an explicit endorsement of people claiming to be innocent who would rather plead guilty than risk their fates at trial. Nonetheless, a demonstration of factual guilt is required for guilty pleas to be valid. In our criminal justice system of pleas (*Lafler v. Cooper* 2012), for the most part, the willingness to plead guilty is often conflated with factual guilt itself (Bibas 2014).

One example case of an innocent pleading guilty is Robert H., described by Alschuler and Deiss (1994). Robert H. spent 6 months in jail without being charged before first meeting his public defender, who reportedly handled more than 500 cases that year. The lawyer explained that if he pleaded guilty, Robert could go home that day or he could go to trial but spend up to another year in jail waiting. Upon the lawyer's recommendation, Robert pleaded guilty. However, it was later discovered that Robert had been confused with someone else and was innocent.

There are numerous other cases like Robert's. As of April 2015, the National Registry of Exonerations has catalogued 211 substantiated false guilty pleas since 1989, accounting for about 13 % of the exonerations in their sample. There are many reasons to believe that the true number of false guilty pleas is much larger than 211, though.

First, given that guilty pleas account for almost all (95 %) convictions, it would be difficult to fathom that the number of true false guilty pleas is not substantially larger than presently known. For this not to be the case, police and prosecutors would have to have identified and prosecuted the right person in almost every case, which is unlikely. Second, whereas guilty pleas occur most often for less serious crimes, the majority of identified exonerations are for the serious crimes of rape and murder. However, for all crimes except murder, plea rates exceed 90 % (Cohen and Reaves 2006). And still, about two-thirds of murderers plead guilty (Redlich and Ozdogru 2009). Bowers (2008) argues that most false guilty pleas occur for low-level crimes. Third, because plea bargains do generally shorten the time of incarceration or even eliminate it, the motivation to correct the injustice and be set free is not present. Fourth, guilty pleas are very difficult to withdraw and appeal (Weaver 2001–2002), especially post-sentencing, and thus the wrongful conviction may never be recognized and righted. Indeed, Malloy et al. (2014) found that only one-quarter of self-reported innocent guilty pleaders attempted to retract their plea.

Finally, there are likely to be larger numbers of false guilty pleas than those presently known because of the lack of safeguards in the plea process. When a defendant pleads guilty, institutionalized trial safeguards like burden of proof and cross-examination are absent. As stated by Redlich (2010; p. 56),

> A defendant who pleads guilty may know that an eyewitness exists against him and will consider this risk when opting to plead guilty; a defendant who goes to trial may also hold this knowledge but has the opportunity to challenge the eyewitness, bring in expert witnesses, and have a jury or judge decide the credibility of the witness…. To be sure, in the wrongful conviction cases that went to trial, the safeguards meant to identify the causal factors (e.g., eyewitness misidentification, false confessions) failed. However, in wrongful conviction cases that culminated in a plea bargain, the causal factors never had the opportunity to undergo scrutiny or challenge. As such, the factors identified as contributing to wrongful convictions by trial are likely to be even more prevalent in wrongful convictions by guilty plea.

For these reasons and others, estimating rates of false guilty pleas using samples of officially exonerated individuals is dangerously misleading (Redlich 2010). Because persons who plead guilty are, by definition, convicted, asserting actual innocence after entering a plea is notoriously difficult. Thus, those who plead falsely may never be recognized as exonerated, such as the case of Kerry Max Cook (2007). Cook spent over 20 years on death row in total. After going through three trials (in two of which he was convicted) and being faced with a fourth, the prosecutor offered Cook a deal: plead no contest in exchange for time served. The prosecutor commenting on this deal stated "The important thing for us was to insure that he got a conviction for murder that would follow him for the rest of his life" (see http://www.pbs.org/wgbh/pages/frontline/criminal-justice/cleared-

by-dna-texas-man-still-fighting-to-clear-his-name/). The conviction did follow Cook—because of his no contest plea, Cook technically remains a convicted murderer even though a mere 2 months after the plea, DNA evidence exonerated him. His case, highlighted in the popular play and movie *The Exonerated*, is not included in the National Registry of Exonerations sample of more than 1800 exonerees because only official declarations of exoneration are included. Because of Cook's no contest plea, his case did not qualify.

At the end of the day, it is likely impossible to get a true and precise count of false guilty pleas. However, because we know that some people falsely plead guilty, the important question may not be 'how many' but 'why do innocent people plead guilty'? The short and uncomplicated answer to why is that there are powerful inducements to plead guilty, particularly when pitted against remaining in jail pretrial and/or risking a conviction with a stiffer sentence at trial. As stated by Bowers (2008), "It is hardly a new observation that guilty pleas may prove attractive to the innocent" (p. 1120). For the most part, pleas are attractive because they do reduce the charges, the time in jail or prison, and potentially other consequences (e.g., registering as a sex offender). Indeed, the Minnesota statewide tender-of-plea form includes the statement "I do/do not make the claim that the fact I have been held in jail since my arrest and could not post bail caused me to decide to plead guilty in order to get the thing over with rather than waiting my turn at trial" (Redlich 2010).

Laboratory studies have also delved into the question of why innocent people plead guilty. These studies have found rates of false guilty pleas as high as 61 % (Dervan and Edkins 2013; see also Norris and Redlich 2014). In one of the first studies, Gregory et al. (1978) asked male college students to imagine they were guilty or innocent of armed robbery, listen to a tape of the defense attorney's arguments, and then reject or accept a plea offer. When the number of charges was high versus low (four vs. one charge) and when the sanctions were high versus low (10–15 vs. 1–2 years in prison), 100 % of the "guilty" participants and 33 % of the "innocent" participants pled guilty. More recent lab studies (i.e., Dervan and Edkins 2013; Russano et al. 2005) have also found that the more attractive the "deal," the more likely one is (innocent and guilty alike) to plead guilty. For example, Dervan and Edkins (2013) found that when the expected sentence at trial was lenient as opposed to harsh, fewer guilty and innocent participants were willing to accept the plea. Thus, it is clear that actual and hypothetical defendants will plead guilty to crimes they did not commit, which should, in theory, invalidate their pleas.

Though Federal Rule 11, Part 3 requires that before entering judgment on a guilty plea, the court must determine that there is a factual basis for the plea, in practice judges may or may not inquire into defendants' factual guilt. In the sample of written tender-of-plea forms, Redlich and Bonventre (2015) reported that less than half (43.8 %) included mention of factual guilt, such that defendants who checked the box and signed the forms acknowledged their guilt. Significantly, fewer forms (4.3 %) mentioned innocence explicitly. The available evidence indicates that courts either do not inquire or simply take the defendants' 'word for it' (Redlich et al. in press). In the study of 720 plea hearings from the mid-1970s,

factual guilt was reportedly "de-emphasized" (p. 277) in judges' colloquies; only 14 defendants were coded as maintaining their innocence and of these 14, the judge refused to accept the plea in only five. As stated by Miller et al. (1978):

> One court held that it was not necessary for the trial judge to interrogate a defendant in determining the factual basis. Another said the purpose of a factual basis inquiry was to determine that the defendant committed a crime at least as serious as that to which he was going to plead. … Another court held it to be a sufficient inquiry if a judge requests the prosecutor to set forth the facts and subsequently draws the attention of the defendant to the prosecutor's statement before accepting the plea. In yet another case it was acceptable for the trial judge to ask the defendant leading and suggestive questions requiring only a yes or no answer (p. 281, references omitted).

Some courts may also use the acceptance of the guilty plea itself as proof of factual guilt, and not inquire further. As stated by Justice White in the *Alford* decision, "[W]hile most pleas of guilty consist of both a waiver of trial and an express admission of guilt, the latter element is not a constitutional requisite to the imposition of criminal penalty. An individual accused of crime may voluntarily, knowingly, and understandingly consent to the imposition of a prison sentence even if he is unwilling or unable to admit his participation in the acts constituting the crime" (p. 37). However, as mentioned, judges choosing to accept *Alford* pleas must establish a factual basis of guilt via other means than the defendant's admissions. Shipley (1987) argues that typically judges only receive summaries of the evidence from the prosecution.

In summary, given the extremely high prevalence of guilty pleas in our criminal justice system, there are numerous reasons to suspect that innocent defendants accept guilt and take the deal being offered to them. By their very nature, pleas occur in secret with little oversight. But by the very fact that courts must verify factual bases of guilt, the courts acknowledge that innocent defendants might plead guilty.

Needed Research on Factual Guilt and Plea Decision-Making

In the mid-1990s, notable scholars like Kassin (1997; Kassin and Kiechel 1996) and Leo (1996) published articles on interrogations and (false) confessions. Although there certainly had been some articles on confessions prior to these, arguably it was these now seminal publications from the mid-1990s that sparked a revolution of research on police interrogation and false confessions. In 1997, Kassin had described the study of false confessions as "meager" (p. 231). However, by 2010, the corpus of research had grown so much as to support a scientific "white" paper, sponsored by the American Psychology-Law Society (Kassin et al. 2010). Recently, Redlich et al. (in press) argued that a similar amount of attention be paid to guilty pleas, generally, and false guilty pleas, specifically. Following the trajectory of the false confession literature, scholars should investigate both the situational and dispositional factors that lead to defendants accepting guilt (via pleas) for crimes

they did not commit (see Malloy et al. 2014; Redlich 2010). Moreover, creative methodologies are needed to study the phenomenon of false guilty pleas. Although self-report studies (e.g., Malloy et al. 2014; Redlich et al. 2010) and hypothetical vignettes (e.g., Bordens 1984) have advanced our understanding of the circumstances under which false guilty pleas can occur, innovative, ethical, and plausible methods are needed to take this nascent field to the next level.

Conclusion

Despite the popularity of long-standing television shows like "Law and Order," that depict our adversarial justice system in action, the stark reality is that only a very small proportion of defendants opt for trial and actually experience this process. Because ours is a system of pleas, ensuring that pleas are indeed knowing, intelligent, voluntary, and based on factual guilt as envisioned by the U.S. Supreme Court, becomes all the more imperative. If we ask "Are we measuring what we think we are measuring" in regard to courts' assessments of knowingness, intelligence, voluntariness, and factual guilt, the answer is most likely a categorical "no." Because of deficient capabilities, lack of knowledge and complete information, incomprehensible questions and written documents, undue pressure, and the fact that the plea bargaining system incentivizes some innocent defendants to plead guilty, it is likely that the number of people entering invalid pleas is much higher than known or appreciated.

Psychological scientists are well-positioned to address the validity of pleading guilty. There is a wealth of research on juror/jury decision-making, interviewing victims, witnesses, and suspects, and adjudicative competence at many different points on the criminal justice system continuum. These extensive bodies of literature, described in this volume and elsewhere, would surely inform and apply to research questions on guilty pleas, the process, and the people involved.

References

Acker, J., & Redlich, A. D. (2011). *Wrongful convictions: Law, social science, and policy*. Durham, NC: Carolina Academic Press.

Advokat, C. D., Guidry, D., Burnett, D. M. R., Manguno-Mire, G., & Thompson, J. W. (2012). Competency restoration treatment: Differences between defendants declared competent and incompetent to stand trial. *Journal of the American Academy of Psychiatry and Law, 40*, 89–97.

Alschuler, A. W., & Deiss, A. G. (1994). A brief history of the criminal jury in the United States. *The University of Chicago Law Review, 61*, 867–928.

American Bar Association. (1999). *ABA standards for criminal justice, pleas of guilty* (3rd ed.). Washington, DC: American Bar Association.

Anderson, S. D., & Hewitt, J. (2002). The effect of competency restoration training on defendants with mental retardation found not competent to proceed. *Law and Human Behavior, 26*, 343–351.

Benchbook for U.S. district court judges. (2013). 6th Edn. Federal Judicial Center. Accessible at http://www.fjc.gov/public/pdf.nsf/lookup/Benchbook-US-District-Judges-6TH-FJC-MAR-2013-Public.pdf/$file/Benchbook-US-District-Judges-6TH-FJC-MAR-2013-Public.pdf

Bibas, S. (2003). Harmonizing substantive criminal law values and criminal procedure: The case of Alford and Nolo Contendere pleas. *Cornell Law Review, 88*, 1361–1412.

Bibas, S. (2004). Plea bargaining outside the shadow of trial. *Harvard Law Review, 117*, 2463–2547. doi:10.2307/4093404

Bibas, S. (2011). *The myth of the fully informed rational actor* (pp. 79–86). XXXI: Saint Louis University Public Law Review.

Bibas, S. (2014). Plea bargaining's role in wrongful convictions. In A. D. Redlich, J. R. Acker, R. J. Norris, & C. L. Bonventre (Eds.), *Examining wrongful convictions: Stepping back, moving forward* (pp. 157–167). Durham, NC: Carolina Academic Press.

Blackwood, H. L. (2009). *A comparison of Miranda procedures: The effects of oral and written administrations on Miranda comprehension*. Denton, Texas. UNT Digital Library. http://digital.library.unt.edu/ark:/67531/metadc11057/. Accessed July 29, 2014.

Bonnie, R. J. (1992). The competence of criminal defendants: A theoretical reformulation. *Behavioral Sciences and the Law, 10*, 291–316. doi:10.1002/bsl.2370100303

Bordenkircher v. Hayes, 434 U.S. 357. (1978).

Bordens, K. S. (1984). The effects of likelihood of conviction, threatened punishment, and assumed role on mock plea bargain decisions. *Basic and Applied Social Psychology, 5*, 59–74.

Bordens, K. S., & Bassett, J. (1985). The plea bargaining process from the defendant's perspective: A field investigation. *Basic and Applied Social Psychology, 6*, 93–110. doi:10.1207/s15324834basp0602_1

Boruchowitz, R. C., Brink, M., Dimino, M. (2009). *Minor crimes, massive waste: The terrible toll of America's broken misdemeanor courts.* Retrieved from https://www.nacdl.ord/reports/misdemeanor/

Bowers, J. (2008). Punishing the innocent. *University of Pennsylvania Law Review, 156*, 1117–1179. Retrieved from http://www.law.upenn.edu/lrev/

Boykin v. Alabama, 395 U.S. 238. (1969).

Brady v. United States, 373 U.S. 83. (1963).

Bureau of Justice Statistics. (2010). *Felony defendants in large urban counties, 2006*. Washington, DC: Office of Justice Programs, U.S. Department of Justice. Available at: http://bjs.ojp.usdoj.gov/content/pub/pdf/fdluc06.pdf

Bushway, S., & Redlich, A. (2012). Is plea bargaining in the "shadow of the trial" a mirage? *Journal of Quantitative Criminology, 28*(3), 437–454. doi:10.1007/s10940-011-9147-5

Bushway, S., Redlich, A., & Norris, R. J. (2014). An explicit test of "bargaining in the shadow of trial." *Criminology, 52*, 723–754. doi:10.1111/1745-9125.12054

Cialdini, R. B. (2001). *Influence: Science and practice*. New York: Harper Collins.

Cohen, T. H., & Reaves, B. A. (2006). *Felony defendants in large urban counties, 2002*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.

Coleman, C. (2002). *PBS Frontline: A crime of insanity. Interview with Prosecutor Cheryl Coleman.* Accessed May 24, 2006, http://www.pbs.org/wgbh/pages/frontline/shows/crime/interviews/coleman.html

Cook, K. M. (2007). *Chasing justice: My story of freeing myself after two decades on death row for a crime I did not commit*. New York: Harper Collins.

Dervan, L. & Edkins, V. A. (2013). Pleading innocents: Laboratory evidence of plea bargaining's innocence problem. *Journal of Criminal Law and Criminology*, *103*, 1–48.

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

Edkins, V. A. (2011). Defense attorney plea recommendations and client race: Does zealous representation apply equally to all? *Law and Human Behavior, 35*, 413–425. doi:10.1007/s10979-010-9254-0

Everington, C., & Dunn, C. (1995). A second validation study of the competence to stand trial for defendants with mental retardation (CAST-MR). *Criminal Justice and Behavior, 22*, 44–59.

Fisher, G. (2003). *Plea bargaining's triumph: A history of plea bargaining in America*. CA: Stanford University Press.

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

Gregory, W. L., Mowen, J. C., & Linder, D. E. (1978). Social psychology and plea bargaining: Applications, methodology, and theory. *Journal of Personality and Social Psychology, 36*, 1521–1530.

Grisso, T., Steinberg, L., Woolard, J., Cauffman, E., Scott, E., Graham, S., Lexcen, F., Reppucci, N.D., & Schwartz, R. (2003). Juveniles' competence to stand trial: A comparison of adolescents' and adults' capacities as trial defendants. *Law and Human Behavior, 27*, 333–364. doi:10.1023/A:1024065015717

Guthrie, C., Rachlinski, J., & Wistrich, A. (2001). Inside the judicial mind. *Cornell Law Review, 86*, 777.

Haigler, K. O., Harlow, C., O'Connor, P., & Campbell, A. (1994). *Literacy behind prison walls: Profiles of the prison population from the national adult literacy survey, (NCES 94102)*. U.S: Department of Education.

Hodges, J., Guiliotti, N., & Porpotage, F. M. (1994). Improving literacy skills of juvenile detainees. *OJJDP Juvenile Justice Bulletin*. Washington, DC: Office of Justice Programs, U.S. Department of Justice.

James M. Crumb, Jr. v. The People of the State of Colorado, Case No. 08SC884. (2010).

Kaban, B., & Quinlan, J. C. (2004). Rethinking a "knowing, intelligent and voluntary waiver" in Massachusetts' juvenile courts. *Journal of the Center for Families, Children, and the Courts, 5*, 35–55.

Kalven, H., & Zeisel, H. (1966). *The American jury*. Chicago: University of Chicago Press.

Kassin, S. M. (1997). The psychology of confession evidence. *American Psychologist*, *52*, 221–233. doi:10.1037/0003-066X.52.3.221

Kassin, S. M., Drizin, S., Grisso, T., Gudjonsson, G., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions: Risk factors and recommendations. *Law and Human Behavior, 34*, 3–38. doi:10.1007/s10979-010-9217-5

Kassin, S. M., & Kiechel, K. (1996). The social psychology of false confessions: Compliance, internalization, and confabulation. *Psychological Science, 7*, 125–128.

Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research* (4th ed.). Belmont, CA: Cengage Learning.

Kutateladze, B. L., Andiloro, N. R., & Johnson, B. D. (2014). Opening Pandora's box: How does defendant race influence plea bargaining? *Justice Quarterly.*

Lafler v. Cooper, 132 S. Ct. 1376. (2012).

Landes, W. (1971). An economic analysis of the courts. *Journal of Labor Economics, 14*(1), 61–107.

Lang, D. M. (2012). *Padilla v. Kentucky*: The effect of plea colloquy warnings on defendants' ability to being successful *Padilla* claims. *The Yale Law Journal, 121*, 944–997.

Langbein, J. H. (1992). On the myth of written constitutions: The disappearance of criminal jury trial. *Harvard Journal of Law and Public Policy, 15*, 119–127.

Leo, R. A. (1996). Inside the interrogation room. *Journal of Criminal Law and Criminology, 86*, 266–303. doi:10.2307/1144028

Lynch, T. (2003). The case against plea bargaining. *Regulation, fall*, 24–27.

Malloy, L. C., Shulman, E. P., & Cauffman, E. (2014). Interrogations, confessions, and guilty pleas among serious adolescent offenders. *Law and Human Behavior, 38*, 181–193.

Miller, H. S., McDonald, W. F., & Cramer, J. A. (1978). *Plea bargaining in the United States*. In Final report to the National Institute of Law Enforcement and Criminal Justice, Washington, DC.

Mossman, D. (2007). Predicting restorability of incompetent criminal defendants. *The Journal of the American Academy of Psychiatry and the Law, 35*, 34–43.

Mossman, D., Noffsinger, S. G., Ash, P., Frierson, R. L., Gerbasi, J., Hackett, M., Lewis, C F., Pinals, D. A., Scott, C. L., Sieg, K. G., Wall, B. W., & Zonana, H. V. (2007). AAPL practice guideline for the forensic psychiatric evaluation of competence to stand trial. *Journal of the American Academy of Psychiatry and Law, 35*(Suppl.) S3–S72.

Nagel, S. S., & Neef, M. G. (1979). *Decision theory and the legal process*. Lexington, MA: Lexington.
Norris, R., & Redlich, A. D. (2014). Seeking justice, compromising truth? Criminal admissions and the Prisoner's Dilemma. *Albany Law Review*.
North Carolina v. Alford, 400 U.S. 25. (1970).
Padilla v. Kentucky, 130 S. Ct. 1473. (2010).
Pinals, D. A. (2005). Where two roads meet: Restoration of competence to stand trial from a clinical perspective. *New England School of Law, 31*(1), 81–108.
Pirelli, G., Gottdiener, W. H., & Zapf, P. A. (2011). A meta-analytic review of competency to stand trial research. *Psychology, Public Policy, and Law, 17*, 1–53.
Plous, S. (1993). *The psychology of judgment and decision-making*. New York, NY: McGraw-Hill Series in Social Psychology.
Poythress, N. G., Bonnie, R. J., Hoge, S. K., Monahan, J., & Oberlander, L. B. (1994). Client abilities to assist counsel and make decisions in criminal cases. *Law and Human Behavior, 18*, 437–452. doi:10.1007/BF01499049
Poythress, N., Monahan, J., Bonnie, R., Otto, R. K., & Hoge, S. K. (2002). *Adjudicative competence: The MacArthur Studies*. New York: Kluwer/Plenum.
Rakoff, J. (2014). Why innocent people plead guilty. *The New York Review of Books*. Accessible at, http://www.nybooks.com/articles/archives/2014/nov/20/why-innocent-people-plead-guilty/
Redlich, A. D. (2010). False confessions and false guilty pleas: Similarities and differences. In G. D. Lassiter & C. Meissner (Eds.), *Interrogations and confessions: Current research, practice and policy* (pp. 49–66). Washington, DC: APA Books.
Redlich, A. D., Bibas, S., Edkins, V., & Madon, S. (in press). *The psychology of defendant plea decision-making*. American Psychologist.
Redlich, A. D., & Bonventre, C. (2015). Content and comprehensibility of adult and juvenile tender-of-plea forms: Implications for knowing, intelligent, and voluntary guilty pleas. *Law and Human Behavior*, *39*, 162–176.
Redlich, A. D., Bushway, S., Norris, R., & Yan, S. (2013). *Bargaining in the shadow of trial: Exploring the reach of evidence outside the jury box*. Final report to the National Institute of Justice, 2009-IJ-CX-0035. Washington, DC: Office of Justice Programs.
Redlich, A. D., & Cusack, K. (2010). Mental health treatment in criminal justice settings. In B. L. Levin, K. D. Hennessey, & J. Petrila (Eds.), *Mental health services: A public health perspective* (3rd ed., pp. 421–440). New York: Oxford University Press.
Redlich, A. D., Hoover, S., Summers, A., & Steadman, H. J. (2010a). Enrollment in mental health courts: Voluntariness, knowingness, and adjudicative competence. *Law and Human Behavior, 34*, 91–104.
Redlich, A. D., & Ozdogru, A. A. (2009). Alford pleas in the age of innocence. *Behavioral Sciences and the Law, 27*, 467–488.
Redlich, A. D., & Shteynberg, R. (2016). To plead or not to plead: A comparison of juvenile and adult true and false plea decisions. *Law and Human Behavior*. Online first.
Redlich, A. D., & Summers, A. (2012). Voluntary, knowing, and intelligent pleas: Understanding plea inquiries. *Psychology, Public Policy, and Law, 18*, 626–643.
Redlich, A. D., Summers, A., & Hoover, S. (2010b). Self-reported false confessions and false guilty pleas among offenders with mental illness. *Law and Human Behavior, 34*, 79–90.
Roberts, J. (2009). Ignorance is effectively bliss: Collateral consequences, silence, and misinformation in the guilty-plea process. *Iowa Law Review, 95*, 119–194.
Russano, M. B., Meissner, C. A., Narchet, F. M., & Kassin, S. M. (2005). Investigating true and false confessions within a novel experimental paradigm. *Psychological Science, 16*, 481–486.
Sanborn, J. B. (1992). Pleading guilty in juvenile court: Much ado about something very important to young defendants. *Justice Quarterly, 9*, 127–150.
Shipley, C. J. (1987). The Alford plea: A necessary but unpredictable tool for the criminal defendant. *Iowa Law Review, 72*, 1063–1089.
Smith, D. A. (1986). The plea bargain controversy. *The Journal of Criminal Law and Criminology, 77*, 949–968.

State v. Kaufman, 51 Iowa 578. (1879).
Steadman, H. J., Osher, F., Robbins, P. C., Case, B., & Samuels, S. (2009). Prevalence of serious mental illness among jail inmates. *Psychiatric Services, 60*, 761–765.
Stuntz, W. J. (2004). Plea bargaining and criminal law's disappearing shadow. *Harvard Law Review, 117*, 2548–2569.
Taslitz, A. E. (2008). Prosecutorial preconditions to plea negotiations: "Voluntary" waivers of constitutional rights. *Criminal Justice, 23*, 14–26.
Tor, A., Gazal-Ayal, O., & Garcia, S. M. (2010). Fairness and willingness to accept plea bargain offers. *Journal of Empirical Legal Studies, 7*, 97–116.
Tucker, A. (1950). A two-person dilemma. In E. Rasmusen (Ed.), *Readings in games and information*. Oxford, UK: Blackwell Publishers.
Turner, J. A. (2006). Judicial participation in plea negotiations: A comparative view. *American Journal of Comparative Law, 54*, 199–267.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124.
Ulmer, J. T., Eisenstein, J., & Johnson, B. D. (2010). Trial penalties in federal sentencing: Extra-guidelines factors and district variation. *Justice Quarterly, 27*, 560–592.
United States v. Ruiz, 536 U.S. 622. (2002).
United States v. Speed Joyeros, S.A., 204 F.Supp.2d 412 (E.D.N.Y. 2002).
Wall, B. W., & Christopher, P. P. (2012). A training program for defendants with intellectual disabilities who are found incompetent to stand trial. *Journal of the American Academy of Psychiatry and Law, 40*, 366–373.
Weaver, K. D. (2001–2002). A change of heart or a change of law? Withdrawing a guilty plea under federal rule of criminal procedure 32(e). *Journal of Criminal Law and Criminology, 92*, 273–306.
Wolf Harlow, C. (2003). Educational and correctional populations. Washington, DC: Office of Justice Programs, Bureau of Justice Statistics. Available at: http://www.bjs.gov/content/pub/ascii/ecp.txt
Zapf, P. A., & Viljoen, J. L. (2003). Issues and considerations regarding the use of assessment instruments in the evaluation of competency to stand trial. *Behavioral Sciences and the Law, 21,* 351–367. doi:10.1002/bsl.535

Capital-izing Jurors: How Death Qualification Relates to Jury Composition, Jurors' Perceptions, and Trial Outcomes

Logan A. Yelderman, Monica K. Miller and Clayton D. Peoples

During the seventeenth and eighteenth centuries, Colonial America imposed the death penalty for various capital crimes. These capital crimes were specific acts that deemed the defendant eligible for the death penalty, including murder, treason, and burglary, among others (Banner 2002). Most of these death penalties were carried out via public hangings. After the U.S. gained independence, the legal system and punishment—including the death penalty—evolved (Banner 2002; Vidmar and Hans 2007). Currently, capital crimes are largely limited only to murder,[1] and death row inmates are generally executed in a private setting by lethal injection, which is considered to be more humane than previous methods such as hanging, beheading, electrocution, gas chamber, and firing squad.

The legal process involved in capital trials has also changed in that during sentencing decisions in death penalty trials, jurors are the fact finders while judges are the ultimate decision-makers based on the jury's recommendation (*Ring v. Arizona* 2002). Capital trials typically involve a bifurcated trial (i.e., two-phase) process. In the first phase, called the guilt phase, the jury weighs evidence and decides on the guilt of the defendant, similar to non-death penalty criminal trials. If the defendant is found guilty, the trial moves to the second phase, called the sentencing phase. In this phase, the jury must decide on the punishment for the defendant; options typically include life in prison without the possibility of parole or the death penalty (Vidmar

[1]Several states have statutes allowing the death penalty for the rape of a minor/child; however, this was considered unconstitutional in *Kennedy v. Louisiana* (2008; see also Kirchmeier 2006). Similarly, the federal death penalty guidelines allow for the death penalty for crimes such as wrongful deaths (in specific contexts), treason, and espionage (Snell 2014).

L.A. Yelderman (✉)
Prairie View A&M University, Prairie View, TX, USA
e-mail: layelderman@pvamu.edu

M.K. Miller · C.D. Peoples
University of Nevada, Reno, Reno, NV, USA

B.H. Bornstein and M.K. Miller (eds.), *Advances in Psychology and Law*,
Advances in Psychology and Law 2, DOI 10.1007/978-3-319-43083-6_2

and Hans 2007). In most states, the jurors "weigh" or compare aggravating and mitigating circumstances (Devine 2012) when making sentencing decisions. Aggravating circumstances are characteristics of a crime that make the defendant more worthy of the death penalty (e.g., the crime is especially heinous). Mitigating circumstances are characteristics of the defendant or crime that make the defendant more deserving of a "life in prison without parole" sentence (e.g., the defendant cooperated with the police; Butler and Moran 2002). At least one aggravating circumstance must be present for the death penalty to be an option. Typically, when aggravators are weighed more heavily or are more substantial than mitigators, then the death penalty is given to the defendant (see Latzer and McCord 2011).

Another aspect of death penalty trials that has evolved over the past century is the death qualification process. Death qualification is a process of jury selection unique to death penalty trials, and it occurs during voir dire, the general jury selection process by which potential jurors are excluded if they are unable to render a fair verdict (Vidmar and Hans 2007). Based on potential death penalty trial jurors' responses to judges' and attorneys' questions, some of them are excluded from jury service during death qualification (Haney 2005). Even though this process is in place, presumably to help choose a fair and impartial jury, death qualification might be inherently flawed because it systematically excludes some groups and increases the tendency for jurors to choose death sentences over life in prison sentences (Butler and Moran 2007b; Cowan et al. 1984; Haney 1984; Mauro 1992).

The purpose of this chapter is to: (1) give a brief overview, describing the evolution of the death penalty and death qualification process; (2) examine how the death qualification process might influence jury composition, jurors' perceptions, and jurors' verdicts and sentences; and (3) offer theoretical applications to help explain why death qualification has these effects and propose future research.

The Death Penalty

All but 16 states have sentenced a defendant to death or executed an inmate in the past 40 years; since then, several states have abolished the death penalty (Smith 2012). Currently, 31 states, the federal government, and the U.S. military allow the death penalty. Of the states that have executed inmates in the past 40 years, ten have executed one (or more) inmate per year and three states have executed two or more inmates per year (Smith 2012). This suggests that, although some states have abandoned their use of execution, a majority of states have retained it and many states use it somewhat regularly. However, the death penalty has not always been considered an acceptable method of punishment or criminological strategy. This section provides a brief case history of the death penalty and capital trial processes.

Although the death penalty was broadly applied to common criminals (e.g., thieves, arsonists, burglars, robbers, and petty treason offenders) in Colonial America, not all Americans supported execution for some of these lesser crimes (Banner 2002). For example, William Penn insisted that the death penalty be

reserved only for murderers, an idea that is reflected in modern law (Banner 2002; Palmer 2014). In the early twentieth century, support for and use of the death penalty began declining; then, in *Furman v. Georgia* (1972), the Supreme Court placed a moratorium on the death penalty stating that it violated the cruel and unusual punishment clause in the Eighth Amendment of the U.S. Constitution (Latzer and McCord 2011; Palmer 2014). This claim of cruel and unusual punishment was unrelated to the *method* of execution, but instead was based on the inconsistency and disparity in its *application*; specifically, death sentences were disproportionately given to individuals of certain races and social classes (Palmer 2014). The death penalty was then reintroduced and deemed constitutional in *Gregg v. Georgia* (1976) with the requirement that its application must include more objective criteria, outlined by each state that allows it. This case led to the adoption of several sentencing schemes; the most common is a system of weighing aggravating and mitigating circumstances.

In the decades leading up to *Gregg v. Georgia* (1976), Americans and legal professionals were introduced to the counterintuitive belief that criminal offenses might not necessarily occur under complete volition (Banner 2002). This belief had significant effects on individuals' determination of whether defendants were capable of assuming responsibility for the crime, and thus, eligible for the death penalty. Following *Gregg v. Georgia* (1976), several Supreme Court cases reflected the idea that the mental state of the defendant mattered. Ultimately, the death penalty was ruled to be unconstitutional for the criminally insane (*Ford v. Wainwright* 1986), the mentally retarded (*Atkins v. Virginia* 2002), and juveniles who committed a crime under the age of 18 (*Roper v. Simmons* 2005). Moreover, since 2000, seven states have abolished the death penalty (four states since 2010), and four of the 31 states that have not abolished the death penalty are currently on a governor-imposed moratorium (retrieved from http://www.deathpenaltyinfo.org). This provides evidence that the landscape of death penalty legislation across the U.S. is continuing to evolve.

This brief history highlights how prior controversies (e.g., disparity in application and the constitutionality of executing individuals with compromised mental states) manifest in changes in death penalty policies and legislation. Though lawmakers, indeed, consult public opinion when making legislative decisions, evidence of any causal association is tenuous (McGarrell and Sandys 1996). While the death penalty was abandoned for a brief period of time, it was reintroduced in 1976 (after *Gregg v. Georgia* 1976) with new reforms that required changes in the trial processes. This generated a new controversy related to changes that affected the death qualification process and jurors' discretion over who was sentenced to death and for what reasons. Specifically, the death qualification process was (and for the most part still is) critiqued to unduly influence capital juries, increasing conviction proneness. Similarly, capital juries were critiqued for using targeted, or channeled, discretion convicting specific groups of individuals at disproportionate rates. Alterations in the process of death qualifying jurors, as a result of evolving death penalty legislation, is identified as potentially underlying increased biases related to jurors' perceptions during capital trials, verdicts, and sentencing decisions (Haney

2005). The potential biasing effects of the death qualification process is the primary focus of the rest of this chapter.

A History of Death Qualification

The process of death qualification is considered to be a necessary component in capital trials because it minimizes the possibility that a jury would consist of individuals who are unable to evaluate evidence appropriately and impartially; however, research suggests that the process itself might bias various aspects of the trial (Banner 2002; Haney 2005). Beginning in the 1980s, researchers have examined the potential influence of death qualification on juries. This section includes a brief legal history of death qualification and then provides an extensive examination of the research related to how it impacts various aspects of the trial process.

Several cases outline the evolution of death qualification in the jury selection process. These changes mark attempts by the courts to best ensure a process that selects jurors who are not so biased that they cannot properly follow the law (Vidmar and Hans 2007). The following cases highlighted changes made in death penalty cases specific to death qualification.

In *Witherspoon v. Illinois* (1968), the Supreme Court ruled that, during voir dire, attorneys cannot systematically exclude jurors who did not endorse pro-death penalty attitudes (leaving a jury of only pro-death penalty jurors); this was considered a violation of the Sixth Amendment (impartial jury clause). Then, in the 1970s, after the *Furman v. Georgia* (1972) and *Gregg v. Georgia* (1976) cases, most states that allowed the death penalty adopted the bifurcated trial process (Vidmar and Hans 2007). Bifurcated trials changed the trial process such that jurors participated in two phases: a verdict phase and a sentencing phase (see Banner 2002; Haney 2005; Latzer and McCord 2011; Palmer 2014; Vidmar and Hans 2007). Because jurors now heard essentially two sets of arguments and evaluated case facts at two separate times, the jury selection process required changes. Potential jurors had to be screened for specific attitudes and beliefs that might influence their guilt judgments *and* sentencing judgments (see Banner 2002; Haney 2005; Latzer and McCord 2011; Palmer 2014; Vidmar and Hans 2007).

Nearly two decades after *Witherspoon v. Illinois* (1968), the Supreme Court ruled in *Wainwright v. Witt* (1985) that jurors can be excluded based on personal beliefs that influence their ability to sentence a person to death properly under the statute, and such exclusions do not violate the Sixth Amendment (impartial jury clause) or Fourteenth Amendment (due process clause). This ruling broadened the criteria for which potential jurors could be excluded (see Devine 2012; Haney 2005; Vidmar and Hans 2007), such as economic attitudes, religious beliefs, death penalty attitudes, beliefs about criminality, etc.

Under the *Witherspoon v. Illinois* (1968) death qualification criteria, jurors could be removed based on whether or not they would *automatically* choose one verdict over the other, with jurors typically conveying their inability to give a death

sentence. Jurors might be asked if they would vote for the death penalty if the defendant was found guilty. Jurors who answered that they would never be able to give a death sentence (under any circumstances) were likely removed. However, this process of removal could not be systematic such that it resulted in an only pro-death penalty jury. Under the *Wainwright v. Witt* (1985) death qualification criteria, jurors could be removed if their personal or religious attitudes rendered them unable to be impartial, despite whether or not they would automatically vote for or against sentencing the defendant to death. In this case, jurors might be asked if they had any personal attitudes or beliefs that related to the case or the defendant and whether or not these attitudes or beliefs would influence their sentencing decision. If jurors responded that their attitudes or beliefs would influence their sentencing decisions, or even if they simply *thought* that their attitudes might influence their sentencing decisions, then they were eligible for removal.

Challenges to the constitutionality of death qualification surfaced again in *Lockhart v. McCree* (1986), but the Supreme Court held that death-qualified juries were constitutional. *Lockhart v. McCree* (1986) was an important case because attorneys for the defendant McCree provided social science research supporting the notion that death-qualified juries were more "conviction prone," meaning that a jury made up of death-qualified jurors was more likely to give a conviction than a jury that was made up of a broader selection of jurors. The Supreme Court stated that this evidence was too weak to firmly conclude that that death-qualified juries were unconstitutional. Although this issue has yet to be revisited by the Supreme Court, research conducted after *Lockhart v. McCree* (1986) suggests that McCree's claim that death-qualified juries are conviction prone was valid. Several accounts (e.g., Butler and Moran 2007b; Cowan et al. 1984; Haney 1984; Horowitz and Seguin 1986; Mauro 1992) support the notion that death-qualified juries are more conviction prone and death sentence prone than non-death-qualified juries (which will be discussed in the following sections of the chapter).

Later, in *Morgan v. Illinois* (1992), the Supreme Court ruled that individuals who would automatically impose the death penalty every time for a convicted murderer eligible for the death penalty (called automatic death penalty jurors, or ADP jurors) could be removed as well as those who were absolutely opposed to the death penalty. Although it could be argued that removing ADP jurors offsets the removal of jurors who are strongly opposed, this is not necessarily the case. In a study by Miller and Hayward (2008), death qualification eliminated 11 ADP jurors and 147 jurors who were absolutely opposed to the death penalty. Thus, death qualification might still result in juries that are, as a whole, more supportive of the death penalty than the public in general.

More recently, the U.S. Supreme Court ruled that appellate courts cannot decide whether a lower court should have removed a juror during death qualification. In *Uttecht v. Brown* (2007), an appeals court reversed a trial court's decision based on its finding that the trial court's exclusion of a juror biased the jury against the defendant. However, the Supreme Court ruled that the trial court's decision to remove the juror must be upheld by the appellate courts. This ruling gives trial court judges and attorneys the sole discretion to decide which jurors are biased and which are not biased.

Each of these cases addresses the issue of the constitutionality of jury selection and death qualification in capital trials. Because attorneys exclude potential jurors who might exhibit bias (based on the jurors' attitudes, beliefs, or perceptions related to the case) during the death qualification process, attorneys might systematically remove specific social and demographic groups. The next section focuses on current research that addresses these issues.

The Biasing Effects of Death Qualification

Research has shown that death qualification procedures can result in biased juries. This bias is demonstrated in research regarding three general areas: (1) jury composition; (2) jurors' perceptions; and (3) verdict and sentencing decisions. Each of these is discussed in turn below, after a general discussion of the voir dire process. Then theoretical implications are discussed at the end of the chapter.

Death Qualification and Jury Composition

All of the case rulings discussed above argued that the death qualification process potentially results in biased juries, because it affects the jury's composition (*Morgan v. Illinois* 1992; *Uttecht v. Brown* 2007; *Wainwright v. Witt* 1985; *Witherspoon v. Illinois* 1968). This section discusses how the death qualification process affects the jury that is selected. Not all people who are summoned for jury duty will ultimately serve on the jury, as the process of voir dire (i.e., jury selection) eliminates some potential jurors. Concerns arise when this process results in some groups (e.g., Catholics or women) being eliminated more than their counterparts.

During the general voir dire process (unrelated to death qualification), lawyers and judges exclude potential jurors through two types of challenges: challenges for cause and peremptory challenges (Lieberman and Olson 2009). Judges and attorneys use challenges for cause to remove jurors for a specific reason that can be related or unrelated to the case. For instance, potential jurors who state that they have been raped are likely to be excluded from serving as a juror on a rape case through the use of a challenge for cause. These challenges are unlimited but must be confirmed by the judge (Lieberman and Olson 2009). Challenges for cause tend to be associated with more apparent biases or circumstances that could contribute to bias, such as family ties, personal experience, personal circumstances that make it difficult for a person to serve, or prior exposure to case facts or events.

Attorneys use peremptory challenges to exclude potential jurors based on their suspicion that the juror might be biased—even if they do not necessarily have evidence of this bias (Lieberman and Olson 2009). For example, a prosecuting attorney might use a peremptory challenge to exclude a Democrat potential juror in a case involving a shooting because of a belief that Democrats are less supportive of

gun rights than Republicans, even if the potential juror specifies that he does not have strong feelings about guns. Peremptory challenges can be used to exclude potential jurors for any reason other than race (*Batson v. Kentucky* 1986) or gender (*J.E.B. v. Alabama* 1994). Though it is likely that attorneys attempt to circumvent these two case laws (e.g., *Foster v. Chatman* 2016) and use death qualification criteria to remove jurors based on race or gender.

Potential jurors in death penalty trials can be removed through death qualification if the judge determines that they are unable to be impartial (Devine 2012). Typically, death qualification excludes individuals who are either extremely opposed to *or* extremely in favor of the death penalty, particularly after *Morgan v. Illinois* (1992). For example, during death qualification, attorneys and judges might ask potential jurors if they are able to give the death penalty at all. Those who answer "yes" might then be asked if they believe the death penalty should always be given to individuals who commit capital crimes. Using these types of questions, jurors who have extreme attitudes toward the death penalty are identified and then subsequent questions are used to help judges and attorneys decide whether these extreme attitudes will influence potential jurors' ability to be impartial. Jurors who are excluded (i.e., "excludables") might tend to share common identities, traits, or group affiliations. Therefore, death qualification might exclude specific groups of people that share common attitudes or beliefs that are associated with their opposition to or support for the death penalty.

It is possible that, if the death qualification process systematically eliminates certain groups or individuals with specific traits, then the process might result in a biased jury. For example, if the death qualification process systematically excludes religious conservatives who hold the law in high authority (i.e., legal authoritarians) and tend to be of a racial minority group, then the jury might be disproportionately made up of White, nonreligious liberals who believe the law is malleable and that a man's instinct has the highest authority. As a result, any biases inherent in these groups would then affect the jury verdict. This is also true if the exclusionary result was the opposite; then the prejudices embedded within minority religious conservatives who were legal authoritarians would permeate death-qualified juries. Research examining the effects of death qualification on jury composition suggests that death qualification often results in juries that are biased in ways that consistently disadvantage capital defendants (e.g., Haney et al. 1994). It is possible to argue that as long as those who are extremely against *and* extremely for the death penalty are excluded, that risk would be minimal. However, this is not necessarily true because, as mentioned above, the process likely excludes those who strongly oppose the death penalty at a higher rate than those who strongly support the death penalty. There are two aspects of death qualification that support this notion. First, one could argue that it is likely easier, in practice, to identify and exclude those who are extremely against the death penalty than those who are extremely in favor. Second, there may simply be more people in society who are extremely *opposed* to the death penalty than those extremely in *favor* based on a reading of recent Gallup polls (see also Miller and Hayward 2008).

We would contend that identifying those who are extremely against the death penalty is probably relatively easy in practice given that, procedurally, it involves asking a simple yes/no question (upon which the prospective juror could be categorized as "not death-qualified" and be excluded from the jury). For instance, the question might be phrased, "Would you be willing to consider the death penalty as a possible sentence if the defendant were found guilty?" If the answer is "no," then the prospective juror is identified as failing to meet the death qualification criteria and therefore is excluded from the jury. Even if follow-up questions are posed, these, too, would likely involve dichotomous yes/no answers that would result in exclusion. For instance, the judge may follow up with, "Are you sure there are no circumstances in which you would consider the death penalty as a possible sentence?" If the answer is, "yes," the prospective juror is likely to be dismissed and excluded from the jury.

In theory, identifying jurors who are extremely in favor of the death penalty should be as easy as identifying those against the death penalty. A simple yes/no question could be asked, such as "Would you be willing to consider life without parole as a possible sentence if the defendant was found guilty?" Identifying them in practice, however, may be more challenging. Sandys and Trahan (2008) discuss the difficulties associated with identifying those they refer to as "latent ADP" jurors. Latent ADP jurors are those whose attitudes and beliefs would lead them to always vote for the death penalty if a defendant was found guilty of a capital crime, but who are difficult to identify as ADP jurors during the death qualification process. As far as the authors are aware, there is no equivalent category of "latent anti-death penalty jurors" addressed in the literature. Moreover, because death penalty attitude questions are often framed to assess death penalty opposition, it is possible that latent ADP jurors are more likely to be misidentified or unidentified during the voir dire process (e.g., Dillehay and Sandys 1996).

Although admittedly anecdotal, the third author of this paper observed firsthand the difficulties associated with identifying ADP jurors during the voir dire phase of a murder trial. Jurors who were against the death penalty were very quickly identified and excluded within just one or two yes/no questions; it took much longer to identify and exclude jurors who were extremely in favor of the death penalty. In one instance, a prospective juror expressed his belief in the Biblical dictum "an eye for an eye" and made statements indicating strong support for the death penalty. He was only successfully identified as ADP and excluded from the jury when he stated that, in a hypothetical scenario, even his wife should receive the death penalty if she were found guilty of murder.

Although the above paragraphs provide insight into the death qualification process and the difficulties associated with identifying potential jurors with extreme death penalty attitudes, it is also important to understand these attitudes from a broader societal perspective. When examining public sentiment concerning the death penalty, a striking pattern emerges: approximately 37 % of the population is opposed to the death penalty (see, e.g., Gallup Polls on the topic, available here: http://www.gallup.com/poll/1606/death-penalty.aspx), while just 21 % of the population would be unwilling to impose a sentence other than death (Sandys and

Trahan 2008; in an analysis of Bowers 1995). Given the larger percentage of the population adamantly opposed to the death penalty relative to those strongly in favor, it is little wonder that statisticians have determined that even after various court decisions allowed for the exclusion of ADP jurors, their exclusion would do little to curb a jury's bias on juries (Kadane 1984). Simply put, removing the 21 % of ADP jurors is not enough to counter the effect of removing the 33 % of jurors who oppose the death penalty.

If the death qualification process resulted in the exclusion of more jurors in opposition to the death penalty than jurors in favor, death-qualified juries might exhibit specific biases. Though this bias might directly relate to death penalty attitudes and beliefs, it might also result in a disproportionate representation of jurors with certain demographic characteristics, legal attitudes, and justice beliefs. Juries that are not representative of all demographics, attitudes, and beliefs are not necessarily unconstitutional, but when the death qualification process consistently and systematically excludes individuals with particular attitudes and beliefs, problems can arise. For instance, issues arose during death qualification and jury selection in the trial of the Boston marathon bombing suspect Dzhokhar Tsarnaev in 2015 (MacDonald 2015). Catholics were disproportionately removed relative to other religious groups. This was particularly an issue because the greater Boston area is 46 % Catholic (MacDonald 2015).

The Tsarnaev case illustrates the concern that death qualification has potential to systematically exclude specific social groups. Research supports the notion that religion relates to death penalty support; Protestants are more in favor of the death penalty than Catholics or Jews (Bornstein and Miller 2009; Summers et al. 2010). Thus, it is not surprising that Catholics were being eliminated. Other groups are also at risk of being disproportionately excluded; analyses of data from the General Social Survey (GSS) show that men have been consistently more in favor of the death penalty than women over the past 30 years by an average of about 11 % (Haney 2005). There are also racial/ethnic differences such that Whites are more in favor of the death penalty than ethnic minorities. For instance, analyses of GSS data reveal a consistent difference between Whites and Blacks of approximately 28 % over the past 30 years (Haney 2005). Thus, prosecutors who know such statistics are likely to exclude women, minorities, Catholics, and Jews.

Some research more specifically addresses the question of whether death qualification systematically excludes jurors who hold certain religious beliefs. In a study assessing relationships between death qualification status and individual differences, Catholics were more likely to be excluded than non-Catholics (Summers et al. 2010). Further, religious fundamentalists and individuals who interpreted the Bible literally were less likely to be excluded, compared to their counterparts, and both fundamentalism and literal interpretation were related to attitudes in favor of the death penalty (Miller and Hayward 2008; Summers et al. 2010). This suggests that, through death qualification, juries are more likely to be composed of jurors with certain types of religious beliefs—beliefs that are related to attitudes in favor of the death penalty.

While this might be the case, attorneys and judges are not instructed to compose a jury based on religious beliefs but on death penalty attitudes and beliefs (*Witherspoon v. Illinois* 1968). Despite this understanding, several prosecutors still select jurors, using religious beliefs and affiliations (i.e., Jews) as exclusion criteria (Bornstein and Miller 2009; Darrow 1936; Dieter 2005). It is possible that attorneys and judges align their jury selection practices with official religious stances on the death penalty. For example, juries that constitute a prosecutorial advantage would *include* jurors who affiliate with the Lutheran, Evangelical, and Southern Baptist traditions because these groups officially support the death penalty. In contrast, juries that constitute a prosecutorial advantage would *exclude* jurors who affiliate with the Catholic, Episcopalian, Jewish, Presbyterian, Methodist, Unitarian Universalist, and Church of Christ traditions because these groups officially oppose the death penalty (Bornstein and Miller 2009). Moreover, research suggests Protestants, specifically fundamentalists and Evangelicals, tend to be more supportive of the death penalty than Catholics (Grasmick et al. 1993; Miller and Hayward 2008; O'Neil et al. 2004; Stack 2003; Young 1992, 2000).

Death qualification also systematically excludes jurors based on race, gender, political affiliation, and social class. Specifically, women, Democrats, and low income individuals are more likely to be excluded than their counterparts (Fitzgerald and Ellsworth 1984; Summers et al. 2010). Finally, death qualification affects the racial composition of juries, specifically excluding Blacks and other racial/ethnic minorities (Dieter 2005; Fitzgerald and Ellsworth 1984; Gonzalez-Perez 2001, 2002; Swafford 2011). In one example, a 12-person jury formed from a jury pool that was 44 % Black did not include a single Black juror; this likely was the result of death qualification and strategic peremptory challenges (Swafford 2011; see also Price 2009; Sarma 2012). Such findings support the idea that the death qualification process systematically eliminates jurors who belong to certain social and demographic groups, many of which are typically underprivileged (e.g., low income, Black). This is because members of these groups tend to have stronger oppositional attitudes toward the death penalty (Sullivan 2014). Changes in a jury's racial composition can also change the way in which case facts are interpreted and discussed by a jury (Lynch and Haney 2011). In sum, research suggests that death qualification is more likely to eliminate members of some groups than others. Whether or not death qualification influences jurors' perceptions of trial-related information and individuals involved in death penalty trials—possibly as a result of altered jury composition—is discussed in the following section.

Death Qualification and Jurors' Perceptions

Because death qualification systematically excludes jurors who strongly oppose the death penalty, death-qualified jurors might interpret trial-related information differently than excludables. These differences in interpretation potentially stem from differences in both death penalty attitudes and the death qualification process. Death

qualification is intended to exclude only those whose ability to do their duties as jurors would be affected by their attitudes toward the death penalty (as specified in the *Wainwright v. Witt* 1985 decision). Thus, death-qualified and excludable jurors should weigh evidence similarly, hold similar attitudes about defendants, attorneys, and judges, and perceive witnesses in a similar manner. However, this is not necessarily the case, as discussed in this section.

Over the past few decades, researchers have studied the effects of death qualification on jurors' interpretation of evidence, perceptions of trial participants, and beliefs related to the trial (Butler 2007a; Haney 1984; Haney et al. 1994). This research suggests that death qualification might affect jurors' mental frameworks, establishing biased information processing schemas and strengthening punitive attitudes.

In a seminal study on death qualification, Haney (1984) demonstrated that the process itself had an effect on perceptions of defendants and other individuals in the courtroom, attitudes toward the death penalty, and perceptions of others' attitudes. In this study, half of the participants were shown a video of a death qualification voir dire, and the other jurors were shown the same voir dire but with no death qualification. Then, both groups answered questions about a trial. Individuals who watched the death qualification process believed (more so than individuals who did not watch the death qualification process) that the judge, prosecuting attorney, and defense attorney thought the defendant was guilty. Also, those who watched the death qualification video were more likely to believe that the defendant would be convicted and sentenced to death for murder, and more likely to be convicted of *something* when compared to those who did not watch the death qualification portion of the video. Lastly, those who watched the death qualification video expressed more punitive attitudes toward the judge and prosecuting attorney and reported a stronger belief that the law disapproves of those who oppose the death penalty (Haney 1984). This suggests that simply being exposed to the death qualification process can have substantial effects on a juror's attitudes, perceptions of the defendant, judge, and attorneys, and perceptions of others' attitudes.

Butler (2007a) found that death-qualified jurors also displayed more general prejudicial attitudes than excludables. Specifically, death-qualified jurors reported higher levels of homophobia, racism, and sexism, which can potentially lead to prejudicial decisions regarding defendants who identify with various minority groups.

Death qualification also influences the interpretation of victim testimony and perceptions of victims or victims' survivors. Compared to excludables, death-qualified individuals were more likely to perceive themselves as similar to the victim and perceive the victim as more valuable to friends and family members (Butler 2008). They were also more likely to hold favorable attitudes toward the victim's survivors, feel similar to the survivors, and believe the survivors suffered psychologically. Finally, they were more likely to perceive the defendant unfavorably, perceive the defendant as not able to be rehabilitated in prison, perceive the defendant as less likeable, and believe that information about the impact on family was relevant in the trial and should be used by the jury (Butler 2008). This suggests

that death qualification influences the way in which individuals perceive and evaluate various people during the trial process (see also Haney 2005). Because death qualification both increases *negative* perceptions of the *defendant* and increases *positive* perceptions of the *victim(s)*, it is relatively easy to conclude that death qualification biases jurors' attitudes against the defendant.

Another important aspect of death qualification's effect on jurors in capital trials is jurors' information processing and evaluation of evidence. Since the ruling in *Gregg v. Georgia* (1976), juries are required to evaluate evidence in death penalty cases based on aggravating and mitigating circumstances related to the crime (Butler and Moran 2007a; Latzer and McCord 2011). Compared to excludables, death-qualified individuals endorsed aggravating circumstances more strongly and endorsed mitigators less strongly (Butler and Moran 2002, 2007a; see also Haney et al. 1994). Similarly, death-qualified individuals, under both the *Witherspoon* and *Witt* standards, found aggravating evidence to be more aggravating and mitigating evidence to be less mitigating, and some juries interpreted mitigating evidence as being aggravating during deliberations (Haney et al. 1994; Luginbuhl and Middendorf 1988; Stevenson et al. 2010).

Death qualification impacts the evaluation of aggravating and mitigating circumstances directly; however, it might also have an indirect impact. The effect of death qualification on jury composition likely shapes the ways in which jurors view and evaluate aggravating and mitigating circumstances, such that increasing White representation and decreasing Black representation leads to biased stereotype-driven interpretation of case facts and evidence (Swafford 2011). This stereotype-driven differentiation in evidence evaluation also stems from what is called the "empathic divide," which describes the process of dehumanizing defendants based on demographic characteristics (i.e., race). Dehumanizing defendants then leads to improper or biased interpretation of aggravating and mitigating evidence (Ghoshray 2013; Lynch and Haney 2011). Specifically, when jurors view defendants as depraved and dispositionally criminal, rather than committing crime as a result of psychological and environmental antecedents, jurors are more willing to endorse aggravating circumstances because they fit their criminal stereotype; jurors are also more willing to reject mitigating evidence because it does not fit their criminal stereotype. Juror and victim race can also perpetuate biases among these evaluations, and racial stereotypes can be used in sentencing arguments during deliberations (Devine and Kelly 2015; Foley and Powell 1982; Haney et al. 1994; Lynch and Haney 2015).

It is possible that stereotype-based evaluations of aggravating and mitigating circumstances might be a structural issue in which jurors are presented with an abundance of information they cannot effectively comprehend and complex instructions difficult to understand (Cho 1994; Hans 1995; Luginbuhl and Howe 1995; Morgan and Mannheimer 2009). This information overload results in intuitive or heuristic decision-making and improper consideration and use of aggravating and mitigating information (Morgan and Mannheimer 2009). It can also increase the likelihood that jurors rely on their own misinformed facts regarding the outcomes of their decisions. Jurors rely on misinformation during capital

sentencing, such as perceived parole likelihood and risk of future dangerousness, despite inaccuracies in those perceptions (see Costanzo and Costanzo 1994; Cunningham et al. 2009, 2011; Dayan et al. 1989; Reidy et al. 2013). Overall, death-qualified juries seem to disproportionately evaluate aggravating and mitigating circumstances with a punitive lens, often leading to prosecutorial advantage. These biases likely reflect both death qualification and the capital trial structure.

Death qualification also influences the evaluation of scientific testimony in a death penalty case. For example, death-qualified individuals were more likely to perceive ambiguous scientific testimony evidence as more valid, important, unbiased, and high quality, compared to excludables (Butler and Moran 2007b). While it seems apparent as to how death qualification influences perceptions of trial evidence, one issue unresolved by social science research is whether death qualification influences the accuracy of jurors' ability to recall evidence. While Butler (2007b) suggested that death qualification improved evidence recall and identification, Cowan et al. (1984) suggested that mock jurors in juries that included both death-qualified individuals and excludables were better able to recall evidence than juries of only death-qualified mock jurors. This research suggests that the role of death qualification in considering evidence recall remains unclear. Future research should address these inconsistencies.

Jurors' tendencies to use certain cognitive or emotional tools when evaluating and interpreting trial information likely guides their attitudes and verdict decisions. Because death penalty attitudes are often used as an exclusionary tool during verdict selection, death qualification might result in a jury of individuals who process information similarly. If homogeneity in information processing among jurors correlates with verdict decisions, death qualification might lead to biased jury decisions indirectly through jurors' information processing tendencies. Jurors who process information analytically, systematically, and strategically, but without intuition or emotion, are considered to process information rationally. In contrast, jurors who tend to process information intuitively or based on emotion are considered to process information experientially (Miller et al. 2014). Both rational and experiential processing are often measured as traits, indicating a relatively stable inclination to process information one way or the other. In this regard, Butler and Moran (2007b) found that death qualification was related to jurors' information processing traits. Specifically, death-qualification resulted in retaining jurors who tended to process information less rationally, using less systematic and analytical cognitive strategies when compared to excludables; this suggests that death-qualified jurors might be less likely to strategically and accurately incorporate trial evidence into their decision-making. This further supports the argument that the death qualification process affects jurors' decisions by excluding individuals who might otherwise provide a more rational understanding and interpretation of trial evidence.

Finally, death qualification not only influences jurors' perceptions of information related to the trial, but it also influences jurors' perceptions of information outside of the trial. Death penalty cases tend to receive more attention from the news and media than other cases, partly because of the gruesome nature of the crimes

(heinous murders, torture, rape, etc.). Therefore, jurors might be exposed to this publicity prior to serving on the jury for that case. Information about a defendant or case that is communicated publicly through various media prior to the beginning of the trial is called pretrial publicity. Pretrial publicity typically evokes juror biases against the defendant, potentially a result of the media's priority and emphasis on pro-death penalty stories (Butler 2007b; Haney 2005; Mowen and Schroeder 2011). Compared to excludables, death-qualified individuals attended to more news television and believed that pretrial publicity would not compromise the defendant's right to a fair trial (Butler 2007b). Therefore, pretrial publicity in death penalty cases might be excessively influential to death-qualified jurors, biasing them against the defendant.

In sum, it appears that death-qualified individuals not only perceive *individuals* differently in death penalty trials, but they also perceive and process the *evidence* differently. These studies suggest that death qualification might favor the prosecution through altered perceptions of evidence and defendants. Because jurors' decisions reflect capital trial structures and processes, biases related to the evaluation of evidence manifest in verdicts and sentence recommendations. The extent to which death qualification impacts jury verdicts and sentences is discussed in the next section.

Death Qualification, Verdicts, and Sentences

The most important issue related to death qualification is whether death qualification actually affects verdicts and sentences. Allen et al. (1998) conducted a meta-analysis, including 14 studies, examining the relationship between death penalty attitudes and the likelihood of conviction in a capital case. Overall, favorable attitudes toward the death penalty were positively related to the likelihood of conviction in a capital case, and this relationship was more pronounced among the studies that included a death qualification process. Including current research (approximately five additional studies not included in the meta-analysis) and the studies in the Allen et al. (1998) meta-analysis, the bulk of the research investigating this issue has confirmed that death-qualified juries are indeed more punitive in terms of their guilt verdicts and sentences (Butler 2007c; Butler and Moran 2007b; Cowan et al. 1984; Horowitz and Seguin 1986; Mauro 1992).

In context, it can be easily concluded that juries composed of individuals able to give the death penalty are more likely to give the death penalty, given conviction. However, the effects of death qualification come under stricter scrutiny when examining *Buchanan v. Kentucky* (1987), in which the defendant, David Buchanan, was tried for a noncapital offense by a death-qualified jury in a joint trial (Whisler 1988). In this case, the jury was death-qualified because his codefendant was being tried for capital murder. Thus, the issue of whether or not death-qualified juries are conviction prone played a substantial role. For Buchanan, the jurors' ability to give the death penalty was unrelated to his charges, yet these criteria were used when

selecting his jury; however, the U.S. Supreme Court ruled that this did not violate Buchanan's rights (*Buchanan v. Kentucky* 1987). This case highlights the importance of research that investigates the link between death qualification and conviction proneness.

As Haney (2005) pointed out, death-qualified jurors are asked to confirm that they are able to give the death penalty, and this confirmation likely biases them toward convicting the defendant. By thinking, "given the defendant is guilty, I will be able to give the death penalty," jurors may have already committed to a verdict choice, one they believe the judge wants them to pursue. This might bias jurors against the defendant before hearing any evidence. This mindset might then serve as a foundation in which the jurors interpret and comprehend trial information (Haney 2005). As such, the death qualification process is a suggestive and influential process that leads to conviction prone juries by creating juries that are less likely to share attitudes in opposition of giving a death penalty and reaffirming jurors' willingness and expectations to convict the defendant and sentence him to death.

In the verdict phase of the trial, juries must decide if the defendant is guilty of the crime. The verdict options typically are guilty or not guilty, but some jurisdictions allow for others (e.g., verdict of not guilty by reason of insanity or guilty but mentally ill). Several studies have tested the effects of death qualification on choosing a guilty verdict over a not guilty verdict. In a majority of the studies, death-qualified jurors were more likely than excludables to choose a guilty verdict over a not guilty verdict (Butler and Moran 2007b; Cowan et al. 1984; Horowitz and Seguin 1986).[2] Although it is reasonable to assume that excluding jurors who oppose the death penalty will result in juries that favor the death penalty more than juries that included jurors who oppose the death penalty, Haney (1984) highlights the independent effects of exposure to the formal death qualification process on conviction proneness.

Haney (1984) showed death-qualified participants a video of a jury selection. Half of the participants watched a noncapital voir dire and half of the participants watched the same voir dire with an additional 30-min segment of a death qualification process. The juries who had watched the death qualification process were more likely to convict the defendant even though they experienced the same trial materials as the group that did not watch the death qualification video. It is notable that the participants did not themselves participate in the death qualification process—they merely watched people in a video go through the process. It is possible that personally experiencing death qualification (e.g., answering the judge's questions) would influence a juror to be even more likely to convict the defendant. It is also notable that all participants in this study were death-qualified; thus, there are not only differences *between* death-qualified juries and excludables—but also

[2]One important note is that some of these studies used different exclusion criteria based on the year they were conducted. Different exclusion variables can lead to different jury compositions (Neises and Dillehay 1987; Seltzer et al. 1986; Dillehay and Sandys 1996).

differences *among* death-qualified jurors (Haney 1984). The results of this study suggest that being exposed to the death qualification process increases the likelihood of conviction *independent* of individual attitudes related to death penalty support (see also Allen et al. 1998). Thus, the process itself has an effect on jurors' verdict decisions over and above individual characteristics and attitudes.

Similar to its effects on *verdicts*, death qualification might also influence *sentences*. By thinking "I'm death-qualified," jurors might believe they should give the death sentence because they are qualified to do so. Death qualification likely increases the chance that a death penalty will be chosen because finding the defendant guilty of a crime worthy of the death penalty implies its imposition. In several studies, death qualification increased the likelihood of sentencing the defendant to death over life in prison (Haney 1984; Horowitz and Seguin 1986; Jurow 1971; Mauro 1992). For instance, Butler (2007c) found that jurors who were death-qualified were more likely to give a death sentence compared to jurors who were excludable. As with verdicts, jurors' sentencing behaviors are affected by the death qualification process itself.

Death qualification systematically excludes jurors who are considered unable to be impartial. However, this process inherently possesses several mechanisms that increase punitive attitudes and facilitate a higher likelihood of convicting and sentencing a defendant to death (Haney 2005). During death qualification, jurors are included based on their willingness to give the death penalty, reinforcing the desirability of pro-death penalty attitudes and behaviors. Also, jurors affirm their position by answering several questions about their death penalty attitudes with statements that reinforce and potentially solidify their willingness to give the death penalty upon a conviction. Death-qualified jurors are reoriented with a pro-death penalty mental framework during the trial, leading to increased receptivity to guilt confirming evidence and aggravating factors while simultaneously rejecting innocence confirming evidence and mitigating factors. During the decision-making process, jurors are primed with death penalty cognitions, leading to the desire to deliberate about the defendant's sentence. This results in hastened verdict decisions in which guilt is ultimately rewarded because it is required in order to advance the jury to deliberations about the death penalty, a penalty they were *qualified* and *authorized* to give (Haney 2005). These mechanisms, in combination, appear to link structural components of death qualification to increased convictions and death sentences.

Theoretical Applications and Extensions

Though research tends to support the notions that death qualification alters jury composition, influences trial perceptions, and affects verdicts and sentences, it is critical to consult social psychological theory to better understand these phenomena. The field of social psychology and law has provided various potential theoretical explanations as to how and why death qualification impacts various stages of

the trial process. Each of these theoretical approaches contributes to the current field's understanding of death qualification and its effects on jurors and juries, and combined, they offer significant insight into these effects and provide a foundation for future research.

Regarding jury composition, when attorneys exclude potential jurors, they might use statistics (e.g., showing that men are more in favor of the death penalty than women) to increase the possibility of creating a jury that will decide in their favor. Often, however, attorneys are likely to base their exclusions on intuitions related to certain attitudes and beliefs associated with individuals and their group statuses. This method of exclusion might have social psychological theoretical underpinnings. Individuals often hold lay theories about the world (Furnham 1988; Wegener and Petty 1998). For instance, jurors might believe that people do not deserve to be put to death no matter the crime. In their perceptions and understanding of the world, life might be considered a fundamental right and that no human has the right to determine life and death of others. Other jurors might believe that murderers deserve to die in order to deter crime, despite insufficient evidence supporting the effectiveness of the death penalty as a deterrent (Donohue and Wolfers 2005, 2009; Katz et al. 2003; Kirchgassner 2011; Land et al. 2012; Siennick 2012). These implicit beliefs and theories about crime and punishment might stem from religious beliefs, personal beliefs, or other determinants.

Understanding jurors' lay beliefs provides information to judges and attorneys about how jurors might perceive the defendant, victims, or witnesses. Based on this information, judges and attorneys might attempt to predict how jurors will vote in that particular trial. Specifically, attorneys might try to identify and exclude jurors who would vote against their side.

Judges and attorneys might also use their own lay theories to exclude jurors during voir dire. Darrow (1936) once suggested that prosecuting attorneys should seek jurors of specific Christian denominations because they are more likely to render guilty verdicts. Further, because attorneys might assume that Catholics are strongly opposed to the death penalty (and thus should be excluded from death penalty cases), attorneys might be more likely to exclude all Catholics. This might occur despite the fact that many Catholics support the death penalty (Carroll 2004).

From this research, we might suggest that there is some systematic evaluation and exclusion of jurors based on social or demographic characteristics related to shared lay theories and beliefs—especially when no statistics are available. However, these lay theories are not necessarily accurate portrayals of reality[3] or formal scientific theories. Nevertheless, they can still influence jury selection.

Regarding trial processes, one of the main issues related to death-qualified jurors is that the death qualification process itself influences jurors such that jurors who experience the death qualification process hold more pro-prosecution or

[3]Although Catholics tend to be excluded from death-qualified juries because of the Catholic Church's position and because they are believed to be opposed to the death penalty, nearly 60 % favor the death penalty (Kohut et al. 2012).

anti-defendant perceptions of the trial information than jurors who do not experience death qualification (Haney 2005). This argument centers on the notion that death qualification activates a mentality that is biased against the defendant.

Death qualification requires that jurors openly claim that they are able to give the death penalty if the defendant were to be found guilty and deserving of the death penalty under the guidance of the law. This commitment is agentic in nature. Rather than simply claiming impartiality, jurors are required to claim their willingness to consider sentencing a person to death. Thus, they are primed with the idea that the death penalty is an appropriate punishment for defendants found guilty of a capital crime (Haney 2005). They are also primed with the belief that they must decide both verdict and sentence based upon the entire trial. Therefore, death-qualified juries might be more likely to interpret trial-related information within this pro-death mental framework (meaning that when the evidence is presented, jurors are considering both guilt and death). By contemplating sentencing during the verdict phase of the trial, jurors are preparing themselves for a guilty verdict. Thus, death-qualified jurors might perceive the evidence against the defendant as more deserving of the death penalty (and a guilty verdict), when compared to excludables.

This effect might be a result of several different processes. Haney (2005) argues that judges and attorneys repeatedly ask jurors about their attitudes toward the death penalty; this bolsters these attitudes if they are in favor of the death penalty and questions their attitudes if they are opposed. Haney suggests that individuals who initially oppose giving the death penalty are asked repeatedly whether they are able to give the death penalty, *assuming* the defendant is *guilty*. This simple repetition of questions reiterates the statement that the defendant is guilty. This then biases jurors to think the defendant is guilty because they have contemplated an assumption of guilt related to this decision. Additionally, judges and attorneys might ask their death qualification questions repeatedly to make sure that the juror should not be excluded. However, this can often lead to jurors changing their minds (Haney 2005). Shuy (1995) even provides evidence that judges ask these types of questions in ways leading the jurors to answer questions consistent with being death-qualified. This leads to the conclusion that the death qualification process itself increases the likelihood of a potential juror to be death-qualified *and* assume guilt.

Assuming death qualification increases the likelihood of convictions and the assumption of a defendant's guilt, the process might not necessarily have an impact on death sentencing. While it is quite basic to assume that individuals who are excluded because of opposition against the death penalty would lower the likelihood of a jury's decision to sentence a defendant to death, this does not suggest that the death qualification *process* increases the likelihood of choosing a death sentence. However, Haney (1997, 2005) argues just this point. He proposes that death-qualified juries in capital cases experience moral disengagement, distancing jurors from the moral implications of their actions. This occurs through a process he calls "structural aggravation," in which death qualification and various aspects of death penalty trials facilitate and perpetuate more punitive responses toward

convicted defendants by devaluing the defendant's life and enhancing perceptions of future dangerousness (Haney 1997, 2005). Thus, death qualification might alter the ways in which jurors process information during capital trials.

Death qualification might also alter the ways in which jurors process information, such that death-qualified jurors might experience emotions that influence the way they process and interpret trial-related information (see Epstein 2008; Slovic and Peters 2006). Epstein's (1990) cognitive experiential self-theory suggests that individuals process information in two ways, *rationally* (analytical, calculating, and logical) and *experientially* (heuristic, emotional, and intuitive). Experiential processing leads to fewer cognitive evaluations and more decisions based on feelings and emotions (Epstein 1990). Death qualification might exclude jurors who naturally process information more rationally, resulting in a jury that processes information more experientially. This is partially supported in that death-qualified individuals reported lower trait rational processing scores than excludables (Butler and Moran 2007b). Moreover, individuals who experience heightened emotional *states* are more likely to process information experientially at the time of emotional experience (Epstein 1990). Jurors often experience intense emotions during capital trials, including the death qualification process, which suggests that capital jurors might be at an increased risk of processing trial-related information experientially (see Antonio 2008). Experiential processing might influence jurors' interpretation of evidence and perceptions of legal professionals, resulting in more intuitive evaluations driven by a "pro-death" mental framework compared to more impartial and rational evaluations.

In addition, priming and automaticity research suggests that individuals can be primed with certain ideas or beliefs, activating a mental schema or framework on which individuals then base their decisions (Bargh and Chartrand 2000). Death qualification might prime a certain set of beliefs or attitudes, and jurors might seek information confirming these beliefs. For instance, if death-qualified jurors are primed with beliefs that the defendant is guilty, they might seek information that confirms these beliefs and disregard information that opposes these beliefs, an effect called confirmatory bias (Rabin and Schrag 1999). Moreover, priming specific beliefs might also lead to different decisions. For example, Yelderman and Miller (2015) found that individuals who were primed with religious beliefs related to punitive attitudes and decisions were more punitive in their verdict and sentencing decisions (compared to individuals who were not primed). Although similar to the influence of emotion on information evaluation, priming beliefs might also lead jurors to disregard certain information altogether, rather than simply weighing it differently. This process would then lead to a greater propensity to convict the defendant and give a death sentence—as discussed next.

Regarding verdicts and sentences, these decisions might be influenced by several of the same processes that influence information processing related to perceptions of judges, attorneys, defendants, victims, testimony, and evidence (discussed in the last section). However, two new theoretical applications suggest that verdicts and sentences might also be influenced by unique processes including cognitive dissonance and subjective decision-making thresholds.

Springer and Lalasz (2014) suggest that jurors who are death-qualified experience psychological dissonance throughout the process. When jurors are initially death-qualified, they answer questions in ways that attest to the fact that they are able to give a death penalty based on their attitudes toward the death penalty. They reiterate this ability by stating that they are able to give the death penalty, assuming the guilt of the defendant. As Haney (2005) stated, this answering period emphasizes the assumption of guilt even though the trial has not started. Thus, death-qualified jurors enter the trial with a *biased* presumption of innocence instead of a complete presumption of innocence. Death-qualified individuals are more likely to assume the defendant is guilty than excludable jurors prior to the trial. Because jurors know that they are expected to assume innocence, yet might also assume guilt due to participation in the death qualification process, they experience stress or inner turmoil (Springer and Lalasz 2014). This then leaves them with an unpleasant feeling of cognitive dissonance. To eliminate this feeling, they might justify their beliefs of guilt by convicting the defendant and subsequently voting for the death penalty (Springer and Lalasz 2014).

Other approaches relate to individual differences between death-qualified and excludable jurors. Thompson et al. (1984) suggest that death-qualified individuals have a lower threshold for conviction. Because of the biased information processes and perceptions of trials, death-qualified jurors need less proof to return a guilty verdict and death sentence. This was based on the finding that death-qualified individuals anticipated being less regretful for erroneous convictions and more regretful for erroneous acquittals (Thompson et al. 1984). This might suggest that individuals with lower standards for conviction are more likely to be death-qualified or that, once death-qualified, standards for guilt and giving a death penalty decrease. Ellsworth et al. (1984) suggest that excludables are more likely to value due process over efficient crime control, thus leading them to be less punitive. This comports with Butler and Moran's (2007a) finding that death-qualified individuals tend to have higher legal authoritarian and just world beliefs. Such traits might lead to a higher propensity to convict and favor the death penalty.

Lastly, the experience of the death qualification process might lead to higher convictions and more death penalties because jurors are doing what they believe the judge wants them to do. Simply labeling jury service as jury "duty" implies the moral and honorable commitment of serving on a jury, which is often reinforced by judges and attorneys (Dillehay and Sandys 1996). Death-qualified jurors might fulfill this *moral* and *honorable* obligation by giving the death penalty. Also, individuals tend to obey authorities (Burger 2009; Milgram 1963); as judges are the ultimate authority in the courtroom, jurors are likely to take their cues from the judge. During the process, the judge could ask jurors if they could set aside their objections and give a death penalty (assuming the defendant is found guilty), if they could do their civic duty and be an unbiased juror, and if they are *so* strongly opposed to it that they could not set aside their opinions. Such questions imply that the "correct" thing to do is to find the defendant guilty and give the death penalty, even if they have to ignore their own opinions (though these opinions might resurface during the trial or deliberations; Dillehay and Sandys 1996).

Implications, Limitations, and Future Directions

Death qualification appears to be a process that systematically biases the jury against the defendant. Because of this, defendants in death penalty cases might be at an initial disadvantage, facing an increased likelihood of being found guilty and sentenced to death. This result seems to occur for three reasons. First, individuals with characteristics related to punitiveness are more likely to be selected for capital juries than their counterparts. Second, once they are death-qualified, jurors tend to interpret trial information differently compared to those who are not selected. This interpretation includes biased processing of evidence and biased perceptions of the defendant and victims. Third, the process of death qualification and the individual traits of those selected through death qualification both relate to higher rates of convictions and death sentences. In sum, it appears as if the death qualification process facilitates convictions and death sentences (see also Haney 2005).

Implications for judges and attorneys begin with the voir dire process. Understanding how the process of death qualification during voir dire can lead to the systematic exclusion of potential jurors might inform judges' and attorneys' current practices so that potential jurors are interviewed and screened with the intention to reduce or minimize bias toward conviction and death. Though judges and attorneys might assume that their selection process is unbiased, judges and attorneys should think introspectively and address the extent to which their own lay theories drive exclusion decisions.

Also, understanding how death qualification might influence jurors' information processing during trials might provide insights into attorney arguments and evidence presentation. Research suggests that death qualification relates to jurors' weighing aggravating factors more and mitigating factors less. Because of this research, attorneys might choose to emphasize aggravating and mitigating circumstances differently. For example, defense attorneys might use strategies to clarify and maximize understanding of various mitigating factors or use strategies to discredit or call into question the aggravating factors as likely they already do. Attorneys could point out jurors' biases and ask them to actively correct for these biases—assuming they can do so. In contrast, prosecuting attorneys might frame mitigating factors as aggravating, taking advantage of death-qualified jurors' tendency to interpret mitigating information skeptically and negatively. Enhancing mitigating evidence and minimizing aggravating evidence might also reduce jurors' moral disengagement and humanize the defendant. Several academic and legal scholars have suggested altering various aspects of the death qualification and voir dire process in capital trials to increase jury impartiality (e.g., Cox and Tanford 1989; Harvard Law Review 2014), and integrating current scientific research on the effect of death qualification on juries' decisions might help, at the very least, approach this goal.

Although each study reviewed here has limitations (e.g., fictional trials, non-representative samples, and lack of deliberation), overall this research is quite persuasive. Several of these relationships have been replicated in multiple studies,

and meta-analyses support many of these death qualification effects. Further, findings converge to show the breadth of the ways in which death qualification can affect trial processes and trial outcomes. However, general limitations still exist. The current literature needs to address the lack of verisimilitude by using more realistic trials. Experimental and survey studies are often criticized for not resembling real trials (Finch and Ferraro 1986). Moreover, the legitimacy of social science research's ability to influence court decisions has been contested as a result of perceived weak methodological validity and rigor (i.e., *Lockhart v. McCree* 1986; see also Thompson 1989). Studies need to use actual trials, trial videos, or reenactments in order to increase the realism related to participating in a capital trial. Although ecological validity is not always a major issue in research, it is still important to determine whether results hold with more realistic scenarios (see Bornstein 1999). Research also needs to include generalizable samples (Wiener et al. 2011). Although researchers often acquire venirepersons (people from actual jurisdictional jury pools) for death qualification research, many samples still consist of convenience and student samples, which may or may not influence the results (Bornstein 1999; Chomos and Miller 2015).

Another limitation is that there is no clear conclusion about exactly how the death qualification process affects jurors' information processing. Because research results are still mixed, future research should address this. Lastly, there is almost no inclusion of deliberations in death qualification research. Jury deliberations involve several aspects that can influence jury verdicts, including deliberation quality, structure, and content; foreperson characteristics and stance; majority and minority shifts; and polling procedures (Devine et al. 2001, 2007). Some studies suggest that jury deliberations might produce a leniency effect, decreasing the likelihood of conviction (e.g., MacCoun and Kerr 1988; Miller et al. 2011); however, other studies suggest that juries, specifically capital juries, do not show a leniency bias but often become more punitive after deliberation (e.g., Devine et al. 2004; Lynch and Haney 2009, 2015). This might result from a tendency to conform to the majority or the foreperson's stance or differences in participation, both of which might disproportionately reflect a pro-death view (Devine et al. 2007; see also, Cornwell and Hans 2011). Future research should strive to understand how death-qualified jury deliberations differ from non-death-qualified jury deliberations in decision-making and the independent effects of deliberations on verdict and sentencing decisions.

Future directions for this line of research include testing many of the theoretical explanations provided here. Researchers might utilize new techniques, such as priming techniques, to expand on the existing literature. They might also use statistical methods to determine whether perceptions of trial information (e.g., evidence of aggravation and mitigation) mediate the relationship between death qualification and verdict and sentencing decisions. Researchers might more directly use theories of cognitive dissonance, confirmation bias, processing states and traits. Finally, researchers should use diverse samples of mock jurors and actual jurors, naturalistic settings, and involve realistic trial stimuli in future research to increase the work's ecological validity.

Conclusion

Overall, death qualification has been studied quite extensively (though much of the research was conducted in the 1980s and 1990s), leading to results that suggest the process itself has an effect on jury composition, jurors' perceptions of trial-related information, and verdict and sentence decisions. Though the purpose of death qualification might be to assist judges and attorneys in choosing a jury absent any extreme death penalty biases, it might instead create or perpetuate such biases.

Death-qualified juries are different from non-death-qualified juries in ways beyond the inherent differences in death penalty attitudes. These differences stem from the systematic exclusion of specific social groups and individuals with shared attitudes or beliefs that are thought of as prejudicial toward defendants in death penalty trials. Because death-qualified juries differ on these types of characteristics, they might actually be more susceptible to systematic biases toward defendants.

Death qualification also appears to influence the ways in which individuals interpret trial-related information. By simply being exposed to the death qualification process, jurors perceive various trial-related information in ways that disadvantage the defendant. Thus, beyond individual characteristics, the death qualification process itself creates biases in jurors' understanding of the trial and perceptions of the defendant.

Even more concerning are findings that demonstrate that the death qualification process produces conviction prone juries, independent of jurors' individual death penalty attitudes, that are more willing to give a death sentence to a guilty defendant. Thus, from this research, it is fairly easy to conclude that the death qualification process has important effects on juries and trial outcomes, and these effects should continue to be studied.

References

Allen, M., Mabry, E., & McKelton, D. M. (1998). Impact of juror attitudes about the death penalty on juror evaluations of guilt and punishment: A meta-analysis. *Law and Human Behavior, 22*, 715–731.

Antonio, M. E. (2008). Stress and the capital jury: How male and female jurors react to serving on a murder trial. *The Justice System Journal, 29*(3), 396–407.

Atkins v. Virginia, 536 U.S. 304 (2002).

Banner, S. (2002). *The death penalty: An American history*. Cambridge, MA: Harvard University Press.

Bargh, J. A., & Chartrand, T. L. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 253–285). New York, NY: Cambridge University Press.

Batson v. Kentucky, 476 U.S. 79 (1986).

Bornstein, B. H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior, 23*, 75–91.

Bornstein, B. H., & Miller, M. K. (2009). *God in the courtroom: Religion's role at trial*. New York, NY: Oxford University Press.

Bowers, W. J. (1995). The capital jury project: Rationale, design, and a preview of early findings. *Indiana Law Journal, 70*, 1043–1102.

Buchanan v. Kentucky, 483 U.S. 402 (1987).

Burger, J. M. (2009). Replicating Milgram: Would people still obey today? *American Psychologist, 64*, 1–11. doi:10.1037/a0010932

Butler, B. (2007a). Death qualification and prejudice: The effect of implicit racism, sexism, and homophobia on capital defendants' right to due process. *Behavioral Sciences and the Law, 25*, 857–867. doi:10.1002/bsl.791

Butler, B. (2007b). The role of death qualification in jurors' susceptibility to pre-trial publicity. *Journal of Applied Social Psychology, 37*, 115–123. doi:10.1111/j.0021-9029.2007.00150.x

Butler, B. (2007c). The role of death qualification in capital trials involving juvenile defendants. *Journal of Applied Social Psychology, 37*, 549–560. doi:10.1037/t04485-000

Butler, B. (2008). The role of death qualification in venirepersons' susceptibility to victim impact statements. *Psychology, Crime, & Law, 14*, 133–141. doi:10.1080/10683160701483534

Butler, B. M., & Moran, G. (2002). The role of death qualification in venirepersons' evaluations of aggravating and mitigating circumstances in capital trials. *Law and Human Behavior, 26*(2), 175–184. doi:10.1023/A:1014640025871

Butler, B., & Moran, G. (2007a). The impact of death qualification, belief in a just world, legal authoritarianism, and locus of control on venirepersons' evaluations of aggravating and mitigating circumstances in capital trials. *Behavioral Sciences and the Law, 25*, 57–68. doi:10.1002/bsl.734

Butler, B., & Moran, G. (2007b). The role of death qualification and need for cognition in venirepersons' evaluations of expert scientific testimony in capital trials. *Behavioral Sciences and the Law, 25*, 561–571. doi:10.1002/bsl.758

Carroll, J. (2004). *Who supports the death penalty? Gallup poll*. The Gallup Organization. Retrieved from http://www.deathpenaltyinfo.org/gallup-poll-who-supports-death-penalty

Cho, S. (1994). Capital confusion: The effect of jury instructions on the decision to impose death. *The Journal of Criminal Law and Criminology*, *85*(2), 532–561.

Chomos, J. C., & Miller, M. K. (2015). Understanding how individual differences are related to community sentiment toward safe haven laws using a student sample. In M. K. Miller, J. A. Blumenthal, & J. Chamberlain (Eds.), *Handbook of community sentiment* (pp. 83–97). New York, NY: Springer.

Cornwell, E. Y., & Hans, V. P. (2011). Representation through participation: A multilevel analysis of jury deliberations. *Law & Society Review, 45*(3), 667–698. doi:10.1111/j.1540-5893.2011.00447.x

Costanzo, S., & Costanzo, M. (1994). Life or death decisions: An analysis of capital jury decision making under the special issues sentencing framework. *Law and Human Behavior, 18*(2), 151–170. doi:10.1007/BF01499013

Cowan, C. L., Thompson, W. C., & Ellsworth, P. C. (1984). The effects of death qualification on jurors' predisposition to convict and on the quality of deliberation. *Law and Human Behavior, 8*, 53–79. doi:10.1007/BF01044351

Cox, M., & Tanford, S. (1989). An alternative method of capital jury selection. *Law and Human Behavior, 13*(2), 167–183. doi:10.1007/BF01055922

Cunningham, M. D., Sorensen, J. R., & Reidy, T. J. (2009). Capital jury decision-making: The limitations of predictions of future violence. *Psychology, Public Policy, and Law, 15*(4), 223–256. doi:10.1037/a0017296

Cunningham, M. D., Sorensen, J. R., Vigen, M. P., & Woods, S. O. (2011). Life and death in the lone star state: Three decades of violence predictions by capital juries. *Behavioral Sciences & the Law, 29*(1), 1–22. doi:10.1002/bsl.963

Darrow, C. (1936). Attorney for the defense. *Esquire, 5*, 36.

Dayan, M., Mahler, R. S., & Widenhouse, M. G, Jr. (1989). Searching for an impartial sentencer through jury selection in capital trials. *Loyola of Los Angeles Law Review, 23*(1), 151–192.

Devine, D. J. (2012). *Jury decision making: The state of the science*. New York, NY: New York University Press.

Devine, D. J., Buddenbaum, J., & Houp, S. (2007). Deliberation quality: A preliminary examination in criminal juries. *Journal of Empirical Legal Studies, 4*(2), 273–303. doi:10.1111/j.1740-1461.2007.00089.x

Devine, D. J., Clayton, L. D., Dunford, B. B., Seying, R., & Pryce, J. (2001). Jury decision making: 45 years of empirical research on deliberating groups. *Psychology, Public Policy, and Law, 7*(3), 622–727. doi:10.1037//1076-8971.7.3.622

Devine, D. J., & Kelly, C. (2015). Life or death: An examination of jury sentencing with the capital jury project database. *Psychology, Public Policy, and Law, 21*(4), 393–406. doi:10.1037/law0000060

Devine, D. J., Olafson, K. M., Jarvis, L. L., Bott, J. P., Clayton, L. D., & Wolfe, J. M. T. (2004). Explaining jury verdicts: Is leniency bias for real? *Journal of Applied Social Psychology, 34* (10), 2069–2098. doi:10.1111/j.1559-1816.2004.tb02691.x

Dieter, R. C. (2005, October). *Blind justice: Juries deciding life and death with only half truth: How death penalty jurors are unfairly selected, manipulated, and kept in the dark*. Death Penalty Information Center. Retrieved from http://www.deathpenaltyinfo.org/

Dillehay, R. C., & Sandys, M. R. (1996). Life under *Wainwright v. Witt*: Juror disposition and death qualification. *Law and Human Behavior, 20*, 147–165.

Donohue, J. J., & Wolfers, J. (2005). Use and abuses of the empirical evidence in the death penalty debate. *Stanford Law Review, 58*, 791–846.

Donohue, J. J., & Wolfers, J. (2009). Estimating the impact of the death penalty on murder. *American Law and Economics Review, 11*, 249–309.

Ellsworth, P. C., Bukaty, R. M., Cowan, C. L., & Thompson, W. C. (1984). The death-qualified jury and the defense of insanity. *Law and Human Behavior, 8*(1/2), 81–93. doi:10.1007/BF01044352

Epstein, S. (1990). Cognitive-experiential self-theory. In H. Plessner, C. Betsch, & T. Betsch (Eds.), *Intuition in judgment and decision making* (pp. 23–37). New York, NY: Erlbaum.

Epstein, S. (2008). Intuition from the perspective of cognitive-experiential self-theory. In H. Plessner, C. Betsch, & T. Betsch (Eds.), *Intuition in judgment and decision making* (pp.23–37). New York, NY: Erlbaum.

Finch, M., & Ferraro, M. (1986). The empirical challenge to death-qualified juries: On further examination. *Nebraska Law Review, 65*, 21–74.

Fitzgerald, R., & Ellsworth, P. C. (1984). Due process vs. crime control: Death qualification and jury attitudes. *Law and Human Behavior, 8*, 31–51. doi:10.1007/BF01044350

Foley, L. A., & Powell, R. S. (1982). The discretion of prosecutors, judges, and juries in capital cases. *Criminal Justice Review, 7*(2), 16–22. doi:10.1177/073401688200700203

Ford v. Wainwright, 477 U.S. 399 (1986).

Foster v. Chatman, 578 U.S. 399 (2016).

Furman v. Georgia, 408 U.S. 238 (1972).

Furnham, A. (1988). *Lay theories*. New York, NY: Pergamon Press.

Ghoshray, S. (2013). Capital jury decision making: Looking through the prism of social conformity and seduction to symmetry. *University of Miami Law Review, 67*(2), 477–519.

Gonzalez-Perez, M. (2001). A model of decision making in capital juries. *International Social Science Review, 76*(3/4), 79–91.

Gonzalez-Perez, M. (2002). The potential for bias in capital juries. *The Justice System Journal, 23* (2), 235–247.

Grasmick, H. G., Cochran, J. K., Bursik, R. J., & Kimpel, M. (1993). Religion, punitive justice and support for the death penalty. *Justice Quarterly, 10*, 289–314.

Gregg v. Georgia, 428 U.S. 153 (1976).

Haney, C. (1984). Examining death qualification: Further analysis of the process effect. *Law and Human Behavior, 8*, 133–151. doi:10.1007/BF01044356

Haney, C. (1997). Violence and the capital jury: Mechanisms of moral disengagement and the impulse to condemn to death. *Stanford Law Review, 49*(6), 1447–1486.

Haney, C. (2005). *Death by design: Capital punishment as a social psychological system*. New York, NY: Oxford University Press.

Haney, C., Hurtado, A., & Vega, L. (1994). "Modern" death qualification: New data on its biasing effects. *Law and Human Behavior, 18*, 619–633. doi:10.1007/BF01499328

Hans, V. P. (1995). How juries decide death: The contributions of the capital jury project. *Indiana Law Journal, 70*(4), 1233–1240.

Harvard Law Review. (2014). Live free and nullify: Against purging capital juries of death penalty opponents. Note. *Harvard Law Review, 127*(7), 2092–2112.

Horowitz, I. A., & Seguin, D. G. (1986). The effects of bifurcation and death qualification on assignment of penalty in capital crimes. *Journal of Applied Social Psychology, 16*(2), 165.

J.E.B. v. Alabama, 511 U.S. 127, 1994.

Jurow, G. L. (1971). New data on the effect of a "death qualified" jury on the guilt determination process. *Harvard Law Review, 84*(3), 567–611.

Kadane, J. B. (1984). After Hovey: A note on taking account of the automatic death penalty jurors. *Law and Human Behavior*, *8*, 115–120.

Katz, L., Levitt, S. D., & Shustorovich, E. (2003). Prison conditions, capital punishment, and deterrence. *American Law and Economics Review, 5*, 318–343. doi:10.1093/aler/ahg014

Kennedy v. Lousiana, 554 U.S. 407 (2008)

Kirchgassner, G. (2011). Econometric estimates of deterrence of the death penalty: Facts or ideology? *KYKLOS, 64*, 448–478. doi:10.1111/j.1467-6435.2011.00515.x

Kirchmeier, J. (2006). Casting a wider net: Another decade of legislative expansion of the death penalty in the United States. *Pepperdine Law Review, 1*, 1–40.

Kohut, A., Doherty, C., Dimock, M., & Smith, G. (2012). *More concern among opponents about wrongful convictions: Continued majority support for death penalty*. Pew Research Center. Retrieved from http://www.people-press.org/files/legacy-pdf/1-6-12%20Death%20penalty%20release.pdf

Land, K. C., Teske, R. H. C., & Zheng, H. (2012). The differential short-term impacts of executions on felony and non-felony homicides. *Criminology & Public Policy, 11*, 541–563. doi:10.1111/j.1745-9133.2012.00834.x

Latzer, B., & McCord, D. (2011). *Death penalty cases: Leading U.S. Supreme Court cases on capital punishment*. Burlington, MA: Elsevier.

Lieberman, J. D., & Olson, J. (2009). The psychology of jury selection. In J. D. Lieberman & D. A. Krauss (Eds.), *Jury psychology: Social aspects of trial processes* (pp. 97–128). Burlington, VT: Ashgate Publishing Company.

Lockhart v. McCree, 476 U.S. 162 (1986).

Luginbuhl, J., & Howe, J. (1995). Discretion in capital sentencing instructions: Guided or misguided? *Indiana Law Journal, 70*(4), 1161–1181.

Luginbuhl, J., & Middendorf, K. (1988). Death penalty beliefs and jurors' responses to aggravating and mitigating circumstances in capital trials. *Law and Human Behavior, 12*(3), 263–281. doi:10.1007/BF01044384

Lynch, M., & Haney, C. (2009). Capital jury deliberation: Effects on death sentencing, comprehension, and discrimination. *Law and Human Behavior, 33*(6), 481–496. doi:10.1007/s10979-008-9168-2

Lynch, M. & Haney, C. (2011). Mapping the racial bias of the white male capital juror: Jury composition and the "empathic divide." *Law & Society Review, 45*(1), 69–101. doi:10.1111/j.1540-5893.2011.00428.x

Lynch, M., & Haney, C. (2015). Emotion, authority, and death: (raced) Negotiations in mock capital jury deliberations: Emotion, authority, and death. *Law & Social Inquiry, 40*(2), 377–405. doi:10.1111/lsi.12099

MacCoun, R. J., & Kerr, N. L. (1988). Asymmetric influence in mock jury deliberation: Jurors' bias for leniency. *Journal of Personality and Social Psychology, 54*(1), 21–33. doi:10.1037/0022-3514.54.1.21

MacDonald, G. J. (2015, January 25). *Boston bombing jury excludes some Catholics*. USA Today. Retrieved from http://www.usatoday.com/story/news/nation/2015/01/25/boston-bombing-jury-selection-excludes-observant-catholics/22121061/

Mauro, R. (1992). Tipping the scales toward death: The biasing effects of death qualification. In P. Suedfeld & P. Tetlock (Eds.), *Psychology and social policy* (pp. 243–254). New York, NY: Hemisphere Press.

McGarrell, E. F., & Sandys, M. (1996). The misperception of public opinion toward capital punishment: Examining the spuriousness explanation of death penalty support. *American Behavioral Scientist, 39*(4), 500–513. doi:10.1177/0002764296039004012

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology, 67*, 371–378.

Miller, M. K., & Hayward, R. D. (2008). Religious characteristics and the death penalty. *Law and Human Behavior, 32*, 113–123. doi:10.1007/s10979-007-9090-z

Miller, M. K., Maskaly, J., Green, M., & Peoples, C. D. (2011). The effects of deliberations and religious identity on mock jurors' verdicts. *Group Processes & Intergroup Relations, 14*(4), 517–532. doi:10.1177/1368430210377458

Miller, M. K., Wood, S. M., & Chomos, J. C. (2014). Relationships between support for the death penalty and cognitive processing: A comparison of students and community members. *Criminal Justice and Behavior, 41*, 732–750. doi:10.1177/0093854813509369

Morgan, K., & Mannheimer, M. J. Z. (2009). The impact of information overload on the capital jury's ability to assess aggravating and mitigating factors. *The William and Mary Bill of Rights Journal, 17*(4), 1089–1137.

Morgan v. Illinois, 504 U.S. 719 (1992).

Mowen, T. J., & Schroeder, R. D. (2011). Not in my name: An investigation of victims' family clemency movements and court appointed closure. *Western Criminology Review, 12*, 65–81.

Neises, M. L., & Dillehay, R. C. (1987). Death qualification and conviction proneness: Witt and Witherspoon compared. *Behavioral Sciences & the Law, 5*, 479–494.

O'Neil, K. M., Patry, M. W., & Penrod, S. D. (2004). Exploring the effects of attitudes toward the death penalty on capital sentencing verdicts. *Psychology, Public Policy, and Law, 10*, 443–470.

Palmer, L. J, Jr. (2014). *The death penalty in the United States: A complete guide to federal and state laws*. Jefferson, NC: McFarland & Company Inc.

Price, M. J. (2009). Performing discretion or performing discrimination: Race, ritual, and peremptory challenges in capital jury selection. *Michigan Journal of Race & Law, 15*(1), 57–107.

Rabin, M., & Schrag, J. L. (1999). First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics, 114*, 37–82.

Reidy, T. J., Sorensen, J. R., & Cunningham, M. D. (2013). Probability of criminal acts of violence: A test of jury predictive accuracy: Capital jury predictive accuracy. *Behavioral Sciences & the Law, 31*, 286–305. doi:10.1002/bsl.2064

Ring v. Arizona, 536 U.S. 584 (2002).

Roper v. Simmons, 543 U.S. 551 (2005).

Sandys, M., & Trahan, A. (2008). Life qualification, automatic death penalty voter status, and juror decision making in capital cases. *Justice Systems Journal, 29*, 385–395.

Sarma, B. J. (2012). Challenges and opportunities in bringing the lessons of cultural competence to bear on capital jury selection. *The University of Memphis Law Review, 42*(4), 907–934.

Seltzer, R., Lopes, G. M., Dayan, M., & Cana, R. F. (1986). The effect of death qualification on the propensity of jurors to convict: The Maryland example. *Howard Law Journal, 29*, 571–608.

Shuy, R. (1995). How a judge's voir dire can teach a jury what to say. *Discourse & Society, 6*, 207–222. doi:10.1177/0957926595006002004

Siennick, S. E. (2012). Deterrence and the death penalty: A new look at the some recent findings. *Criminology & Public Policy, 11*, 535–537. doi:10.1111/j.1745-9133.2012.00833.x

Slovic, P., & Peters, E. (2006). Risk perception and affect. *Current Directions in Psychological Science, 15*, 322–325. doi:10.1111/j.1467-8721.2006.00461.x

Smith, R. J. (2012). The geography of the death penalty and its ramifications. *Boston University Law Review, 92*, 227–289.

Snell, T. L. (2014). *Capital Punishment, 2012—Statistical Tables*. U.S. Department of Justice, Bureau of Justice Statistics, Office of Justice Programs. Retrieved January 18, 2016 from http://www.bjs.gov/

Springer, V., & Lalasz, C. B. (2014). Death-qualified jurors and the assumption of innocence: A cognitive dissonance perspective on conviction-prone verdicts. *The Social Science Journal, 51* (2), 287–294. doi:10.1016/j.soscij.2013.08.010.

Stack, S. (2003). Authoritarianism and support for the death penalty: A multivariate analysis. *Sociological Focus, 36*, 333–352.

Stevenson, M. C., Bottoms, B. L., & Diamond, S. S. S. (2010). Jurors' discussions of a defendant's history of child abuse and alcohol abuse in capital sentencing deliberations. *Psychology, Public Policy, and Law, 16*(1), 1–38. doi:10.1037/a0018404

Sullivan, J. T. (2014). The demographic dilemma in death qualification of capital jurors. *Wake Forest Law Review, 49*(4), 1107–1172.

Summers, A., Hayward, R. D., & Miller, M. K. (2010). Death qualification as systematic exclusion of jurors with certain religious and other characteristics. *Journal of Applied Social Psychology, 40*, 3218–3234. doi:10.1111/j.1559-1816.2010.00698.x

Swafford, A. T. (2011). Qualified support: Death qualification, equal protection, and race. *American Journal of Criminal Law, 39*(1), 147.

Thompson, W. C. (1989). Death qualification after Wainwright v. Witt and Lockhart v. McCree. *Law and Human Behavior, 13*, 185–215.

Thompson, W. C., Cowan, C. L., Ellsworth, P. C., & Harrington, J. C. (1984). Death penalty attitudes and conviction proneness: The translation of attitudes into verdicts. *Law and Human Behavior, 8*, 95–113. doi:10.1007/BF01044353

Uttecht v. Brown, 551 U.S. 1 (2007).

Vidmar, N., & Hans, V. P. (2007). *American juries: The verdict*. Amherst, NY: Prometheus Books.

Wainwright v. Witt, 470 U.S. 1039 (1985).

Wegener, D. T., & Petty, R. E. (1998). The naïve scientist revisited: Naïve theories and social judgment. *Social Cognition, 16*, 1–7. doi:10.1521/soco.1998.16.1.1

Whisler, B. J. (1988). Sixth amendment: Death qualification of the jury: Process is permissible where defendant does not face death penalty. *The Journal of Criminal Law and Criminology*, *78*(4), 954–983.

Wiener, R. L., Krauss, D. A., & Lieberman, J. D. (2011). Mock jury research: Where do we go from here? *Behavioral Sciences & the Law, 29*(3), 467–479. doi:10.1002/bsl.989

Witherspoon v. Illinois, 391 U.S. 510 (1968).

Yelderman, L. A., & Miller, M. K. (2015, February). *Religious fundamentalism, religious priming, and mock juror verdicts involving the insanity defense*. Paper session presented at the American Psychology and Law Society annual conference in San Diego, CA.

Young, R. L. (1992). Religious orientation, race and support for the death penalty. *Journal for the Scientific Study of Religion, 31*, 76–88.

Young, R. L. (2000). Punishment at all costs: On religion, convicting the innocent, and supporting the death penalty. *William & Mary Bill of Rights Journal, 9*, 237–245.

The Impact of Emotions on Juror Judgments and Decision Making

Narina Nunez, Victoria Estrada-Reynolds, Kimberly Schweitzer and Bryan Myers

In the final days of the case against Dzhokhar Tsarnaev for the Boston Marathon bombing, the prosecution presented the testimony of Dr. Henry Nields, a medical examiner, who spoke at length and presented grisly photographs into evidence depicting the charred and broken remains of 8-year-old Martin Richard, who died in the explosion. While watching the accompanying photographs, the witness spoke in vivid detail about the boy's body, his broken bones and burned skin covered with nails and wood fragments, and how his intestines had spilled outside his body. In hearing the testimony, jurors "hung their heads in their hands. At least three were crying. Two or three looked as if they were staring into the abyss" (Seelye 2015, A-16).

The above example illustrates the contradiction that jurors face in their decision making. A trial can be a highly emotionally charged event, yet jurors are often asked to put aside their emotions in rendering their decisions. The role of emotions in legal decision making has been debated extensively (e.g., see Bandes 1996), with the traditionally held opinion by legal scholars that emotions largely impede rational decisions and sound judgments (Laster and O'Malley 1996; Posner 1999). Recent work in emotion and decision making literatures, however, suggests that such sweeping generalizations are incorrect, or at the very least, are incomplete. In fact, there is a growing body of literature suggesting that certain emotions might have a detrimental effect on decision making, while others may not.

Blame and responsibility are central to jurors' decision making in both civil and criminal contexts. In civil contexts, these attributions feature most prominently in

N. Nunez (✉) · V. Estrada-Reynolds · K. Schweitzer
Department of Psychology, University of Wyoming, 1000 E. University Ave., Laramie, WY 82071, USA
e-mail: narina@uwyo.edu

B. Myers
Department of Psychology, University of North Carolina – Wilmington, Teaching Laboratory Building, 601 S. College Rd., Wilmington, NC 28403-5612, USA

B.H. Bornstein and M.K. Miller (eds.), *Advances in Psychology and Law*, Advances in Psychology and Law 2, DOI 10.1007/978-3-319-43083-6_3

determining liability, though they can also factor into damages, especially punitive damages. In criminal contexts, they feature most prominently in determining the defendant's culpability, which has implications both for guilt judgments and sentencing (especially in capital cases). The present chapter incorporates research on emotional influences on both civil and criminal juror and will outline what we know about emotions in the courtroom, how they affect decisions, and what we need to focus on in future research.

Emotions have not received a great deal of attention among psycholegal researchers, but interest has grown of late. In 2006, *Law and Human Behavior* devoted a special issue to emotions and legal decision making. In one article, Bornstein and Wiener (2006) noted that the law at times promotes the use of emotions in decisions (e.g., determining culpability of a hate crime or determining mental suffering in a civil trial), and at other times eschews the use of emotions in decision making (e.g., requiring jurors to view gruesome crime scene photos dispassionately). They further pointed out that psycholegal researchers had only "scratched the surface of this vibrant field" (Bornstein and Wiener 2006, p. 117). Since this publication, there has been a burgeoning interest in studying how emotions affect legal decisions. As scholars, we have learned much about emotions and legal decision making, yet much remains unclear, and so we make some recommendations for future areas of inquiry.

In the first section of this chapter, we focus on the basic theories surrounding emotions and judgments, as well as the degree to which the general theories of emotions and decision making are relevant to decisions made in the courtroom. Some theories predict that emotions can alter jurors' processing styles, and may indirectly affect decisions (Tiedens and Linton 2001), yet others suggest that emotions are likely to affect decisions when a particular information processing style is active (Forgas 1995). Further, emotions have been theorized to impact the value we assign to particular information and what information we attend to (Tetlock et al. 2007) or even recall (Fishfader et al. 1996). Finally, theories of emotion suggest that emotions may be informative, telling us how we feel about things, and therefore acting as direct information for judgments and decision making (Clore et al. 2001).

Next, we examine factors during the course of the trial that may prompt specific emotional reactions from jurors. The trial itself is a dramatic event, in which prosecutors and defense attorneys seek to evoke emotions in jurors (e.g., sympathy, anger) in the hopes of making the jury more receptive to their particular case. Throughout the trial, there are various occasions in which emotional responses by jurors are likely to be substantial. For example, the graphic nature by which a particularly brutal act is described or the presentation of gruesome crime scene photos might elicit disgust in jurors. When sentencing the defendant in capital trials, jurors may learn of the suffering the victim's family has experienced as a result of the loss of a loved one, and this victim impact evidence might elicit strong reactions such as anger or sadness. Additionally, the demeanor of the witness can prompt corresponding emotional changes in jurors, relatively automatically and outside their awareness. Trials are powerful events, full of emotion, and this charged atmosphere has great potential to elicit emotions in jurors.

In the third section of the chapter, we discuss the psycholegal research that has examined the effect jurors' emotions have on the decisions they make during trial. Specifically, we discuss research regarding how juror anger, disgust, sadness, and other negative emotions impact juror decisions. Certain emotions have little impact on jurors' decisions, whereas others have a more profound impact. As will become clear, the next direction in the study of emotions and juror decision making will be to better understand these differential effects, and successfully link research on jurors with existing theories on emotions and judgment.

Last, we examine potential moderators that might alter the relation between emotions and decisions. Our findings that certain emotions affect decisions are important, but they leave out other important considerations. For example, individuals may vary in their sensitivity to certain emotions. More sensitive individuals might be impacted more by emotional evidence than those who are less sensitive. Further, actual trials are complex and may include circumstances that can mitigate or enhance emotional effects. For example, juror instructions are often used to caution jurors about their biases, and these instructions could moderate the impact of emotions. Additionally, jury deliberations could alter the effects of emotions on decisions, as jurors might caution each other about using emotions to make decisions, or during discussions they may compartmentalize emotions from their final decisions. There are also instances in which discussion might exacerbate the role of emotions on decisions (e.g., group polarization). We examine how juror instructions, individual differences in emotionality, and jury deliberations might minimize (or exacerbate) the impact of emotions on decisions. As will be seen, there is a paucity of research on these potential moderators, and what research has been done suggests that certain moderators (e.g., jury instructions) might be less influential than others (e.g., deliberations).

In each section of the chapter, we highlight what we know and what we have learned, but we also highlight what we do not know and make suggestions for future research. Our treatment of this research area is not exhaustive, but we attempt to highlight important accomplishments in the field, and encourage additional research where gaps exist.

How Emotions Influence Judgments of Blame and Responsibility: Theories and Empirical Evidence

Theories concerning how emotions may impact judgments have examined constructs such as affect, moods, and emotions (Forgas 1995; Schwartz and Clore 1983). In some cases, little distinction has been made between the three. Indeed, in some instances, we will see similar patterns of findings, regardless of which specific construct is identified. But, researchers generally tend to use affect as the more general category and subsume moods and emotions within it. The chief discrepancy between moods and emotions is that moods are generally considered to be more diffuse and longer lasting than emotions (Forgas 1995). In this chapter we will use the term emotions, but when an

important distinction needs to be made, we will point it out. We therefore begin with a description of some influential theories that guide the emotions and judgment literature in order to establish a number of predictions regarding how jurors may be impacted by these states.

Affect Infusion Model and Juror Decisions

Theories of emotions offer various explanations as to how emotions affect jurors' attention and processing of information presented as trial evidence. One prominent theory, known as the Affect Infusion Model (AIM; Forgas 1995), has specific predictions as to when emotions are more likely to influence the decision making process. For instance, emotions more likely affect judgments when one engages in heuristic or substantive information processing (compared to the alternatives of direct access and motivated processing). Heuristic processing is a strategy in which one needs to make some judgment and does so using the least amount of effort necessary. Heuristic processing is active when the situation does not require accuracy, motivation is low, the observer is under cognitive strain, and the target is considered simple. Substantive information processing, on the other hand, is a strategy in which the decision-maker must attend to, learn, and understand new information when making a judgment. This strategy most likely occurs when one is motivated toward accuracy, has the cognitive capacity to make a judgment, and the target is more complex. The similarity between these judgment strategies is that heuristic and substantive processing are not guided by preexisting patterns of searching for information. As a result, it is believed that emotions are more likely to affect judgments when one engages in heuristic or substantive information processing (Forgas 1995).

Under both heuristic and substantive processing, it is expected that judgments follow a mood-congruency effect, in which the judgment matches the mood of the individual. Emotions can act as information regarding how one feels about a particular stimulus (affect-as-information; Schwarz and Clore 1983) or can activate related cognitions that then affect the judgment in a congruent manner (affective priming; Bower 1981). Additionally, and specifically under substantive processing, the emotion experienced by the individual is also expected to aid in recalling mood-congruent information (Forgas 1995). This suggests that when jurors engage in either heuristic or substantive information processing, their judgments are more likely to "match" their mood and they will recall more information congruent with their current mood.

As Pennington and Hastie (1992) have noted, constructive processing is typical of jurors. Consequently, jurors likely adopt a substantive processing style because substantive processing occurs when motivation to process information is high and individuals are engaged in constructive processing tasks (Forgas 1992). According to the AIM, emotions should affect juror judgments through affective priming. It is also possible that under some trial circumstances, jurors might engage in heuristic

processing or some aspects of a trial might be processed heuristically. For example, jurors might process gruesome crime scene photos heuristically and become disgusted by what they see. In turn, the disgust could be directed toward the defendant, and make jurors more punitive. Thus, the AIM would predict that emotions would also be likely to impact juror decisions; however, research has not focused on identifying the processing styles of jurors or if some parts of trial are more likely to be processed heuristically.

Appraisal Theory and Information Processing

Other theories of emotion suggest that emotions can signal the use of a particular processing style, which subsequently affects decisions (Tiedens and Linton 2001). According to Tiedens and Linton's (2001) appraisal theory, specific emotions are associated with specific appraisal patterns and elicit feelings of certainty or uncertainty. Emotions related to certainty (e.g., anger, disgust, happiness, and contentment) might signal to the individual that his/her information is correct and he/she is right, ending any further search for information. Conversely, emotions related to uncertainty (e.g., hope, surprise, fear, worry, and, in some contexts, sadness) will signal to the decision-maker to persist in searching for information (Feigenson 2010; Tiedens and Linton 2001).

As such, certain emotions elicit heuristic information processing and other emotions elicit systematic information processing. Specifically, angry participants make more stereotypic judgments (e.g., believed a student athlete had cheated) and are more likely to rely on heuristic cues (e.g., agreed more with information provided by an expert versus someone perceived to be low in expertise) compared to sad or neutral participants (Bodenhausen et al. 1994). Fearful participants are more uncertain and, as a result, use more systematic processing strategies (Lerner and Keltner 2000; Parker and Isbell 2010), such as increased consideration of circumstances in ethical decisions (Kligyte et al. 2013). In addition to affecting information processing strategies, appraisal theories suggest that emotions can also impact attributions of blame or responsibility. To illustrate, Keltner et al. (1993) found that angry participants were more likely to believe individuals caused future events, whereas sad participants believed future events to be the result of situational factors.

In the context of juror decisions, unlike the AIM model discussed above, being exposed to emotion-provoking evidence (from witness testimony, gruesome photographs, or victim impact statements to name a few) may trigger a particular processing style (i.e., systematic or heuristic), which could affect attributions of guilt and other legal decisions.

Value and Application of Legal Information

While appraisal theories suggest how emotions can affect depth of processing information (i.e., systematic versus heuristic), Tetlock et al. (Goldberg et al. 1999; Tetlock 2002; Tetlock et al. 2007) describe more specifically how anger can affect the value of certain information. Similar models have been proposed by other researchers whereby emotions trigger a goal-directed decisional process (e.g., Alicke 2000; Haidt 2001). The Intuitive Prosecutor Model posits that anger elicits moral outrage and can shift a decision-maker from being an "intuitive scientist" to being an "intuitive prosecutor." In other words, a motivational shift occurs, in which one who previously sought an explanation for certain events now seeks to blame and punish (Goldberg et al. 1999). Generally speaking, emotions are believed to automatically trigger cognitive and behavioral responses towards some goal (e.g., fear triggers avoidance, while anger triggers retribution). The emotion then directs one's cognitive resources (e.g., attention, memory) to solving a particular goal (e.g., punishment and justice in the case of anger).

Although appraisal theorists might hypothesize that anger narrows attention and information processing strategies (i.e., increased use of heuristics), the Intuitive Prosecutor Model suggests that anger focuses attention and other cognitive resources toward relevant information to help achieve a punishment goal. As a result, angry individuals tend to have increased motivation to blame individuals and be more punitive in their judgments, among other consequences. To support this model, Goldberg et al. (1999) found that participants previously primed with an angry scenario were more punitive in subsequent judgments when they believed the individual in the previous transgression went unpunished compared to angry participants who believed the individual did receive some punishment, an effect that has been replicated in subsequent research (Tetlock et al. 2007).

Affect-as-Information Informing Juror Decisions

Finally, emotions can play a direct role in judgments and decision making. According to the affect-as-information hypothesis (Clore et al. 2001), the emotional response one has toward a target of judgment is a source of information for how one feels about that target. Due to the quick nature of emotions informing judgments, they are believed to have a large impact on decisions and can typically reach the observer faster than information about the target (Pham et al. 2001). As this process is described in automatic terms, emotions then act directly on judgments and decision making in a heuristic fashion. According to this hypothesis, the direct effects emotions have on cognition are described in positive or negative terms. When making a judgment, positive emotions inform the observer that the target has some positive value, while negative emotions signal that the target has no value, and so positive and negative evaluations are assigned, respectively (Clore and Huntsinger 2007).

One main tenet of the affect-as-information hypothesis is that the observer perceives that the target elicited the felt emotion (i.e., the emotion is integral rather than incidental). Once one recognizes that there is another source for his/her emotional experience, the judgment will no longer be affected by the emotion. Additionally, emotions will not affect judgments if the observer believes that the emotional information is irrelevant for the current evaluation/judgment (Schwarz and Clore 1983). For example, in their classic study, Schwarz and Clore (1983) found that participants reported more positive assessments of their life satisfaction on a sunny day but more negative assessments on a rainy day (this effect, however, disappeared when the weather was brought to their attention).

Summary

Though by no means exhaustive, the theories reviewed here describe when emotions are most likely to affect juror decisions (AIM; Forgas 1995) or when emotions elicit particular information processing styles (appraisal theory; Tiedens and Linton 2001). Emotions can also affect recall of information, as well as the weight we assign a particular piece of information (Bower 1981; Tetlock et al. 2007). Finally, emotions can have a more direct impact on evaluative judgments by acting as information regarding the target of the judgment (Schwarz and Clore 1983). Though many of these theories have not been formally tested in the context of juror decisions, the research summarized in the sections to follow generally supports the theories described in the present section.

Trial Evidence Eliciting Emotions in Jurors

The trial has been characterized as a melodrama, in which the story of the case is "simplified, personalized…[and] moralized (Feigenson 2000, p. 155)." Numerous elements of the trial generate strong emotional responses in jurors. In some cases the emotions are the result of intentional efforts to arouse these reactions, and in other instances emotions arise naturally and unavoidably from the evidence. The use of emotion to persuade individuals is an area long studied in the fields of applied social psychology, in business schools, and in advertising (Taute et al. 2011). And for good reason; persuasive appeals tend to produce greater change in attitude when the argument is congruent with the affective state of the target audience (Maio and Haddock 2007). Studies have demonstrated that eliciting emotions such as sympathy leads to increased prosocial behavior (Bagozzi and Moore 1994; Batson et al. 1997; Kemp et al. 2013).

Throughout the course of a trial, there are countless instances when strong emotions such as fear, sadness, disgust, and anger may be experienced. Trials are often dramatic, as individuals' lives and welfare hang in the balance, and acts of

great cruelty are frequently depicted in detail. Attorneys realize the opportunity to move jurors emotionally in order to better persuade them toward a particular decision. In many instances, the emotional nature of the evidence at trial may not represent a deliberate appeal on the part of the attorneys, but rather, come as a result of particular aspects of the evidence. The heinous and depraved nature of the crime or exposure to gruesome photographs of the crime victims can serve to generate strong emotional responses in jurors.

Graphic Photographs

As the recent case of the Boston Marathon bomber illustrates, jurors may sometimes be overwhelmed by witnessing highly graphic crime scene photographs, as the brutality of the crime and the suffering of the victim become apparent. Indeed, images such as the beating of Rodney King remained burned in the public's psyche more than two decades later. As Feigenson and Spiesel (2009) note, the capacity to capture events on video has increased with technological advances, and so this type of evidence may occur with greater frequency in the future. A number of studies have investigated the effects of graphic photographs on juror judgments. In two early studies, participants made larger damage awards after witnessing gruesome crime scene photos as compared to participants who did not (Oliver and Griffitt 1976; Whalen and Blanchard 1982). In a later study, the presence of a graphic crime video led mock jurors to lower their criterion for assigning a guilty verdict as compared to mock jurors who were not exposed to the crime video (Kassin and Garfield 1991). While the authors in each instance noted the role strong emotions may have played in the outcomes, Bright and Goodman-Delahunty (2004) point out that none of these studies directly assessed the emotional response of jurors. However, Douglas et al. (1997) varied whether the crime scene photograph was in color or black and white and found that color photographs led to higher emotionality ratings (i.e., increased anxiety, anguish, disturbance, shock) and more guilty verdicts. By contrast, in a later unpublished dissertation, Nemeth (2002) varied the gruesomeness of both the photographs and the verbal testimony. Emotions, as measured by the Profile of Mood States (POMS; McNair et al. 1981), did not predict verdict decisions for any of the variations in the photographs.

Bright and Goodman-Delahunty (2006) presented mock jurors with 20-page trial summaries and photographs that were either color or black and white. Additionally, the pictures were either neutral (e.g., close-up photo of a necklace) or gruesome (e.g., postmortem photos of the victim depicting deep neck wounds). Participants completed the Juror Negative Affect Scale (JUNAS) which assessed fear/anxiety, sadness, anger, and disgust, and pre–post difference scores were computed on this measure. Type of photographic evidence significantly affected the difference scores on emotion, as the total JUNAS difference scores for the gruesome photograph group differed significantly from the no photograph group. Moreover, anger at the defendant (but not disgust) mediated the relationship between gruesome

photographs and the degree to which participants perceived the prosecution evidence as sufficient for conviction.

In a more recent study, Bright and Goodman-Delahunty (2011) asked participants to read a 15-page transcript of a negligence case in which a pedestrian had been struck and injured by an automobile. The researchers crossed the severity of the injury with the gruesomeness of the crime scene photographs. Participants rated the emotional impact significantly higher when they were given gruesome photographs than when given neutral photographs. Multivariate effects for ratings on a combination of fear, anger, sadness, and disgust were significant for the gruesomeness of the photographs, such that participants showed higher levels of each emotion compared to those who saw no/neutral photographs.

To summarize, the images of corpses and graphic images of suffering depicted in trial photographs have the potential to take an emotional toll on jurors. Mock jury research has shown that this visual information has the potential to elicit emotional responses in participants such as anxiety, disgust, and anger. One reason why these images may impact jurors is that they communicate the brutality of the act and the suffering of the victim. However, both the brutality of the crime and the suffering of the victim can also be communicated to the jury by the graphic nature by which the crime is described to the jury, and so this area merits further description.

Graphic Verbal Descriptions and Crime Heinousness

Particularly brutal acts may arouse a variety of emotions in jurors: from disgust and/or anger directed at the defendant, to sadness for the suffering of the victim. Consequently, researchers have studied how crime heinousness can impact legal judgments. But, as Bornstein and Nemeth (1999) have noted, researchers have failed to define heinousness consistently. Sometimes the brutality of the act itself has been varied, sometimes graphic descriptions of the state of the body have been provided, and in some instances the innocence of the victim has been highlighted to emphasize the depravity of the crime. For example, Hester and Smith (1973) varied whether the victim of a gang shooting was a rival gang member (nonheinous) or a 10-year-old girl (heinous). Jurors judged the killer in the heinous condition as more dangerous. Hendrick and Shaffer (1975) varied heinousness by manipulating the degree to which the corpse was mutilated. The more heinous the crime, the more approving participants were toward applying the death penalty. Similar findings related to the brutality of the act and the subsequent punishment were observed by Konecni et al. (1996).

Yet, the degree to which emotions were aroused by these acts was not addressed in these previous studies, thus making it unclear what emotional toll resulted from these crimes. However, one study in particular did find that mock jurors showed an increase in stress when asked to read a graphic description of violent evidence (Thompson and Dennison 2004). More recently, Myers et al. (2013) varied the brutality of the crime and measured emotional responses along with sentencing judgments. In the nonheinous condition, the victim died of a single bullet wound to

the heart. In the heinous condition, the victim suffered for a period of time as his kneecaps and groin were shot, and he was pistol whipped repeatedly in his face and head. Though heinousness impacted sentencing, it had a rather weak effect that only approached significance on anger ratings.

In many instances, the detail and graphic nature of the act and the suffering of the victim has been manipulated to determine how jurors might respond. Bright and Goodman-Delahunty (2006) have proposed that when strong emotions such as anger are aroused by graphic visual evidence, the emotion will prompt jurors to alter their judgments concerning each individual piece of evidence presented at trial. This model is consistent with the prosecutorial mindset argued by the Intuitive Prosecutor Model (Goldberg et al. 1999). The researchers further postulate that graphic content may lead jurors to misjudge the probability that the defendant committed the crime (i.e., probability of commission) or lower their threshold for conviction (i.e., reasonable doubt; see Kassin and Garfield 1991).

Yet, findings regarding the effects of vivid descriptions of suffering, brutality or heinousness of the crime have been rather inconsistent. For example, Nemeth (2002) did not find evidence of significant effects when both the verbal descriptions and crime scene photos were varied. Similarly, null findings were also reported by Bright and Goodman-Delahunty (2006) regarding graphic verbal descriptions. Here, the researchers introduced gruesome and detailed descriptions of the wounds to the victim's neck. Contrary to predictions, this manipulation failed to produce differences in the emotional responses of mock jurors or in their ratings of defendant culpability. By contrast, Thompson and Dennison (2004) obtained significant effects by varying the graphic language (e.g., "cut" versus "slashed") associated with the crime that was inserted throughout the lengthy (i.e., 40-page) summary. Graphic details led to significantly greater emotional responses (e.g., stressed, anguished, disturbed, shocked) compared to the nongraphic control group.

Consequently, we see that graphic visual images appear to have a stronger and more robust effect on jurors' emotional responses. By contrast, the effects of graphic verbal descriptions or heinous and brutal crime information are less consistently impactful or less directly associated with emotional changes in jurors. Of course, factors other than the crime, such as the demeanor of the witness or defendant, may also influence legal judgments throughout the course of the trial.

Witness Demeanor

When a witness testifies, jurors pay attention to both the content of the message as well as how the message is delivered. One important aspect of the message delivery is the emotional expression the witness presents. The emotions of others may sometimes be transferred to the audience through a process known as emotional contagion. Additionally, jurors may focus on the affective displays of witnesses in order to assess the credibility of the testimony. Below we discuss the effects of a witness' demeanor on jurors and their decisions.

How emotional displays may be transferred. The testimony of a highly distraught victim who communicates how he/she has suffered as a result of the defendant's acts is emotionally compelling and capable of generating emotions such as sympathy toward the victim or anger toward the defendant. In addition, emotions may also be elicited as a result of observing an emotional witness. These emotional displays from the witness can transfer similar emotional states to the jury. This transference is both automatic and immediate, thus leaving little opportunity for jurors to establish effective defenses to prevent it. We refer specifically to the process of emotional contagion.

Goldie (1999) refers to emotional contagion as a form of emotional engagement that does not require that one understand the emotions of another. Hatfield et al. (e.g., Hatfield et al. 1993, 1994) describe emotional contagion as an automatic process whereby people mimic and synchronize their facial expressions, postures, movements, and even vocalizations with others. Therefore, the sad expression, morose tone, and slumped posture of person A may elicit synchronized related behaviors in person B. As a consequence of person B adopting this behavior, he or she may then begin to experience the sad emotion felt by person A through an afferent feedback loop, as the brain interprets the tone, posture, and facial expression as evidence that they must be feeling sad. According to the facial feedback hypothesis, manipulating our facial expressions leads to differences in our affective states (Strack et al. 1988). Manipulated expressions of a frown cause us to be unhappy, whereas a smile produces happiness (Laird and Bresler 1992). Our brain appears to be hardwired for the tendency to mirror the expressions and behaviors of others (Rizzolatti et al. 2002). The concepts of emotional contagion and facial feedback are not new, as they can be traced to Darwin (1872) and James' (1922) beliefs that facial expressions are not just the product of emotions, but can also affect the emotional experience. Yet, much of the research on this phenomenon is recent. Researchers have proposed similar concepts (e.g., emotion contagion; Sullins 1991; mood contagion; Neuman and Strack 2000; depressive drift; Coyne 1976). Regardless of the term preferred, emotional contagion is supported by a number of studies which demonstrate that we adopt the emotions of others, and that this process can occur relatively automatically (Neumann and Strack 2000).

Emotional demeanor and juror reactions. Jurors may glean critical information about a witness or defendant based on his/her emotional demeanor. Tsoudis and Smith-Lovin (1998) argue that emotional displays by the actor signal to the perceiver the identity of an actor who is in a transient role, such as offering testimony. In support of this, Rose et al. (2006) found that victims are expected to display affect that is consistent with the seriousness of the event they experienced. Rape victims who exhibit strong emotions and signs of distress are judged as more credible than victims who are calm or show muted affect (Baldry 1996). Salekin et al. (1995) found that participants judged a female defendant (but not a male defendant) displaying flat affect to be guilty more frequently than when a moderate amount of affect was displayed. Hackett et al. (2008) similarly found that emotional displays that were in line with expectations led to greater credibility ratings

regarding rape victims. As Salekin et al. (1995) note, it is when displayed emotional expressions violate expectations that jurors may seek explanations for the actor's expressions. Consequently, in instances of severe harm and suffering, the emotionality of a particular witness would be expected to be high, and in such instances a lack of emotional display would be more likely to provoke a response from jurors. Importantly, jurors may be attuned to these displays and make attributions about the credibility of the witness (e.g., Dahl et al. 2007; Kauffman et al. 2003) in addition to the potential for their own emotions to shift as a result.

Defendant Emotion

Another source of emotion at trial that can alter jurors' decisions is the defendant. Peace and Valois (2014) manipulated the emotionality of a defendant in the context of a "he said, she said" sexual assault case to be either low or high through his testimony. The judged appropriateness of the emotional display may depend on the characteristics of the jurors. Results indicated that when the defendant was less emotional during testimony jurors who were experiential processors (i.e., affect oriented, quick, and preconscious processors) were more punitive than when the defendant was more emotional. However, the pattern reversed when jurors were rational processors (i.e., reason oriented, effortful, and conscious processors; Peace and Valois 2014). Here, the more emotional the defendant, the more punitive the juror. Salekin et al. (1995), mentioned earlier, altered the defendant's testimony to be either flat affect (no emotional display), moderate affect (some crying and other emotional displays), or high affect (intense emotional displays). When the defendant was female and displayed either flat affect or high affect, she was seen as significantly more guilty compared to when she displayed moderate affect, whereas perceptions of guilt did not differ significantly across levels of affect if the defendant was male (Salekin et al. 1995).

Some researchers suggest that the emotionality of the defendant alters jurors' verdicts through changes in their perceptions of the defendant (Heath and Grannemann 2014; Heath et al. 2004; Robinson et al. 1994). For example, Heath et al. (2004) used a trial vignette in which the defendant was charged with the murder of her husband. The defendant's testimony was manipulated based on intensity of emotions to be either low affect (no emotion), moderate affect (emotions indicative of sadness), or high affect (visible tears, a lot of crying/emotions). The defendant was judged to be more guilty when she displayed low levels of emotion compared to when she displayed moderate levels of emotion. In a later study, Heath and Grannemann (2014) found no effect of defendant emotion on verdict, but it did affect perceived credibility of the defendant. Jurors perceived the defendant who showed little emotion as less credible, similar to Robinson et al. (1994) findings in which defendant emotion only had an effect on jurors' perceptions of the credibility of the defendant. These studies imply that the effect of defendant emotion on jurors' verdicts is perhaps more nuanced than initially thought, such that a defendant must display the "proper" amount of emotion in order to receive less harsh judgments from jurors.

Overall, we see that jurors are attuned to the emotional displays of the various actors (e.g., witnesses, defendant) in the trial, and the emotional demeanor of these individuals is easily transferred to the jurors. But, the demeanor of a witness may also be used as a guide by which other important judgments (e.g., witness credibility) are made. In other instances, who testifies, and the information they testify about, often form the basis of a direct appeal to juror emotions. In these instances, the testimony may be introduced with the goal of moving jurors emotionally.

Emotional Appeals and Jurors

One approach attorneys may take is to attempt to elicit sympathy or empathy from jurors. Although the research on this is sparse, some studies have tested the effectiveness of these direct appeals. For example, Haegerich and Bottoms (2000) varied whether the attorney directed jurors to empathize with the defendant. Specifically, during the course of reading a 20-page modified transcript based on a patricide case, jurors were encouraged to "reflect on how you would feel if you were in these circumstances" (p. 427). Their six-item measure of empathy revealed that jurors given the empathy instruction exhibited more empathy for the defendant than those who were not presented with empathy appeals.

In other instances, the testimony in question may have little probative value, and its inclusion in the course of a trial may be largely to stir the emotions of the jury. For example, Victim Impact Statements (VIS) are statements given by members of the victim's family and close friends, and most often occur during the sentencing phase of capital murder trials. VIS explain the emotional, financial, psychological, and physical harm that has occurred due to the loss of the victim, and as such, are often inherently emotional, with questionable relevance to the sentencing task with which jurors are confronted (see *Booth v. Maryland* 1987 and *Payne v. Tennessee* 1991). In some instances, these statements are accompanied by photographs, videos, and music.[1] The instances when these emotional statements bring out corresponding anger or sadness in both the jury as well as the judge are legion (see Logan 1999, for a description of many of these accounts).

A number of studies have examined the degree to which VIS may lead to emotional responses in jurors. Greene, Koehring, and Quiat (1998) varied the respectability of the victim portrayed in a VIS and examined the judgments of a

[1]In support of this, Feigenson (2010) notes the proliferation of victim impact videos. Here a slide show of the victim's life, in some instances accompanied by emotionally evocative music, is presented to jurors—in both criminal and civil cases but most often in wrongful death cases. To date, the effects of these emotionally stirring videos on juror decision making have not been empirically tested. However, with the U.S. Supreme Court decision in *Kelly v. California* (2008), which upheld the constitutionality of these videos in capital sentencing proceedings, victim impact videos remain admissible in court. As such, the frequency with which they are used during trial may ultimately increase.

sample of jury-eligible community members. Mock jurors expressed significantly more compassion for the survivors when the victim was high on respectability than when the victim was low on respectability. Similarly, Butler (2008) presented 200 jury-eligible and death-qualified community members a brief summary of a capital case that varied the presence of an emotional victim impact statement. The presence of the victim impact statement led to significantly greater empathy toward the victim as measured by a composite of ratings of likability, compassion, similarity, and perceived suffering of the victim. Boppre and Miller (2014) varied the presence of VIS along with the presence of Execution Impact Evidence (statements made, usually by the defendant's family about how the defendant's execution would impact them). While it did not influence verdicts, individuals did report greater negative affect when they were presented with VIS than when it was absent.

At times, witnesses who deliver the VIS may become visibly distraught, and their emotional demeanor may have the potential to elicit emotional responses in jurors. To examine this, Myers et al. (2002) varied the level of harm expressed in a VIS and crossed that variable with the emotional demeanor of the witness (emotional/stoic) in a videotaped trial simulation. Mock jurors rated their emotions after witnessing the VIS using the PANAS (Watson et al. 1988). Both the severe harm condition and the emotional witness demeanor condition led to significantly greater scores on negative affect. However, changes in negative affectivity in this instance were relatively mild, and subsequent sentencing judgments were unrelated to jurors' emotional responses to the testimony.

In other investigations, the demeanor of the victim testifying was not specifically manipulated, but rather, the degree of emotionality expressed in the statement was varied. Nadler and Rose (2003) varied both level of harm and the level of emotional suffering experienced by the victim. They randomly assigned participants to read a brief vignette that either did not include a VIS, or included a VIS in which the victim expressed either mild emotional injury or severe emotional injury. In the severe emotion VIS, the victim talked about how she now feels vulnerable, afraid, depressed, cannot sleep, and cannot stop thinking about the crime. In the mild emotion VIS, the victim reported being angry initially, but has since been able to return to her normal life. After reading the vignette, participants were asked to indicate what sentence the defendant should receive on a scale from that ranged from "no prison" to "18 years or more." Regardless of crime (robbery or burglary), participants were significantly more likely to recommend a longer sentence when they heard the severe emotion VIS than the mild emotion VIS. And, although the conditions failed to produce differences in participants' ratings of disgust or anger, the severe emotional injury condition led to significantly greater sympathy ratings than the mild emotional injury condition. Consequently, as victims express their emotional suffering, jurors may be more likely to experience feelings of sympathy that are aligned with these expressions of suffering on the part of the witness.

In order to assess whether jurors are sensitive to who expresses the VIS, McGowan and Myers (2004) held the content of the VIS constant and varied the identity of the individual who gave the VIS. Community members read a brief summary of a capital trial in which the victim died in an explosion at work set off

by a disgruntled coworker. They varied the identity of who gave the VIS (a spouse, a coworker, or an emergency first responder) and found that participants did not significantly differ across the three identities in their ratings of how emotionally moving they found the testimony. However, their ratings of how emotionally moving the testimony was did significantly correlate with how much they perceived the witness to suffer as a result of the crime.

Researchers have more recently focused on the specific emotions elicited from witnessing a VIS. Paternoster and Deise (2011) used more ecologically valid stimuli than had been used in previous studies; where jury-eligible and death-qualified community members watched a videotaped trial which included or did not include an actual VIS. The case involved the killing of a police officer, and the VIS included a description of how loving the victim was and how the victim's daughter still missed her father. Individuals who watched the actual VIS were significantly more likely to describe their emotions as "angry" and "vengeful" and were more likely to endorse feelings of "upset" and "angry" than were individuals who did not watch a VIS. Additionally, ratings of sympathy, as well as empathy, were significantly higher when they witnessed a VIS. Importantly, these emotional responses partially mediated their sentencing judgments.

Most recently, Nunez et al. (2015b) manipulated the emotional demeanor of the witness during the delivery of a VIS, and examined how it impacted juror sentencing decisions. They specifically examined two types of negative affect (anger and sadness). As part of the sentencing phase of a capital murder trial, the victim's spouse delivered the identical VIS in either an angry or sad demeanor. Results indicated that participants who heard the angry VIS were significantly more likely to sentence the defendant to death than participants who heard the sad VIS, suggesting that the type of emotion expressed in a VIS can also alter jurors' sentencing decisions.

Overall, we see that the introduction of a VIS has tended to elicit emotional responses in mock jurors. These effects are relatively robust, despite the fact that in many instances jurors are only reading the reactions of the witnesses rather than watching an actual delivery of the VIS. We see stronger emotional effects in cases in which the VIS was real and jurors had the opportunity to see the witness speak (e.g., Paternoster and Deise 2011). One is left to consider how much more impactful these statements would be if jurors were exposed to an actual trial in which they can witness the deceased victim's family in the courtroom throughout the trial.

Psychology and Law Studies of Emotions and Jurors' Decisions

The previous section illustrated how and when emotions are aroused during trial. These emotions can vary in type and intensity and can be experienced by a variety of actors at trial. For example, jurors may experience disgust or sadness after certain evidence is presented, victims may become sad or angry when delivering their

testimony, or the defendant may express sadness or indifference while listening to the testimony. This section presents the current empirical literature examining the various emotions jurors experience at trial and how these emotions alter jurors' decisions.

Juror Emotion

While the courts have tended to identify emotion in general as an undermining factor in reasoned judgments (Posner 1999), it is becoming increasingly clear that juror decision making is differentially influenced by discrete emotional states. Anger, disgust, fear, and sadness cannot be grouped together if one hopes to predict juror judgments of guilt or sentencing. Instead, these emotional states lead to unique patterns of legal judgments that suggest that some emotions are more likely the enemy of reason than are others. Below, we explore several emotions that have a differential impact on juror judgments.

Anger. Anger is arguably the most studied emotion, with much of the recent research finding that anger affects jurors' decisions. The effects of juror anger on subsequent decisions have been examined using a variety of case components, such as pretrial publicity (e.g., Ruva et al. 2011) and victim impact statements (e.g., Myers et al. 2002). In general, findings indicate that as jurors' anger increases, they become more punitive, as evidenced by more guilty verdicts and harsher sentences (e.g., Georges et al. 2013; Nunez et al. 2015a; Ruva et al. 2011).

Recent examinations of the effects of anger have utilized capital murder cases to determine how juror anger may alter sentences. For example, Nunez et al. (2015b) recruited death-qualified jurors to watch an approximately 45 minute reenactment of the sentencing phase of a capital murder trial. The trial consisted of attorneys' opening and closing arguments, testimony from a parole officer regarding aggravating evidence, testimony from a psychologist regarding mitigating evidence, and a VIS from the wife of the victim. Jurors reported their level of emotions both before and after viewing the trial using the PANAS-X (Watson and Clark 1994). Jurors experienced increased levels of anger and sadness, but not fear after watching the trial. However, only increases in juror anger predicted sentence. Specifically, jurors who reported greater changes in anger were significantly more likely to sentence the defendant to death (Nunez et al. 2015a). Consistent with the Intuitive Prosecutor Model described previously, angry participants weighed the prosecution's aggravating evidence as more important, which mediated the effects of anger on death sentences (Nunez et al. 2015a). Thus, as the model predicts, angry individuals with a goal of retribution and punishment are more likely to focus their attention on and weigh factors that are congruent with their goal.

Georges et al. (2013) found similar results, again using a capital murder trial. Jurors' emotions were measured at five different time points throughout the trial, and results indicated that juror anger increased significantly from baseline to the

second point in the trial and then decreased somewhat linearly during the trial. By the end of the trial, anger was lower than the second point in the trial but was still significantly higher than baseline. Jurors who experienced greater increases in anger throughout the trial found the defendant guilty significantly more than jurors who did not experience increases in anger (Georges et al. 2013). As previously mentioned, Paternoster and Deise (2011) used a capital murder context and they also found evidence of anger leading to increased likelihood of sentencing the defendant to death. Consequently, some consensus has emerged that when jurors experience increased anger during trial, more punitive judgments arise as a consequence.

Disgust. As previously mentioned, another emotion that can be aroused during trial is disgust (Bright and Goodman-Delahunty 2006). Although less studied in the legal realm than anger, disgust has also been found to have an effect on jurors' verdicts and sentencing decisions. Findings have been mixed, with some finding no effect of juror disgust on decisions and others finding that juror disgust leads to increased punitiveness. There is some more recent evidence (discussed below), however, that indicates disgust may work through anger to alter jurors' decisions.

To examine the impact of disgust on jurors' decisions, Capestany and Harris (2014) manipulated the level of disgust in a series of crime vignettes. Participants were placed in an fMRI and read 84 brief crime vignettes matched on sentencing recommendations that were either low or high in disgust. After each vignette, they rated on a 7-point Likert scale how much they would punish the defendant. Results indicated a significant main effect of disgust on sentence, such that when the vignette was more disgusting, participants were more likely to suggest a longer sentence. Counterintuitively, the fMRI results indicated that brain regions typically seen as responsible for logical reasoning were less active when the crime vignette was low in disgust, and more active when the crime vignette was high in disgust, implying that disgust can alter jurors' reasoning abilities. Although juror disgust was not directly measured, these results indicate that disgust does impact jurors' decisions. Further, consistent with appraisal theories, described earlier, disgust may shift people's processing styles when aroused.

Directly measuring juror disgust, Georges et al. (2013) found that jurors' reported levels of disgust increased throughout the trial and followed a similar pattern to anger. Additionally, although jurors' disgust increased, changes in disgust did not predict sentencing decisions, only anger did. Though disgust did not predict sentence, it is interesting to note that jurors' disgust and anger changed in concert throughout the trial, providing some evidence that disgust and anger are related.

Recently, researchers have begun to examine the complicated relation between disgust and anger elicited during trial and how these emotions may affect jurors' decisions. For example, Salerno and Peter-Hagene (2013) had participants watch a 20 minute murder trial in which disgust was elicited using graphic pictures of the murder victim. They measured jurors' emotions using a grid with disgust on one axis and anger on the other. Results indicated that disgust was a consistent predictor of moral outrage, and moral outrage mediated the effect of anger on guilty verdict confidence (Salerno and Peter-Hagene 2013). These results provide further support

that disgust and anger are related and further research is necessary to determine their differential effects on verdicts and sentencing decisions.

Recently, Schweitzer et al. (2015) examined the effects of disgust by manipulating the defendant's actions at the crime scene and examined how the disgusting behavior of the defendant might impact decisions. Participants were given a trial summary of a case in which the defendant was being charged with burglary. According to the vignette, surveillance video evidence showed that the defendant broke into the home but did not actually steal anything because the homeowner arrived home. In the no disgust condition, the defendant is seen petting the cat; in the contamination disgust condition, the defendant is seen sifting through the cat litter with his hands; and in the moral disgust condition, the defendant is seen using his finger to penetrate the anus of the cat. The results indicated that participants in the moral disgust condition were significantly more likely to find the defendant guilty of burglary. However, this effect was qualified by a significant double-mediation model, such that participants in the moral disgust condition, compared to participants in both the no and contamination disgust conditions, reported greater increases in disgust, which increased their anger, which then led to more guilty verdicts. These results were replicated in a second study using a different moral disgust manipulation (i.e., the defendant masturbated while in the laundry room with the cat; Schweitzer et al. 2015), providing further support for the notion that disgust alters jurors' verdicts.

Though the research on the effects of disgust on jurors' decisions is just beginning and does not allow for any concrete conclusions to be made, the present findings imply that disgust can alter jurors' decisions, although not always directly. Further, different types of disgust (e.g., moral or contamination) might affect jurors differently.

Fear and Sadness. Compared to anger and disgust, there is very little research examining how fear affects jurors' decisions. Two studies that were discussed earlier (Georges et al. 2013; Nunez et al. 2015a) measured fear before and after exposure to a trial. They found that certain negative emotions (e.g., anger) increased during the trial, but neither found an increase in fear. It is entirely possible that jurors do not feel fearful during trial. The defendant is safely ensconced in the courtroom, and there are officers of the court that ensure safety. Jurors might experience fear in certain types of trials where their own safety might be at risk (e.g., trial involving an organized crime) but it is probably rare. It is also possible that fear is a difficult emotion to arouse in mock jury studies. Participants know they are mock jurors that are participating in an experiment. Though they may experience various emotions during trial, they may not feel fearful. Real jurors, on the other hand, might experience fear during trial, though we know of no study that has attempted to measure fear in actual jurors. It may be that questions about the relationship between fear and legal judgments might be best tested in field studies rather than lab experiments.

The research on sadness is limited and conflicting. It has been hypothesized that sadness leads to a more systematic processing style (Carver and Harmon-Jones

2009), which may lead to more careful examination of trial evidence and potentially less punitive decisions, but this is not always found. For example, Nunez et al. (2015a) measured jurors' emotions and found that although jurors' levels of sadness increased during the sentencing phase of a capital murder trial, their increased sadness did not predict verdict decisions. Also, Semmler and Brewer (2002) directly manipulated sadness by altering the testimony of one of two witnesses during a murder trial. Results indicated that, consistent with appraisal theories, sadness was associated with systematic processing of witness testimony, as sad mock jurors were more accurate in finding inconsistencies in testimony compared to neutral participants (supporting the idea that emotions can affect information processing styles, described earlier). However, results showed no main effect of mood on verdict; mock jurors who received the sad testimony were no more likely to give a guilty verdict than participants who received the neutral testimony (Semmler and Brewer 2002). In contrast, one study that examined a variety of negative emotions found that jurors who reported an increase of sadness (along with vengefulness, outrage, shock, and anxiousness) after reading about the trial gave significantly more guilty verdicts than those who did not (Douglas et al. 1997). The different results of these studies regarding the impact of sadness on jurors' verdicts do not allow for a concise conclusion to be made, and certainly suggest that more research is needed.

Other negative emotions. Other studies have examined the effects of negative emotions in general, without teasing apart the different possible negative emotions. For example, affect-as-information effects were investigated in a civil trial context in which the amount of negative consequences incurred by the plaintiff was varied. Results showed that those who were given information about the plaintiff's negative consequences were more motivated to blame and find the defendant liable (Kadous 2001). However, when these jurors were given an attribution for their negative feelings (i.e., told that jurors can experience anxiety when making these decisions, and they might also have the same experience), these effects disappeared. Although emotion was not directly manipulated, Kadous (2001) inferred from the blame motivation ratings that affect acted as information in these mock jurors' decisions.

Further, Fishfader, Howells, Katz, and Teresi (1996) used a wrongful death suit to examine if mood changes influenced liability assessments and recall of trial information (testing for mood-congruency effects on memory, described earlier). Participants rated their mood before and after the case summary, which included testimony from an eyewitness and expert witness regarding the drowning death of a 12-year-old girl whose hair became stuck in the suction of a pool. Participants who experienced a greater change in negative affect placed significantly less liability on the plaintiff (implying more liability was placed on the defendant). However, there was no significant relationship between change in mood and actual damage awards or recall of information (Fishfader et al. 1996). When examining the effect of a juror's negative affect on criminal trial decisions, Myers et al. (2002) found that jurors' negative affect did not significantly influence their sentencing decisions in a

capital murder trial. Overall, the differential effects of emotions such as sadness and anger suggest that there is merit to investigating a number of other discrete emotions that have, up to this point, been largely ignored in the context of jury decision making.

Summary

The results of the reviewed studies suggest that not all negative affect experienced by jurors leads to more capricious judgments. Specifically, sadness and general negative affect do not consistently predict punitiveness. In fact, the majority of the studies find no effect of sadness or general negative affect on mock jurors' decisions (e.g., Boppre and Miller 2014; Myers et al. 2002; Platania and Berman 2006); however, sad mock jurors were better at detecting inconsistencies in testimony (Semmler and Brewer 2002), suggesting that sadness leads to systematic information processing and perhaps more positive effects on information evaluation. There is more consistent evidence, however, for disgust and anger. Jurors who report experiencing more anger are more punitive than jurors who report less anger (Georges et al. 2013; Nunez et al. 2015a; Paternoster and Deise 2011). Disgust also seems to increase punitiveness, however, it may do so through indirect routes, such as anger or moral outrage (e.g., Salerno and Peter-Hagene 2013; Schweitzer et al. 2015). These studies demonstrate that judges and attorneys do not need to be concerned about the effects of all negative affect experienced by jurors (e.g., sadness), but should use caution when introducing evidence that may increase juror anger or disgust. However, a more extensive examination of sadness is warranted to ensure this particular emotion does not affect juror decisions.

Potential Moderators

In the previous sections, we have reviewed literature that demonstrates which aspects of trial arouse emotions and how emotions affect jurors' decisions. In this section, we propose several potential moderators that should be considered in future research. In particular we examine how jury instructions, individual differences in emotionality, and deliberations might moderate known effects of emotions on decision making.

Juror Instructions

If jurors' emotions are aroused at trial, one possible remedy or moderator might be instructions that caution jurors not to let their emotions interfere with their

decisions. As is true of most areas of psychology, the evidence for the benefits of jury instructions are mixed. In some cases, instructions help, while in others they may make no difference, and in still other instances, they may actually magnify the impact of emotions on decisions.

Limiting instructions. One type of instruction that jurors may receive at trial is known as limiting or cautionary instructions. These instructions often are given when jurors hear or see evidence that might have a prejudicial impact on their decisions. As opposed to typical juror instructions, which are presented at the end of trial, limiting instructions are often given when prejudicial evidence is presented (Cush and Delahunty 2006). One type of limiting instruction is presented when jurors are exposed to gruesome evidence at trial. Cush and Delahunty (2006) exposed mock jurors to gruesome crime scene photographs and tested whether the presence and timing of the instructions would impact decisions. They found that when limiting instructions were provided immediately prior to the presentation of gruesome evidence, jurors were less likely to convict the defendant compared to jurors who received the instructions at the end of trial or who received no instructions. Importantly, jurors who received instructions prior to seeing the gruesome evidence were no more likely to convict the defendant than were mock jurors who did not see the photographs.

Limiting instructions are also provided when evidence presented is challenged and ruled inadmissible. Typically, jurors are given instructions to disregard the inadmissible evidence, though the effectiveness of such instructions is contested.[2] Consistent with the old adage that "a bell, once rung, cannot be unrung," this procedural safeguard may be ineffective for a number of reasons. Generally, instructions to disregard evidence may fail to produce the intended effects because, paradoxically, efforts to suppress thoughts often incur the opposite effect, largely because thought suppression also requires thought monitoring (Wegner 1994). There are also motivational reasons why jurors might be inclined to disregard these decisional restrictions (e.g., reactance, see Brehm and Brehm 1981). In studies examining motions to strike and instructions to disregard evidence, participants are initially presented with information that is probative, but they later learn cannot be used, often for procedural reasons (e.g., hearsay evidence, exclusionary rule). Although the information may be useful for rendering a judgment, it is not information likely to generate strong emotions in jurors. Therefore, the issue of concern here is whether these instructions to disregard have any impact on judgments when the information elicits strong emotions in jurors.

Edwards and Bryan (1997) conducted a study that varied the emotionality of a crucial piece of evidence that was either ruled inadmissible or admissible. When evidence was ruled inadmissible, mock jurors were told to disregard the evidence and not let it affect their verdict decisions. In a series of two studies, the researchers

[2]There is ample evidence that jurors find it difficult to disregard inadmissible evidence (Steblay et al. 2006). But see our extended discussion of this problem in our discussion of jury deliberations at the end of this section.

found that jurors in the inadmissible conditions who were exposed to emotional evidence were not only unlikely to disregard the evidence, they were more likely than any other group (including jurors exposed to nonemotional inadmissible evidence) to find the defendant guilty (Edwards and Bryan 1997). In other words, instructions to disregard evidence exacerbated the influence of the emotional information on decisions.

Similar findings have been reported by researchers who have examined the utility of limiting instructions on jurors who are exposed to pretrial publicity (PTP). PTP has been the subject of much research. The problems caused by PTP have been demonstrated in various studies (e.g., Daftary-Kapur et al. 2014; Ruva and McEvoy 2008), and PTP was the subject of a 1999 meta-analysis (Steblay et al. 1999). Studies typically find that jurors who have been exposed to negative PTP are more likely to find the defendant guilty than are jurors exposed to neutral or no PTP (Steblay et al. 1999). Several researchers have distinguished between factual and emotional PTP (Kerr 2010; Kramer et al. 1990; Ogloff and Vidmar 1994; Wilson and Bornstein 1998). Factual PTP is information about the defendant that is incriminating (e.g., a prior conviction for the same crime or that defendant confessed when arrested). Emotional PTP arouses negative feelings about the defendant but is not incriminating, per se (e.g., lurid descriptions of a victim's injury or negative information about the defendant's lifestyle). In one study, researchers varied the PTP (emotional or factual) and examined whether juror instructions could ameliorate the effects of PTP on decisions (Kramer et al. 1990). They exposed jurors to either factual PTP (the defendant had an extensive prior criminal record and a gun used in the robbery was found at a girlfriend's house), emotional PTP (the defendant hit a young girl a few hours after the robbery and fled the scene, and an interview with the grieving mother after learning her child had died), or no PTP. Half of the mock jurors were instructed not to use any PTP or their reactions to the publicity in their judgments of the case. The researchers found that post-deliberation jurors were the most likely to find the defendant guilty when they were exposed to the emotional PTP (though factual PTP also yielded more guilty verdicts than the no PTP condition), and importantly, juror instructions had no impact on decisions. Others have used juror instructions to attempt to minimize the effects of PTP and have yielded similar nonsignificant results (e.g., Fein et al. 1997). Instructions do not appear to moderate the effects of PTP; this might be particularly true for PTP that arouses negative emotions in potential jurors.

Jury nullification instructions. Jury nullification is the process in which jurors disregard the facts in a case to render a verdict that is consistent with their own sense of justice. For example, a nurse being tried for murder might be found not guilty if his/her actions were at the request of a terminally ill patient who wanted help to die. Although jurors have the power to nullify, courts have been very reluctant to instruct jurors that they have such power. The most notable example is the case of *United States v. Dougherty* (1972), in which the United States Court of Appeals for the Third Circuit ruled that defendants (nine Catholic clergy) found guilty of vandalizing offices at Dow Chemical Company (as a form of protest) were

not entitled to a new trial because the judge refused to allow nullification instructions.

By definition, cases that invite the desire to nullify arouse emotions in jurors. Nullification instructions typically inform mock jurors that their feelings based on their own conscience or sense of justice can be considered when they reach their verdict. Horowitz et al. (2006) have argued that the major concern of the justices in *Dougherty* was that nullification instructions would invite jurors to use emotions rather than evidence to make their decisions, and this would invite "chaos" into the courtroom. Judicial instructions that raise the specter of nullification might encourage jurors to allow their hearts (i.e., emotions) to rule their heads (i.e., reason).

Horowitz et al. (2006) proposed that nullification instructions combined with emotionally charged trial elements might magnify jurors' reliance on emotions in their decision making. They tested this by manipulating characteristics of the victim (either sympathetic: grandfather and generous philanthropist, or nonsympathetic: a mobster with his fortune stored in offshore accounts) to create an emotional bias for or against the victim. They also varied the nature of the case (murder to get access to a victim's fortune, or euthanasia to relieve the suffering of a terminally ill patient). Half of the participants received nullification instructions and half did not. Nullification instructions led to an overall decrease in guilty verdicts. Importantly, mock jurors who received nullification instructions and were exposed to unsympathetic victims were significantly less likely to convict the defendant. Horowitz et al. (2006) suggested that nullification instructions resulted in a misattribution of emotional responses in the nullification relevant trial. Jurors are told they can use their conscience or sense of justice but are not warned that the perceived source of their emotions may be confused. In this study, mock jurors felt negative emotions toward the unsympathetic victim and when given permission to use their feelings to make their decisions, they were less likely to convict the defendant. In a similar study, Kerr et al. (2008) tested standard and modified nullification instructions (instructions that explicitly prohibited jurors from using emotions that they feel at trial in making their decisions). They found that the verdicts in the standard and modified nullification instructions were indistinguishable, and there was an amplification effect of instructions when the victim was unsympathetic. Thus, instructions to nullify amplify bias in cases that raise concerns about the fairness or justice of the law.

As this summary indicates, the use of jury instructions to limit the impact of emotional evidence appears to depend on the case circumstances, the timing of instructions, and what problem the instructions seek to address. However, there is a dearth of literature in this area, and there is much work to be done. As previously mentioned, if instructions could lessen the impact of emotions on jury decisions, this would be of great utility to our judicial system. Providing instructions to jurors would be a relatively easy and inexpensive judicial remedy. We are not naïve enough to believe that instructions will solve all of the problems in this area, but further research could identify the case circumstances that are remedied with instructions and the best timing for those instructions. We may also find that some

emotions (e.g., disgust) are more easily addressed with instructions than others (e.g., anger). In sum, knowing the utility and limitations of juror instructions on emotion should be an important goal of psycholegal scholars.

Individual Differences in Emotionality

There are a number of individual difference characteristics that could moderate the effects of emotional stimuli and subsequent judgments. For example, as mentioned earlier in this chapter, attorneys or witnesses make direct appeals to jurors and seek to elicit emotional responses. As noted, this may come in the form of pictures or videos of the victims, or a particularly moving testimony on the part of a witness. One factor that is relevant to the level of impact these appeals generate is individual differences in empathy. Empathy has been much debated, as researchers are not in universal agreement about whether the construct should be conceptualized as largely an affective process or a cognitive one (Bellas 2011). More complex models of empathy have been proposed. The Perception-Action Model of Emotional Empathy, for example, suggests that empathy reflects a process of automatic mimicry, which gives rise to emotional experiences unless inhibited (e.g., Preston and deWaal 2002). Nevertheless, despite the lack of universal agreement, empathy is regarded by many as a response that shares both cognitive and emotional facets and generally reflects the extent to which one reacts to the observed experiences of another (Davis 1983).

The role of empathy in juror decision making has been noted by legal theorists as critical to how jurors assign blame and sanction (e.g., Bandes 1996). As Arrigo and Williams (2003) have noted, empathy may be a zero-sum game whereby empathy for the victim may come at the expense of empathy for the defendant. In support of this view, Edelman (2006) found that in his study of race and death penalty judgments, White jurors reported more empathy for White victims than they did for Black victims, and subsequently gave less weight to mitigating factors when empathy for the victim was high. In that regard, our capacity to give substantial consideration to mitigating evidence, a critical component of the sentencing decision, is directly related to our capacity to summon empathy either for the victim or for the defendant. Along these lines, Deitz et al. (1982) found that as jurors empathized with the rape victim more, they were more likely to convict the defendant. Attorneys may intuitively understand the importance of empathy in legal judgments and present evidence in a manner that appeals to jurors' empathic nature. Haegerich and Bottoms (2000) remark that attorneys seek to evoke empathy in the jury for their client, and they reference a number of trial manuals that recommend this practice as evidence of this approach (e.g., Smith and Malandro 1985; Wright 1987).

In each of these instances, the degree to which one experienced empathy in the immediate context (i.e., state empathy) was investigated. However, individuals may differ in their tendency to experience empathy across a wide array of contexts (i.e., trait empathy). One of the most widely used measures of trait empathy is the

Interpersonal Reactivity Index (IRI; Davis 1980). This 28-item four-factor scale taps *perspective taking*, *empathic concern* (sympathy for others), *personal distress* (experienced anxiety in the face of distress from others), and *fantasy* (capacity to become imaginatively involved in the experiences of others). Davis (1983) has noted that perspective taking is the most cognitive of the four constructs measured by the IRI, whereas the remaining three constructs should be considered emotional. Other commonly used measures of empathy are more global measures such as the Emotional Empathy Scale (Mehrabian and Epstein 1972) and the Hogan Empathy Scale (Hogan 1969).

Despite the obvious relation between legal judgments and empathy, few studies have specifically examined how individual differences in trait empathy predict legal judgments. In one recent example, Henry et al. (2013) asked participants to sentence the defendant (life/death) after reading the penalty phase of a capital murder trial. Empathic concern, as measured by the IRI, significantly correlated with sentencing judgments, as those with greater empathy were less likely to vote for death, suggesting that participants high in IRI empathized with the defendant. Of course, one could imagine that in other circumstances (e.g., where VIS were given), jurors high in empathy might empathize more with the victim, leading to the opposite effect. In general, greater empathic concern might result in greater leniency for the defendant, as Henry, et al. found that IRI was negatively related to general support for the death penalty.

Another individual difference factor that may play a role in emotions and juror decision making is the tendency for individuals to approach or avoid emotional engagement. Maio and Esses (2001) developed the Need for Affect (NFA) scale, which is a 26-item scale consisting of two related subscales: approach and avoidance. A total NFA score is obtained by subtracting the avoidance score from the approach score. Individuals high in NFA may be inclined to attend to and process emotional information more extensively than those low in NFA (Maio and Esses 2001). Consequently, the role NFA may play in response to emotional trial stimuli merits investigation. In one recent study, Wevodau et al. (2014) varied the presence of victim impact statements containing a high proportion of emotional words. Regardless of the presence of VIS, the researchers found that jurors scoring higher on NFA made lengthier sentence recommendations than jurors scoring low on the scale. In another study, participants viewed a video of a defendant expressing either high or low levels of both verbal and nonverbal remorse and rendered sentencing judgments in a capital murder case. Those high in NFA rendered significantly more lenient sentences than participants scoring low on NFA (Corwin et al. 2012).

Individuals may also differ in their sensitivity and recognition of mood states. For example, interoceptive awareness (IA) is typically defined as the tendency to be aware of one's internal states (Herbert et al. 2011). IA is correlated with the intensity with which emotionally arousing stimuli are experienced (Herbert et al. 2007). More specifically, emotional awareness refers to the ability to recognize and understand one's own emotions. According to Boden and Thompson (2015), this capacity may be divided into the extent to which one *attends to*, and the extent to which one *understands,* his or her own emotions. Importantly, the degree to which

one understands experienced emotions can be distinguished from differences in the emotion itself, including its intensity (Coffey et al. 2003). To illustrate, individuals with alexithemia are unable to identify and describe their own emotions. These individuals are less able to process both verbal and nonverbal emotional stimuli (Herbert et al. 2011; Kano et al. 2003). This "emotion-blind" characteristic has been found to correlate positively with measures of trait aggression (Konrath et al. 2012). Consequently, we would expect that individual differences in emotional awareness/alexithemia would play an important role in the extent to which jurors might be vulnerable to experiencing the emotions of others, as well as the degree to which these emotional responses might be identifiable and perhaps correctable.

Finally, some individual differences might become relevant only for certain types of trials. For example, cases that arouse disgust (e.g., cases with gruesome crime scene photos) may also be moderated by individual differences for that particular emotion. Researchers have noted that there are individual differences in disgust sensitivity (how prone one is to feel disgust easily and intensely). Higher levels of disgust sensitivity have been associated with more conservative values and negative attitudes towards homosexuals (Inbar et al. 2009), as well as with increased health anxiety (Olatunji 2009). More importantly, disgust sensitivity affects legal decisions in cases that arouse disgust. For example, Stevenson et al. (2014) had participants read a case about a 17-year-old boy convicted of aggravated child molestation for soliciting oral sex from an intoxicated 15-year-old girl. They found that participants with higher disgust sensitivity were more likely to recommend that the juvenile be placed on the sex offender registry, and that disgust sensitivity was associated with less empathy for the juvenile and greater belief that the juvenile was a super-predator. In another study, Jones and Fitness (2008) examined the relationship between disgust sensitivity and verdicts in a murder trial. They found that disgust sensitivity was related to increased guilty verdicts, with a one standard deviation increase in disgust sensitivity scores related to a 169 % increase in the odds of voting guilty. In sum, though we have used disgust as an example, other individual differences in emotions (e.g., trait anger) might also be important moderators in specific cases. Specific emotions might be aroused in response to certain case characteristics and individual differences in those specific emotions might be quite relevant to decision making. However, aside from disgust sensitivity, we are not aware of other research that has attempted to document how individual differences in specific cases (e.g., trait anger in an anger provoking case) might moderate juror decisions.

Overall, a more developed understanding of the role emotions may play in juror judgments necessitates the ability to identify and account for individual differences in jurors' tendency to empathize with others, their sensitivity to the emotions aroused at trial, as well as the degree to which they are motivated to experience emotions (i.e., NFA). Understanding these individual differences in how emotions are attended to and perceived by jurors represents an important direction psychological scholars can take when examining the possible effects of emotions on juror decision making, as researchers cannot assume that the presence of a certain emotion at trial will always elicit the same level of response.

Jury Deliberations

Virtually none of the existing literature on emotions and juror decision making has included deliberating jurors. There is no study that we are aware of that has analyzed the content of deliberations, and few that have looked at decisions by pre- and post-deliberating jurors.[3] In 2005, one legal scholar asked the following question: "Might the experience of jury deliberation attenuate any such mood effects?" (Blumenthal 2005, p. 8). In his own study of mood and moral decision making, Blumenthal (2005) did not include deliberations, nor have most other researchers. Thus, the question posed by Blumenthal remains unanswered.

Although the research on jury deliberations is far from extensive, there remains a body of research that can shed light on deliberation's possible moderating effect on juror response to emotional evidence. On the one hand, deliberations allow for the passage of time between the initial response to the emotional evidence presented at trial, and the subsequent decision jurors will ultimately make when they render a verdict. During this period, strong emotions may subside. In addition, emotional jurors may be cautioned by other jurors to put their passions aside and discuss the evidence in a fair and impartial manner. On the other hand, the emotion-provoking evidence may be discussed the most, and therefore amplify the polarizing nature of group decisions (e.g., Myers and Lamm 1976). Subsequent decisions are then likely to reflect the evidence that received the most discussion, thus increasing the likelihood that jurors' emotions play a biasing role in their decisions.

There is a body of research on mock juror decision making that provides some clues as to how the deliberation process might moderate the effects of emotional responses on the part of jurors. For example, McCoy et al. (1999) examined juror reasoning in pre- and post-deliberation jurors and found that post-deliberation jurors showed a higher level of reasoning about the case than did pre-deliberation jurors. Pritchard and Keenan (2002) asked pre- and post-deliberation jurors to recall trial details and found that post-deliberation jurors had a slight but significantly better memory for details than pre-deliberation jurors without an increase in detail distortions. Generally, findings are consistent with what we know about how groups tend to have superior memories compared to individuals because in an ideal deliberation environment, details are shared among the group (e.g., transactive memory systems, see Peltakorpi 2008). Thus, there appears to be a general positive effect for deliberation with regard to memory and reasoning. When one considers a specific area of inquiry, though, this general principle falls apart.

Consider PTP, which, as discussed earlier, has been found to affect jurors' decisions; importantly, emotional PTP appears to be particularly memorable and difficult to minimize (Kerr 2010; Kramer et al. 1990). Ruva et al. have conducted several studies to examine how PTP affects deliberations and post-deliberation verdicts (Ruva et al. 2007; Ruva and Geunther 2014). In one study, Ruva et al.

[3]We would argue that other jury research has also ignored deliberations (see for example, Nuñez et al. 2011), but it is particularly true of the emotions/legal decision making research.

(2007) exposed half of their mock jurors to negative PTP and had them watch a murder trial a week later. Participants were randomly assigned to provide guilt verdicts with or without deliberation. The researchers found that, as expected, jurors exposed to PTP rendered more guilty verdicts than those who were not exposed to PTP, but importantly, there was no difference between deliberating and nondeliberating jurors (i.e., the effect of PTP was the same for both groups). In another study (Ruva and Guenther 2014), they analyzed the content of the jury deliberations and found that jurors exposed to negative PTP were more likely to discuss ambiguous trial facts as more supportive of the prosecutor's case and presented more prosecution evidence and less defense evidence during discussion. Importantly, all juries discussed the PTP, and only 33 % of the time was the discussion challenged by another juror. In sum, the PTP literature suggests that juror biases due to PTP found at the juror level are not corrected with deliberation. In fact, the deliberations appear to reinforce the effects of PTP found at the individual level.

On the other hand, there are examples in the psycholegal literature that find deliberations which lead to better outcomes by minimizing juror level biases. For example, with regard to jurors' ability to disregard inadmissible evidence, there is a large literature that shows jurors are unable to disregard inadmissible evidence and that this evidence bleeds into their verdict decisions (for a meta-analysis see Steblay et al. 2006). London and Nunez (2000) found that they could reverse this very robust effect by including deliberations as a moderating variable. In a series of two studies, they compared pre- and post-deliberating jurors' verdicts. Participants were assigned to either a control (no extra evidence), inadmissible (extra evidence ruled inadmissible), or admissible (extra evidence ruled admissible) condition. Prior to deliberation, a significant number of mock jurors in the inadmissible condition appeared unable to disregard the inadmissible evidence (much like what others have found). Specifically, 33 % of the control, but 60 % of the inadmissible and 96 % of the admissible condition jurors voted guilty. However, post-deliberation, only 17 % of control and 28 % of inadmissible condition jurors voted guilty (while 100 % of the jurors in the admissible condition voted guilty).

The upshot of this brief review is that deliberations might attenuate or exacerbate mood effects. Generally, deliberations appear to help jurors remember details and reason about a case at a higher level (McCoy et al. 1999; Pritchard and Keenan 2002). But, the general positive effects of deliberation may not apply to all cases. It might attenuate some effects (e.g., inadmissible evidence) but not others (e.g., pretrial publicity). What then might we expect to find with regard to emotion?

Kaplan and Miller (1978) published one of the few studies that examined how emotions impact juror and jury decisions. Jurors watched a live reenactment of an attempted manslaughter trial (the defendant was accused of stabbing the victim with a broken bottle in a bar fight). Jurors were assigned to conditions in which either the defense or prosecuting attorney was annoying. The annoying attorney repeated questions, badgered witnesses, and made obnoxious side comments to the jurors. After viewing the reenactment, jurors indicated their verdict on a 20-point scale and then deliberated. Following deliberation, jurors again indicated their verdict. Kaplan and Miller (1978) found that prior to deliberations, mock jurors attributed the

greatest amount of guilt when the defense attorney was annoying and the least amount of guilt when the prosecutor was annoying (i.e., jurors' irritation spilled over into their verdict decisions). After deliberation, the biasing effect of the attorneys' behaviors disappeared. Thus, deliberations attenuated jurors' pre-deliberation bias.

Though the findings are mixed in some areas, with some finding deliberations help and some finding they hurt, all of the literature points to the importance of deliberations as a potential moderator. Pre- and post-deliberation jurors are different, and too few of the emotion and legal decision making literature includes deliberations as a potential moderator.

Summary

The role emotions play in legal judgments is complex, and added to this complexity is the fact that elements are present (and rarely controlled) that can reduce or increase the impact of emotions on legal judgments. We presented three potential moderators (jury instructions, individual differences in emotionality, and jury deliberations) that should be explored more fully in future research; there are no doubt others worth considering.

Discussion

In this chapter, we described theories of emotion that can help guide the psycholegal research. We presented ample evidence to show that emotions are indeed aroused at trial and that some emotions impact jurors' decisions, as is true in the case of anger. Jurors who become angry are more likely to decide against the defendant (e.g., Georges et al. 2013) and are more punitive (e.g., Nunez et al. 2015a). On the other hand, other emotions, such as sadness have not been found to impact decisions (e.g., Nunez et al. 2015a) and some research shows that sadness can promote higher reasoning about trial evidence (Semmler and Brewer 2002). Thus, all negative emotions are not the same. Finally, we described several factors that might moderate the effect of emotions on decisions. These potential moderators (jury instructions, individual differences in emotionality, and jury deliberations) have not received much research attention. However, what research we do have suggests that these moderators are important and worthy of consideration if we want to provide guidance to legal professionals about limiting the effects of emotions in the courtroom.

Suggestions for Future Research

One important variable that is typically examined in the emotion literature, but has not been adequately addressed by psycholegal researchers, is the difference between incidental and integral emotion. One of the difficulties confronting researchers in this area is that studies that systematically vary emotions without altering the trial information (e.g., through mood inductions) will lack ecological validity (i.e., the emotion was not elicited by the trial information). Specifically, it becomes unclear whether emotional misattributions like these adequately capture the experience of actual jurors whose emotions arise as a consequence of the trial. By contrast, studies that vary the content of the trial in a way to elicit specific emotions undoubtedly suffer from the criticism that in manipulating content so that the emotion evoked is integral to the task at hand (e.g., legal decision), emotion is confounded with content in a way that makes causal interpretations of the findings impossible. Here we seek converging evidence to build a reliable model of emotions and legal judgments where studies involving mood manipulations produce findings that are consistent with studies in which the same emotions are elicited through manipulations in trial content. Unfortunately, much of this work has yet to be done, and we remain in the early stages of understanding integral versus incidental emotions and juror judgments.

Most of the research that explores emotions in the courtroom focuses only on verdict and sentencing. We know little about how emotions affect other legal decisions. For example, we mentioned that fear was not an emotion found to be aroused at trial (at least not in mock jury studies). However, fear might be relevant to accused suspects who are offered a plea deal. If fear interferes with rational judgment, it could lead to poor decision making. Anger has been found to affect jurors' decisions, but there is little research on how anger can affect other legally relevant behaviors (e.g., parole decisions). Godbole, Blandon-Gitlin, and Masip (2015) recently examined how induced emotions could impact the decision to use low or high coercive interrogation techniques. They did not find any significant effect of induced emotion on the endorsement of interrogation techniques, but their college student participants were only asked if they might use certain techniques. One might imagine that a more realistic paradigm with higher stakes (e.g., interrogating a student accused of cheating) might yield significant results. One could also imagine that emotions could impact bail decisions, parole decisions, and more. Yet, we know little about these other legal decisions.

Finally, it would be helpful to know what beliefs exist in the legal community (i.e., among judges and lawyers) about emotions and decisions. In a field study of victim impact statements, Myers et al. (2015) found that judges in capital trials were vigilant about overly emotional statements at their trials and were particularly careful not to allow incendiary statements from victims' families. We suspect that these judges were not well versed in the literature presented in this chapter, but rather relied on their own lay theories of emotions. However, we know little about the lay theories of legal professionals and whether they are correct. Are judges

really concerned about overly sad testimony? What about disgust? What judges and other legal professionals believe about emotions undoubtedly affects their courtroom practices, and researchers should determine if there are inconsistencies between judges' beliefs and current research.

Final Thoughts

On the first day of the murder trial of James Holmes who was accused of killing 12 and wounding 70 in an Aurora, Colorado movie theater, District Attorney George Brauchler held up a photo of the back door of the Aurora movie theater and said, "Through this door is horror. Through this door is bullets, blood, brains, and bodies" (Barajas 2015). One by one, DA Brauchler "recounted the people who died, and then showed images of how they had been shot down. "Boom!" he said, echoing the shotgun blasts. As he spoke, members of the jury were seen crying, as were family members of the shooting victims, who came from around the country to observe the trial" (Healy 2015). Frantic 911 calls from some of the survivors were played to jurors, including a call from 13-year-old Kaylan Bailey inside the movie theater. The 911 operator tries to explain to Ms. Bailey how to perform CPR on the 6-year-old child she was babysitting, while the movie in the background continues to play. She's heard screaming "I need help!" and "I can't hear you!" as she tearfully tries to hear what the 911 operator is telling her to do (Sylte 2015). Jurors heard multiple 911 calls, listened to the stories from 12 survivors, and were told of the carnage that confronted police officers when they arrived at the theater. This all occurred in the first week of the trial. The James Holmes trial is one of many high profile cases that illustrate the presence of emotionally charged evidence and testimony at trial that may have subsequently impacted how jurors interpreted information and ultimately found Holmes guilty of 24 counts of murder and 140 counts of attempted murder.

Legal trials are complex and the tangled ways that emotions can affect decisions has yet to be fully examined. Emotions such as disgust might be aroused one day as gruesome crime scene photos are produced, while emotions of sadness might be elicited as jurors hear a victim impact statement. How this emotional roller coaster affects decisions is unknown because most mock jury studies are short and typically focus on the impact of only one emotion. The degree to which emotions may wax and wane throughout the trial, the specific emotions elicited, the tendency for jurors to differ in their emotional responsivity, and the particular legal decision at hand are all factors that have been shown in isolation to impact legal decisions, yet we know they occur in combination and have not begun to study the interactive effects of these important variables. As we said earlier, researchers have only begun to "scratch the surface."

References

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin, 126*, 556–574. doi:10.1037//0033-2909.126.4.556

Arrigo, B. A., & Williams, C. (2003). Victim vices, victim voices, and impact statements: On the place of emotion and the role of restorative justice in capital sentencing. *Crime and Delinquency, 49*, 603–626. doi:10.1177/0011128703252408

Baldry, A. C. (1996). Rape victims' risk of secondary victimization by police officers. *Issues in Criminological and Legal Psychology, 25*, 65–68.

Bandes, S. (1996). Empathy, narrative, and victim impact statements. *University of Chicago Law Review, 63*, 361–412. doi:10.2307/1600234

Bagozzi, R. P., & Moore, D. J. (1994). Public service advertisements: Emotions and empathy guide prosocial behavior. *Journal of Marketing, 58*, 56–70. doi:10.2307/1252251

Barajas, J. (2015, April 27). 'Through this door is horror': Opening statements begin in theater shooting trial. *PBS News Hour*. Retrieved from http://www.pbs.org

Batson, C. D., Sager, K., Garst, E., Kang, M., Rubchinsky, K., & Dawson, K. (1997). Is empathy-induced helping due to self-other merging? *Journal of Personality and Social Psychology, 73*, 495–509. doi:10.1037/0022-3514.73.3.495

Bellas, C. M. (2011). *"I feel your pain": How juror empathy affects death penalty verdicts*. Unpublished doctoral dissertation, Kent State University, Kent, OH. Retrieved from http://search.proquest.com/docview/745598995?accountid=14793

Blumenthal, J. A. (2005). Does mood influence moral judgment? An empirical test with legal and policy implications. *Law and Psychology Review, 29*(1), 1–28. Retrieved from http://search.proquest.com/docview/229395651?accountid=14793

Boden, M. T., & Thompson, R. J. (2015). Facets of emotional awareness and associations with emotion regulation and depression. *Emotion*. Advance Online Publication. doi:10.1037/emo0000057

Bodenhausen, G. V., Sheppard, L. A., & Kramer, G. P. (1994). Negative affect and social judgment: The differential impact of anger and sadness. *European Journal of Social Psychology, 24*(1), 45–62. doi:10.1002/ejsp.2420240104

Booth v. Maryland. (1987). 482 U.S. 496.

Boppre, B., & Miller, M. K. (2014). How victim and execution impact statements affect mock jurors' perceptions, emotions, and verdicts. *Victims and Offenders, 9*, 413–435. doi:10.1080/15564886.2013.845124

Bornstein, B. H., & Nemeth, R. J. (1999). Jurors' perceptions of violence: A framework for inquiry. *Aggression and Violent Behavior, 4*, 77–92. doi:10.1016/S1359-1789(97)00059-1

Bornstein, B. H., & Wiener, R. L. (2006). Introduction to the special issue on emotion in legal judgment and decision making. *Law and Human Behavior, 30*, 115–118. doi:10.1007/s10979-006-9023-2

Bower, G. H. (1981). Mood and memory. *American Psychologist, 36*(2), 129–148. doi:10.1037/0003-066X.36.2.129

Brehm, S., & Brehm, J. W. (1981). *Psychological reactance: A theory of freedom and control*. New York: Academic Press.

Bright, D. A., & Goodman-Delahunty, J. (2004). The influence of gruesome verbal evidence on mock juror verdicts. *Psychiatry, Psychology, and Law, 11*, 154–166. doi:10.1375/pplt.2004.11.1.154

Bright, D. A., & Goodman-Delahunty, J. (2006). Gruesome evidence and emotion: Anger, blame, and jury decision-making. *Law and Human Behavior, 30*, 182–202. doi:10.1007/s10979-006-9027-y

Bright, D. A., & Goodman-Delahunty, J. (2011). Mock juror decision making in a civil negligence trial: The impact of gruesome evidence, injury severity, and information processing route. *Psychiatry, Psychology, and Law, 18*, 439–459. doi:10.1080/13218719.2010.492095

Butler, B. (2008). The role of death qualification in venirepersons' susceptibility to victim impact statements. *Psychology, Crime & Law, 14*(2), 133–141. doi:10.1080/10683160701483534

Capestany, B. H., & Harris, L. T. (2014). Disgust and biological descriptions bias logical reasoning during legal decision-making. *Social Neuroscience, 9*, 265–277. doi:10.1080/17470919.2014.892531

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin, 135*, 183–204. doi:10.1037/a0013965

Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences, 11*(9), 393–399. doi:10.1016/j.tics.2007.08.005

Clore, G. L., Wyer, R. J., Dienes, B., Gasper, K., Gohm, C., & Isbell, L. (2001). Affective feelings as feedback: Some cognitive consequences. In L. L. Martin, G. L. Clore, L. L. Martin, & G. L. Clore (Eds.), *Theories of mood and cognition: A user's guidebook* (pp. 27–62). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Coffey, E., Berenbaum, H., & Kerns, J. K. (2003). The dimensions of emotional intelligence, alexithymia, and mood awareness: Associations with personality and performance on an emotional stroop task. *Cognition and Emotion, 17*, 671–679. doi:10.1080/02699930302304

Corwin, E. P., Cramer, R. J., Griffin, D. A., & Brodsky, S. L. (2012). Defendant remorse, need for affect, and juror sentencing decisions. *Journal of the American Academy of Psychiatry and Law*, *40*, 41–49. Retrieved from http://www.jaapl.org/content/40/1/41.full.pdf

Coyne, J. C. (1976). Depression and the response of others. *Journal of Abnormal Psychology, 85*, 186–193. doi:10.1037/0021-843X.85.2.186

Cush, R. K., & Delahunty, J. G. (2006). The influence of limiting instructions on processing and judgments of emotionally evocative evidence. *Psychiatry, Psychology and Law, 13*, 10–123. doi:10.1375/pplt.13.1.110

Daftary-Kapur, T., Penrod, S. D., O'Connor, M., & Wallace, B. (2014). Examining pretrial publicity in a shadow jury paradigm: Issues of slant, quantity, persistence and generalizability. *Law and Human Behavior, 38*, 462–477. doi:10.1037/lhb0000081

Dahl, J., Enemo, I., Drevland, G. C. B., Wessel, E., Eilertsen, D. E., & Magnussen, S. (2007). Displayed emotions and witness credibility: A comparison of judgments by individuals and mock juries. *Applied Cognitive Psychology, 21*, 1145–1155. doi:10.1002/acp.1320

Darwin, C. (1872). *The expression of the emotions in man and animals*. Chicago: University of Chicago Press. (1965).

Davis, M. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 1980*(10), 85.

Davis, M. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*(1), 113–126. doi:10.1037/0022-3514.44.1.113

Deitz, S. R., Blackwell, K. T., Daley, P. C., & Bentley, B. J. (1982). Measurement of empathy toward rape victims and rapists. *Journal of Personality and Social Psychology, 43*, 372–384. doi:10.1037/0022-3514.43.2.372

Douglas, K. S., Lyon, D. R., & Oglof, J. R. P. (1997). The impact of graphic photographic evidence on mock jurors' decisions in a murder trial: Probative or prejudicial? *Law and Human Behavior, 21*, 485–501. Retrieved from http://www.jstor.org/stable/1394329

Edelman, B. C. (2006). *Racial prejudice, juror empathy, and sentencing in death penalty cases*. New York: LFB Scholarly.

Edwards, K., & Bryan, T. S. (1997). Judgmental biases produced by instructions to disregard: The (paradoxical) case of emotional information. *Personality and Social Psychology Bulletin, 23*, 849–864. doi:10.1177/0146167297238006

Feigenson, N. (2000). *Legal Blame: How jurors think and talk about accidents*. Washington, D.C: American Psychological Association.

Feigenson, N. (2010). Emotional influences on judgments of legal blame: How they happen, whether they should, and what to do about it. In B. H. Bornstein & R. L. Wiener (Eds.), *Emotion and the law: Psychological Perspectives* (pp. 45–96). New York, NY, US: Springer Science+Business Media.

Feigenson, N., & Spiesel, C. (2009). *Law on display: The digital transformation of legal persuasion*. New York, NY: New York University Press.

Fein, S., McCloskey, A. L., & Tomlinson, T. M. (1997). Can the jury disregard that information? The use of suspicion to reduce the effects of pretrial publicity and inadmissible evidence. *Personality and Social Psychology Bulletin, 23*, 1215–1226. doi:10.1177/01461672972311008

Fishfader, V. L., Howells, G. N., Katz, R. C., & Teresi, P. S. (1996). Evidential and extralegal factors in juror decisions: Presentation mode, retention, and level of emotionality. *Law and Human Behavior, 20*(5), 565–572. doi:10.1007/BF01499042

Forgas, J. P. (1992). Affect in social judgments and decisions: A multi-process model. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 227–275). San Diego, CA: Academic Press.

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin, 117*, 39–66. doi:10.1037/0033-2909.117.1.39

Georges, L. C., Wiener, R. L., & Keller, S. R. (2013). The angry juror: Sentencing decisions in first-degree murder. *Applied Cognitive Psychology, 27*, 156–166. doi:10.1002/acp.2880

Godbole, G., Blandon-Gitlin, I., & Masip, J. (2015). *Exploring the role of perceivers' emotional states on investigative decision*. Poster presented at the annual conference of the Association for Psychological Science, May, 2015.

Goldberg, J. H., Lerner, J. S., & Tetlock, P. E. (1999). Rage and reason: The psychology of the intuitive prosecutor. *European Journal of Social Psychology, 29*(5–6), 781–795. doi:10.1002/(SICI)1099-0992(199908/09)29:5/6<781:AID-EJSP960>3.0.CO;2-3

Goldie, P. (1999). How we think of others' emotions. *Mind and Language, 14*, 394–423. doi:10.1111/1468-0017.00118

Greene, E., Koehring, H., & Quiat, M. (1998). Victim impact evidence in capital cases: Does the victims' character matter? *Applied Social Psychology, 28*, 145–156. doi:10.1111/j.1559-1816.1998.tb01697.x

Hackett, L., Day, A., & Mohr, P. (2008). Expectancy violation and perceptions of rape victim credibility. *Legal and Criminological Psychology, 13*, 323–334. doi:10.1348/135532507X228458

Haegerich, T., & Bottoms, B. (2000). Empathy and jurors' decisions in patricide trials involving child sexual assault allegations. *Law and Human Behavior, 24*, 421–430. ISSN: 0147-7307/00/0800-0421

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814–834. doi:10.1037/0033-295X.108.4.814

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional Contagion. *Current Directions in Psychological Science, 2*, 96–99. Retrieved from http://www.jstor.org/stable/20182211

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion*. New York: Cambridge University Press.

Healy, J. (2015, April 27) Colorado shooting trial of James E. Holmes pits a calculated killer against an erratic mind. *The New York Times.* Retrieved from http://www.nytimes.com

Heath, W. P., & Grannemann, B. D. (2014). How video image size interacts with evidence strength, defendant emotion, and the defendant-victim relationship to alter perceptions of the defendant. *Behavioral Sciences and the Law, 32*, 496–507. doi:10.1002/bsl.2120

Heath, W. P., Grannemann, B. D., & Peacock, M. A. (2004). How the defendant's emotion level affects mock jurors' decisions when presentation mode and evidence strength are varied. *Journal of Applied Social Psychology, 34*, 624–664. doi:10.1111/j.1559-1816.2004.tb02563.x

Hester, R. K., & Smith, R. E. (1973). Effects of mandatory death penalty on the decisions of simulated jurors as a function of heinousness of the crime. *Journal of Criminal Justice, 1*, 319–326. doi:10.1016/0047-2352(73)90034-2

Hendrick, C., & Shaffer, D. R. (1975). Murder: Effects of number of killers and victim mutilation on simulated jurors' judgments. *Bulletin of the Psychonomic Society, 6*, 313–316. doi:10.3758/BF03336673

Henry, S., Mitchell, K., Kaniuka, A., Broszkiewicz, N., Myers, B., & Nunez, N. (2013). *Capital Juror's Sentencing and Victim Impact Statements: Empathy Predicts Mitigation Influence and*

Sentencing Judgments. Poster presented at the Annual Conference of the Association for Psychological Science, May 25.

Herbert, B., Herbert, C., & Pollatos, O. (2011). On the relationship between interoceptive awareness and alexithymia: Is interoceptive awareness related to emotional awareness? *Journal of Personality, 79*, 1149–1175. doi:10.1111/j.1467-6494.2011.00717.x

Herbert, B. M., Pollatos, O., & Schandry, R. (2007). Interoceptive sensitivity and emotion processing: An EEG study. *International Journal of Psychophysiology, 65*, 214–227. doi:10.1016/j.ijpsycho.2007.04.007

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology, 33*, 307–316. doi:10.1037/h0027580

Horowitz, I. A., Kerr, N. L., Park, E. S., & Gockel, C. (2006). Chaos in the courtroom reconsidered: Emotional bias and juror nullification. *Law and Human Behavior, 30*, 163–181. doi:10.1007/s10979-006-9028-x

Inbar, Y., Pizarro, D. A., & Bloom, P. (2009). Conservatives are more easily disgusted than liberals. *Cognition and Emotion, 23*, 714–725. doi:10.1080/02699930802110007

James, W. (1922). What is an emotion? In C. G. Lange, W. James, C. G. Lange, W. James (Eds.), *The emotions* (Vol. 1, pp. 11–30). Baltimore, MD, US: Williams & Wilkins Co. doi:10.1037/10735-001

Jones, A., & Fitness, J. (2008). Moral hypervigilance: The influence of disgust sensitivity in the moral domain. *Emotion, 8*, 613–627. doi:10.1037/a0013435

Kadous, K. (2001). Improving jurors' evaluations of auditors in negligence cases. *Contemporary Accounting Research, 18*, 425–444. doi:10.1506/GM8A-HNPH-LL3L-98FY

Kano, M., Fukudo, S., Gyoba, J., Kamachi, M., Tagawa, M., Mochizuki, H., et al. (2003). Specific brain processing of facial expressions in people with alexithymia: An H2 150 PET study. *Brain, 126*, 1474–1484. doi:10.1093/brain/awg131

Kaplan, M. F., & Miller, L. E. (1978). Reducing the effects of juror bias. *Journal of Personality and Social Psychology, 36*, 1443–1455. doi:10.1037/0022-3514.36.12.1443

Kassin, S. M., & Garfield, D. A. (1991). Blood and guts: General and trial-specific effects of videotaped crime scenes on mock jurors. *Journal of Applied Social Psychology, 21*, 1459–1472. doi:10.1111/j.1559-1816.1991.tb00481.x

Kaufmann, G., Drevland, G. C. B., Wessel, E., Overskeid, G., & Magnussen, S. (2003). The importance of being earnest: Displayed emotions and witness credibility. *Applied Cognitive Psychology, 17*, 21–34. doi:10.1002/acp.842

Kelly v. California, 129 S. Ct. 564 (2008).

Keltner, D., Ellsworth, P. C., & Edwards, K. (1993). Beyond simple pessimism: Effects of sadness and anger on social perception. *Journal of Personality and Social Psychology, 64*(5), 740–752. doi:10.1037/0022-3514.64.5.740

Kemp, E., Kennett-Hensel, P. A., & Kees, J. (2013). Pulling on the heartstrings: Examining the effects of emotions and gender in persuasive appeals. *Journal of Advertising, 42*, 69–79. doi:10.1080/00913367.2012.749084

Kerr, N. L., Boster, F., Callen, C. R., Braz, M. E., O'Brien, B., & Horowitz, I. (2008). Jury nullification instructions as amplifiers of bias. *International Commentary on Evidence, 6*, 1554-4567. doi:10.2202/1554-4567.1068

Kerr, N. L. (2010). Explorations in juror emotion and juror judgment. In B. H. Bornstein & R. L. Wiener (Eds.), *Emotions and the law: Psychological perspectives*. Springer: New York, NY.

Kligyte, V., Connelly, S., Thiel, C., & Devenport, L. (2013). The influence of anger, fear, and emotion regulation on ethical decision making. *Human Performance, 26*(4), 297–326. doi:10.1080/08959285.2013.814655

Konecni, V. J., Ebbesen, E. B., & Hock, R. R. (1996). Factors affecting simulated jurors' decisions in capital cases. *Psychology, Crime, and Law, 2*, 269–297. doi:10.1080/10683169608409784

Konrath, S., Novin, S., & Li, T. (2012). Is the relationship between alexithymia and aggression context dependent? Impact of group membership and belief similarity. *Personality and Individual Differences, 53*, 329–334. doi:10.1016/j.paid.2012.04.004

Kramer, G., Kerr, N. L., & Carroll, J. S. (1990). Pretrial publicity, judicial remedies, and juror bias. *Law and Human Behavior, 14*, 409–438. doi:10.1007/BF01044220

Laird, J. D., & Bresler, C. (1992). The process of emotional experience: A self-perception theory. In M. S. Clark (Ed.), *emotion* (pp. 213–234). Thousand Oaks, CA: Sage Publications.

Laster, K., & O'Malley, P. (1996). Sensitive new age laws: The reassertion of emotionality in law. *International Journal of Sociology of Law, 24*, 21–40. doi:10.1006/ijsl.1996.0002

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgment and choice. *Cognition and Emotion, 14*(4), 473–493. doi:10.1080/026999300402763

Logan, W. (1999). Through the past darkly: A survey of the uses and abuses of victim impact evidence in capital trials. *Arizona Law Review, 41*, 143–192. Retrieved from http://heinonline.org.libproxy.uwyo.edu/HOL/Page?handle=hein.journals/arz41&id=157

London, K., & Nunez, N. (2000). The effect of jury deliberations on jurors' propensity to disregard inadmissible evidence. *Journal of Applied Psychology, 85*, 932–939. doi:10.1037/0021-9010.85.6.932

Maio, G. R., & Esses, V. M. (2001). The need for affect: Individual differences in the motivation to approach or avoid emotions. *Journal of Personality, 69*, 583–615. doi:10.1111/14676494.694156

Maio, G. R., & Haddock, G. (2007). Attitude change. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: A handbook of basic principles* (pp. 565–586). New York: Guilford.

McCoy, M. L., Nunez, N., & Dammeyer, M. M. (1999). The effect of jury deliberations on jurors' reasoning skills. *Law and Human Behavior, 23*, 557–574. doi:10.1023/A:1022348229558

McGowan, M. G., & Myers, B. (2004). Who is the victim anyway? The effects of bystander victim impact statements on mock juror sentencing decisions. *Violence and Victims, 19*, 357–374. doi:10.1891/vivi.19.3.357.6576

McNair, D. M., Lorr, M., & Droppleman, L. F. (1981). *EITS manual: Profile of mood states*. San Diego, CA: Educational and Industrial Training Service.

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality, 40*, 525–543. doi:10.1111/j.1467-6494.1972.tb00078.x

Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin, 83*, 602–627. doi:10.1037/0033-2909.83.4.602

Myers, B., Lynn, S. J., & Arbuthnot, J. (2002). Victim impact testimony and juror judgments: The effects of harm information and witness demeanor. *Journal of Applied Social Psychology, 32*, 2393–2412. doi:10.1111/j.1559-1816.2002.tb01869.x

Myers, B., Nunez, N., Kehn, A. & Wilkowski, B. (2015). *The heterogeneity of victim impact statements in capital trial sentencing*. Unpublished manuscript, Department of Psychology, University of North Carolina-Wilmington, Wilmington, NC.

Myers, B., Roop, A., Kalnen, D., & Kehn, A. (2013). Victim impact statements and crime heinousness: A test of the saturation hypothesis. *Psychology, Crime, and Law, 19*, 129–143. doi:10.1080/1068316X.2011.614244

Nadler, J., & Rose, M. R. (2003). Victim impact testimony and the psychology of punishment. *Cornell Law Review, 88,* 419–456. Retrieved from http://go.galegroup.com.libproxy.uwyo.edu/ps/i.do?id=GALE%7CA97175572&v=2.1&u=wylrc_uwyoming&it=r&p=LT&sw=w&asid=c4e9cb7097c048f93d7957dd8456f5ed

Nemeth, R. J. (2002). *The impact of gruesome evidence on mock juror decision making: The role of evidence characteristics and emotional response*. Unpublished doctoral dissertation, Louisiana State University, Baton Rouge. Retrieved from http://search.proquest.com/docview/305555346?accountid=14793

Neumann, R., & Strack, F. (2000) Mood contagion: The automatic transfer of mood between persons. *Journal of Personality and Social Psychology, 79*, 211–223. doi:10.1037//0022-3514.79.2.211

Nuñez, N., McCrea, S. M., & Culhane, S. E. (2011). Jury decision making research: Are researchers focusing on the mouse and not the elephant in the room? *Behavioral Sciences and the Law, 29*, 439–451. doi:10.1002/bsl.967

Nunez, N. Schweitzer, K., Chai, C. A., & Myers, B. (2015a). Negative emotions felt during trial: The effect of fear, anger, and sadness on juror decisions. *Applied Cognitive Psychology, 29*, 200–209. doi:10.1002/acp.3094

Nunez, N., Schweitzer, K., & Myers, B. (2015b, March 21). *The impact of angry versus sad victim impact statements on sentencing decisions in a capital trial*. Paper presented at the annual American Psychological—Legal Society conference, San Diego, CA.

Ogloff, J. R. P., & Vidmar, N. (1994). The impact of pretrial publicity on jurors: a study to compare the relative effects of television and print media in a child sexual abuse case. *Law and Human Behavior, 13*, 507–525. doi:10.1007/BF01499171

Olatunji, B. O. (2009). Incremental specificity of disgust propensity and sensitivity in the prediction of health anxiety dimensions. *Journal of Behavioral Therapy and Experimental Psychiatry, 40*, 230–239. doi:10.1016/j.jbtep.2008.10.003

Oliver, E., & Griffitt, W. (1976). Emotional arousal and objective judgment. *Bulletin of the Psychonomic Society, 8*, 399–400. doi:10.3758/BF03335179

Parker, M. T., & Isbell, L. M. (2010). How I vote depends on how I feel: The differential impact of anger and fear on political information processing. *Psychological Science, 21*(4), 548–550. doi:10.1177/0956797610364006

Paternoster, R., & Deise, J. (2011). A heavy thumb on the scale: The effect of victim impact evidence on capital decision making. *Criminology, 49*, 129–161. doi:10.1111/j.1745-9125.2010.00220.x

Payne v. Tennessee, 501 U.S. 808 (1991).

Peace, K. A., & Valois, R. L. (2014). Trials and tribulations: Psychopathic traits, emotion, and decision-making in an ambiguous case of sexual assault. *Psychology, 5*, 1239–1253. doi:10.4236/psych.2014.510136

Peltakorpi, V. (2008). Transactive memory systems. *Review of General Psychology, 4*, 378–394. doi:10.1037/1089-2680.12.4.378

Pennington, N., & Hastie, R. (1992). Explaining evidence: Tests of the story model for juror decision making. *Journal of Personality and Social Psychology, 62*, 189–206. doi:10.1037/0022-3514.62.2.189

Pham, M. T., Cohen, J. B., Pracejus, J. W., & Hughes, G. D. (2001). Affect monitoring and the primacy of feelings in judgment. *Journal of Consumer Research, 28*(2), 167–188. doi:10.1086/322896

Platania, J., & Berman, G. L. (2006). The moderating effect of judge's instructions on victim impact testimony in capital case. *Applied Psychology in Criminal Justice, 2*, 84–101.

Posner, Richard A. (1999). Emotion versus Emotionalism in Law. In Susan A. Bandes (Ed.), *The passions of law*. New York, NY: New York University Press.

Preston, S. D., & deWaal, F. M. (2002). Empathy: Its ultimate and proximate causes. *Behavioral and Brain Sciences, 25*, 1–20. doi:10.1017/S0140525X02000018

Pritchard, M., & Keenan, J. (2002). Does jury deliberation really improve jurors' memories? *Applied Cognitive Psychology, 16*, 589–601. doi:10.1002/acp.816

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (2002). From mirror neurons to imitation: Facts and speculations. In A. N. Meltzoff & W. Prinz (Eds.), *The imitative mind: Development, evolution, and brain bases* (pp. 247–266). New York, NY: Cambridge University Press.

Robinson, D. T., Smith-Lovin, L., & Tsoudis, O. (1994). Heinous crime or unfortunate accident? The effects of remorse on responses to mock criminal confessions. *Social Forces, 73*, 175–190. doi:10.2307/2579922

Rose, M. R., Nadler, J., & Clark, J. (2006). Appropriately upset? Emotion norms and perceptions of crime victims. *Law and Human Behavior, 30*, 203–219. doi:10.1007/s10979-006-90303

Ruva, C. L., & Guenther, C. C. (2014). From the shadows into the light: How pretrial publicity and deliberation affect mock jurors' decisions, impressions, and memory. *Law and Human Behavior, 39*, 294–310. doi:10.1037/lhb0000117

Ruva, C. L., Guenther, C. C., & Yarbrough, A. (2011). Positive and negative pretrial publicity: The roles of impression formation, emotion, and predecisional distortion. *Criminal Justice and Behavior, 38,* 511–534. doi:10.1037/1076-898X.14.3.226, doi:10.1177/0093854811400823

Ruva, C. L., & McEvoy, C. (2008). Negative and positive pretrial publicity affect juror memory and decision making. *Journal of Experimental Psychology: Applied, 14*, 226–235. doi:10.1037/1076-898X.14.3.226

Ruva, C., McEvoy, C., & Bryant, J. B. (2007). Effects of pretrial publicity and jury deliberation on juror bias and source memory errors. *Applied Cognitive Psychology, 21*, 45–67. doi:10.1002/acp.1254

Salekin, R. T., Ogloff, J. R. P., McFarland, C., & Rogers, R. (1995). Influencing jurors' perceptions of guilt: Expression of emotionality during testimony. *Behavioral Sciences and the Law, 13*, 293–305. doi:10.1002/bsl.2370130208

Salerno, J. M., & Peter-Hagene, L. C. (2013). The interactive effect of anger and disgust on moral outrage judgments. *Psychological Science, 24*, 2069–2078. doi:10.1177/0956797613486988

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45*(3), 513–523. doi:10.1037/0022-3514.45.3.513

Schweitzer, K., Estrada, V., Ferguson, L., & Nunez, N. (2015, March 19). *The effect of disgust on juror/jury decisions: Evidence for a disposal effect*. Paper presented at the annual American Psychological—Legal Society conference, San Diego, CA.

Seelye, K. Q. (2015, March 30). Prosecution in Boston Marathon trial closes with raw details of killings. *The New York Times*, pp. A-16.

Semmler, C., & Brewer, N. (2002). Effects of mood and emotion on juror processing and judgments. *Behavioral Sciences and the Law, 20*, 423–436. doi:10.1002/bsl.502

Smith, L. J., & Malandro, L. A. (1985). *Courtroom communication strategies*. New York: Kluwer.

Steblay, N. M., Besirevic, J., Fulero, S. M., & Jimenez-Lorente, B. (1999). The effects of pretrial publicity on juror verdicts: A meta-analytic review. *Law and Human Behavior, 23*, 219–235. doi:10.1023/A:1022325019080

Steblay, N., Hosch, H. M., Culhane, S. E., & McWethy, A. (2006). The impact on juror verdicts of judicial instruction to disregard inadmissible evidence: A meta-analysis. *Law and Human Behavior, 30*, 469–92. doi:10.1007/s10979-006-9039-7

Stevenson, M. C., Malie, S. E., Totton, R. R., & Reeves, R. D. (2014). Disgust sensitivity predicts punitive treatment of juvenile sex offenders: The role of empathy, dehumanization, and fear. *Analysis of Social Issues and Public Policy, 0*, 1–20. doi:10.1111/asap.12068

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology, 54*, 768–777. doi:10.1037/0022-3514.54.5.768

Sullins, E. S. (1991). Emotional contagion revisited: Effects of social comparison and expressive style on mood convergence. *Personality and Social Psychology Bulletin, 17*, 166–174. doi:10.1177/014616729101700208

Sylte, A. (2015, April 28). Aurora theater trial blog: Day 1. *9 News.* Retrieved from http://www.9news.com

Taute, H. A., McQuitty, S., & Sautter, E. P. (2011). Emotional information management and responses to emotional appeals. *Journal of Advertising, 40*, 31–43. doi:10.2753/JOA0091-3367400303

Tetlock, P. E. (2002). Social functionalist frameworks for judgment and choice: Intuitive politicians, theologians, and prosecutors. *Psychological Review, 109*(3), 451–471. doi:10.1037/0033-295X.109.3.451

Tetlock, P. E., Visser, P. S., Singh, R., Polifroni, M., Scott, A., Elson, S. B., et al. (2007). People as intuitive prosecutors: The impact of social-control goals on attributions of responsibility. *Journal of Experimental Social Psychology, 43*(2), 195–209. doi:10.1016/j.jesp.2006.02.009

Thompson, C. M., & Dennison, S. (2004). Graphic evidence of violence: The impact on juror decision-making, the influence of judicial instructions and the effect of juror biases. *Psychiatry, Psychology and Law, 11*(2), 323–337. doi:10.1375/1321871042707188

Tiedens, L. Z., & Linton, S. (2001). Judgment under emotional certainty and uncertainty: The effects of specific emotions on information processing. *Journal of Personality and Social Psychology, 81*(6), 973–988. doi:10.1037/0022-3514.81.6.973

Tsoudis, O., & Smith-Lovin, L. (1998). How bad was it? The effects of victim and perpetrator emotion on responses to criminal court vignettes. *Social Forces, 77*, 695–722. doi:10.2307/3005544

United States v. Dougherty. (1972). 473 F. 2d 1113.

Watson, D., & Clark, L. A. (1994). *The PANAS-X: Manual for the positive and negative affect schedule-Expanded Form*. Iowa City: University of Iowa.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1074. doi:10.1037/0022-3514.54.6.1063

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review, 101*, 34–52.

Wevodau, A. L., Cramer, R. J., Kehn, A., & Clark, J. W. (2014). Why the impact? Negative affective change as a mediator of the effects of victim impact statements. *Journal of Interpersonal Violence, 29*, 2884–2903. doi:10.1177/0886260514527170

Whalen, D. H., & Blanchard, F. A. (1982). Effects of photographic evidence on mock juror judgment. *Journal of Applied Social Psychology, 12*, 30–41. doi:10.1111/j.1559-1816.1982.tb00846.x

Wilson, J. R., & Bornstein, B. H. (1998). Methodological considerations in pretrial publicity research: Is the medium the message? *Law and Human Behavior, 22*, 585–597. doi:10.1023/A:1025743614951

Wright, E. T. (1987). *How to use courtroom drama to win cases*. Englewood Cliffs, N.J: Prentice Hall.

Scientific Evidence and Juror Decision Making: Theory, Empirical Research, and Future Directions

Dennis J. Devine and Shelbi Macken

There is something very appealing about using science to help resolve legal questions in the courtroom. Unlike eyewitness testimony, now widely recognized as fallible, "scientific" evidence has an aura of precision, reliability, accuracy, and objectivity. It offers the promise of helping legal fact finders make better decisions and thus warrants the attention of legal practitioners, lawmakers, and jury scholars. This chapter is devoted to reviewing theory and empirical research on the impact of scientific evidence on jurors.

But what exactly is "scientific" evidence? This question has garnered surprisingly little attention from scholars. From a legal perspective, scientific evidence could be viewed as evidence "beyond the ken" of the average layperson that is offered by an expert witness who possesses "scientific, technical, or other specialized knowledge" (adapted from Rule 702, Federal Rules of Evidence, 2015). This definition is not particularly helpful, though, because it fails to capture the essence of scientific evidence or adequately distinguish it from other forms of complex evidence offered by an expert witness. From an academic perspective, researchers have long been interested in evidence that is "probabilistic," "complex," or "statistical" (e.g., Goodman 1992; Kaye and Koehler 1991; Thompson 1989; Wells 1992), but one is hard-pressed to find an explicit definition of scientific evidence in the scholarly literature. So, for the purposes of this chapter, *scientific evidence* is considered to be any form of evidence arising from systematic application of the scientific method that yields diagnostic conclusions about the identity or characteristics of a target individual. Hallmarks of the scientific method include a grounding in theory, the use of standardized analytical procedures, and precise measurement. Diagnostic conclusions are judgments intended to be definitive with regard to whether the target individual possesses some focal property. Of note, this

D.J. Devine (✉) · S. Macken
Indiana University-Purdue University Indianapolis, Indianapolis, USA
e-mail: ddevine@iupui.edu

B.H. Bornstein and M.K. Miller (eds.), *Advances in Psychology and Law*,
Advances in Psychology and Law 2, DOI 10.1007/978-3-319-43083-6_4

definition excludes "framework" research on syndromes or behavioral profiles that might be introduced in court to provide background information for jurors.

Many forms of scientific evidence have emerged over the last century and they can be clustered into several major groups that involve: (1) comparing physical impressions made by various objects (e.g., shoeprints, tire tracks, firearms, fingernails, and teeth), (2) comparing structures or products of the human body (e.g., blood type, DNA, fingerprints, hair, saliva, and semen), (3) assessing the concentration of different substances within the body (e.g., alcohol or drugs), and (4) assessing brain structure and/or cognitive functioning via sophisticated technologies such as *computerized axial tomography* (CAT or CT), *magnetic resonance imaging* (MRI), *positron emission tomography* (PET), and *functional magnetic resonance imaging* (fMRI). These neuroimaging techniques yield visual displays of brain structure or some aspect of brain functioning such as blood flow or oxygen utilization (Khoshbin and Khoshbin 2007).

Scientific evidence in its various forms appears regularly in the courtroom and trial court judges are tasked with determining its admissibility (Chin 2014). In essence, scientific evidence must be deemed both relevant and reliable to be admissible. Relevance is defined as making some fact more or less likely by its existence; reliability corresponds to the trustworthiness of the underlying theory and technique that produced the evidence. Jurisdictions in the United States differ in terms of the criteria they employ with regard to reliability. The oldest criterion—general acceptance in the relevant scientific community—dates back to an influential federal case in the 1920s (*Frye v. United States*, 1923) and remains the standard in a minority of states. However, the federal courts and many states adopted an expanded set of criteria in the 1990s in the wake of a trio of U.S. Supreme Court cases (*Daubert v. Merrell Dow Pharmaceuticals, Inc.* 1993; *General Electric Co. v. Joiner* 1997; *Kumho Tire, Ltd., v. Carmichael* 1999). Judges in these *Daubert* jurisdictions generally make admissibility decisions based on set of criteria that include whether the theory and technique underlying the scientific evidence: (1) has been tested, or at least could be, (2) has a known error rate, (3) has been peer-reviewed and published, (4) has controls and operating standards, and (5) has been generally accepted by the relevant scientific community. Even if determined to be relevant and reliable, scientific evidence can still be excluded if the judge views it as unduly prejudicial, confusing, misleading, or wasteful.

When admitted, scientific evidence is usually introduced by an expert witness via direct examination. On the stand, expert witnesses typically describe the theoretical basis of their work, operational procedures, quantitative results, and substantive conclusions. Experts are often prohibited from offering an "ultimate" conclusion regarding whether a defendant has a mental state or condition that constitutes an element of the crime or a defense.

The various types of scientific evidence can also be divided into two categories based on their purpose. The first category consists of procedures (or "tests") primarily intended to *identify* the perpetrator of some wrongful action. Tests in this category generally involve the systematic comparison of a sample from a known

source (e.g., the defendant or the defendant's weapon) with a sample from an unknown source obtained from a particular location (e.g., crime scene). Common exemplars include the comparison of fingerprints, blood type, hair, bite marks, and DNA. The second category consists of procedures used to assess or *classify* a target individual in terms of some physiological condition (e.g., intoxication or brain damage) or psychological characteristic (e.g., deception). Common exemplars include polygraph tests, blood alcohol tests, and neuroimaging techniques such as CT, PET, and fMRI. Structural damage or aberrant brain functioning indicated by neuroimaging techniques can serve as the basis for arguing the defendant lacked the necessary criminal mindset (i.e., *mens rea*) to be found guilty, or should not be held fully responsible for his or her actions.

Anticipated by a few studies on the polygraph in the 1970s, regular research on scientific evidence began in the late 1980s with work on blood type (serology) tests. Research on serology and polygraph tests continued into the 1990s but then tapered off when national attention was drawn to DNA evidence via its salient role in two trials involving O.J. Simpson. The study of DNA evidence began in the mid-1990s and continues today at a steady rate, making DNA the most-researched form of scientific evidence thus far. In the last five years, research on neuroimaging tests has been spurred by concerns about its potential for undue influence on jurors and become something of a hot topic. The trickle of research on scientific evidence that began in the 1970s has now basically turned into a bubbling brook.

This structure of this chapter is as follows. We first review theoretical models that provide a basis for understanding how jurors use scientific evidence, and then review empirical research on the effects of scientific evidence. In the final section of the chapter, we discuss major empirical findings, offer an integrative conceptual framework, and suggest a number of future research directions. In keeping with our definition of scientific evidence, we limit our attention to forms of scientific evidence intended to conclusively identify or classify defendants, as opposed to expert testimony on behavioral science issues offered to provide a framework for jurors, or characteristics of the expert witnesses themselves.

Theoretical Approaches to Understanding the Impact of Scientific Evidence

Almost every theory, model, or framework in psychology could be applied in some fashion to explain how jurors use scientific evidence. In this section, we review two types of theory with the greatest potential to inform our understanding of how jurors process scientific evidence: (1) general theories of cognition, and (2) theories of juror decision making.

General Theories of Cognition

Four general theories of cognition relevant to juror decision making have considerable potential to inform our understanding of jurors' thought processes even though they are not specifically focused on legal decision making—dual-processing models, the Bayesian model, the general linear model, and the parallel constraint satisfaction model.

Dual-Processing Theory. "Dual-processing" refers to a class of models sharing the fundamental proposition that there are two primary modes of human thought (e.g., Chaiken 1980; Petty and Cacioppo 1986a). The two modes represent different levels of conscious scrutiny that may be used in the processing of persuasive communication. In one of the best-known instantiations of dual-processing theory—the elaboration likelihood model—the two modes are referred to as the *central route* and the *peripheral route* (Petty and Cacioppo 1986b). The central route is resource-intensive and effortful, requiring considerable attention. The primary benefit of central route processing is that reason and logic can be brought to bear to make inferences and draw conclusions. In contrast, the peripheral route is much quicker and requires little if any effort. This form of processing is driven by heuristics, or cognitive rules-of-thumb. An underlying premise of dual-processing theory is that people generally desire to limit the amount of effortful central processing they do, but vary in their need for cognition—that is, their willingness to engage in it (Cacioppo et al. 1996). In addition to this dispositional tendency, contextual aspects of decision tasks are presumed to affect the degree to which people will undertake effortful thought. Specifically, individuals are expected to be less likely to engage in central route processing when: (1) the task at hand is unfamiliar or complex, (2) time pressure is great; or (3) the decision maker is tired, confused, or distracted.

Given that scientific evidence is often novel and complex, dual-processing theory suggests jurors will often rely on heuristics, rules of thumb, or other cognitive shortcuts to assess scientific evidence. For example, jurors may draw inferences from the physical appearance or professional credentials of the expert witness. However, dual-processing theory also implies that some jurors—particularly those with high need for cognition—will think more deeply about scientific evidence than others. The complexity of the scientific evidence will also affect the number of jurors who process it using the central route, with progressively less central route processing as scientific evidence becomes more difficult to comprehend. Ultimately, dual-processing models suggest many jurors will use peripheral route processing for scientific evidence, but the extent of peripheral route processing will differ across jurors (and even within jurors over the course of a trial).

General Linear Model. Another approach to juror decision making involves the well-known general linear model (Anderson 1981; Einhorn et al. 1979; Hastie 1993; Pennington and Hastie 1981). The general linear model assumes information is parsed into pieces (or "cues"), weights are assigned to cues to reflect their perceived importance to the task, and the weighted values are summed. At trial,

evidence is presumably decomposed into cues that are weighted by their perceived probative value and then summed to generate an overall number reflecting the defendant's perceived likelihood of guilt. Jurors are assumed to possess the functional equivalent of a mental "meter" that provides a continuous "read-out" of guilt probability. Jurors' processing ceases when there is no further evidence to consider or some critical event causes the meter to move so far in one direction that it "sticks" or "freezes." This results in the juror choosing the corresponding verdict. If processing does not cease in this manner, jurors will ultimately choose a verdict based on whether their final level of perceived culpability surpasses the level of subjective certainty required for a judgment of culpability as conveyed by their legal instructions.

The general linear model also has implications for how jurors evaluate and use scientific evidence. In particular, it suggests jurors can (and will) parse scientific evidence and assign weights to the pieces that reflect their perceived probative value. Scientific evidence with the capability to identify or classify defendants should in theory receive relatively large weights, and definitive test results presented in unequivocal terms might represent a critical event that causes jurors to cease processing and choose a verdict. This could result in subsequent "strong" evidence being discounted or dismissed. Another intriguing implication of evidence processing according to the general linear model is that scientific test results should have strong probative value regardless of the other evidence in the case or how well the test results fit with it.

Bayesian Model. The Bayesian model is similar to the general linear model in being mathematically elegant and based on the notion that jurors can perform some algebraic-like operations in their heads (Edwards et al. 1963; Hastie 1993; Pennington and Hastie 1981). The model features four basic elements: (1) an initial (prior) probability of guilt or culpability, (2) "new" information (i.e., evidence) received after formation of the initial probability, (3) a likelihood ratio that represents the probative value of new information, and (4) an updated (posterior) probability of guilt. In essence, the Bayesian model assumes jurors begin a trial with an initial probability of guilt that is updated as evidence comes in during the trial. Updating involves multiplying the initial (or current) probability by the ratio of two conditional probabilities: the probability of observing the focal evidence given the defendant is truly *culpable*, and the probability of observing the evidence given the defendant is truly *innocent*. This likelihood ratio essentially captures the probative value of the evidence—the higher the value, the more incriminating the evidence is perceived to be. The initial probability is viewed as a function of jurors' existing knowledge and beliefs about the way the world works; this "baseline" probability is then modified in an iterative fashion as evidence is presented. At the trial's conclusion, the final probability of guilt is compared to a threshold of subjective certainty and an appropriate verdict is chosen.

The Bayesian model has several desirable properties that have made it a favorite of jury scholars (Hastie 1993). In particular, it is formal, parsimonious, and prescriptive. Accordingly, the Bayesian model has often been used as a benchmark for assessing how logically and coherently jurors use scientific evidence.

Unfortunately, the model also has two major drawbacks—it invokes complex cognitive operations that do not align well with jurors' descriptions of their thought processes, and comparisons of juror judgments and model-based predictions are often quite discrepant (Kaye et al. 2007). As such, the Bayesian model has considerable normative value but limited descriptive value.

Parallel Constraint Satisfaction. The emergence of the connectionist architecture in the 1980s was a major advance in the study of human cognition, and connectionist models have been used to account for a variety of cognitive phenomena such as impression formation, social categorization, and causal attribution (Read and Miller 1998; Tryon 2012). In essence, information in a connectionist architecture is represented by a set of nodes that are linked in an associative network via connections that can be either excitatory or inhibitory. Information that surpasses a threshold level of activation is treated as being "true." During decision-making, activation spreads out along connections in the network in a rapid, effortless, and unconscious fashion. New information is incorporated into the network based on its correspondence with existing associations, but connections among existing information cues can and do change strength (and even direction) to accommodate new information. In essence, the direction and strength of network connections are continually adjusted to maximize the meaningfulness (or coherence) of the information network. This processing goal has been referred to as parallel constraint satisfaction (Holyoak and Thagard 1989). In the context of a particular decision task, strong positive connections develop between information cues that point towards the same judgment or conclusion, whereas negative connections emerge among information cues that support different judgments. This Gestaltian process whereby the parts are made to fit into a whole allows the decision maker to eventually arrive at a clear preference in complex tasks that involve large amounts of conflicting and/or ambiguous information.

In a legal context, Dan Simon and his colleagues used a model featuring parallel constraint satisfaction to explain how jurors process trial evidence (Simon 2012; Simon et al. 2004). One consequence of parallel constraint satisfaction is *coherence-based reasoning*. At trial, strong positive associations will form among information cues in two separate clusters—one supporting culpability and one supporting acquittal. Over the course of the trial, within-cluster connections and between-cluster connection among the evidentiary cues will be adjusted to maximize their coherence, with one verdict option eventually emerging as the leading verdict. New evidence consistent with the leading verdict is readily assimilated with other evidentiary cues that support that verdict, whereas new evidence inconsistent with the leading verdict may be suppressed (discounted) or reinterpreted to reduce cognitive dissonance. Parallel constraint satisfaction eventually yields a verdict preference based on a tightly connected cluster of supporting evidentiary items.

The parallel constraint satisfaction model has several key implications for the study of juror decision making. First, jurors do not (and cannot) evaluate bits of evidence in isolation. Second, evidentiary cues do not have a fixed probative value —evidence evaluation is a dynamic and continuous process. Third, when jurors are asked to indicate the probative value of any particular piece of evidence (e.g., a

DNA test result), their response will represent only the final stage of this fluid, subconscious process.

Theories of Juror Decision Making

Three additional theories are highly relevant to understanding how jurors' process scientific evidence in that they specifically address how laypersons make sense of the evidence encountered at trial. These three theories are the story model, the director's cut model, and exemplar cueing.

Story Model. In contrast to the formal, quantitative nature of the Bayesian and general linear models, the story model presumes jurors reach their decisions by actively fashioning trial evidence into a narrative framework—a *story*. Stories are cognitive structures that represent linked sequences of events featuring actors who behave based on conscious motives or intentions, and these behaviors in turn have real-world consequences (Bennett 1978; Bennett and Feldman 1981; Hastie 1993; Pennington and Hastie 1993). According to the story model, jurors do not just passively absorb the evidence at trial but rather actively sift through it, focusing on some elements and ignoring others, in order to create a compelling narrative explanation that accounts for events leading up to the trial.

So what makes a story compelling? Jurors presumably evaluate potential stories using three certainty principles: coverage, coherence, and uniqueness (Pennington and Hastie 1993). Coverage and coherence are used to assess the plausibility of a story, whereas uniqueness affects how much confidence a juror will have in a story. *Coverage* refers to how well a story can account for, or explain, the various pieces of evidence presented at trial. *Coherence* is a multidimensional construct that captures the extent to which a story makes sense. Specifically, it refers to the degree to which a story is internally consistent, plausible, and complete. Stories are most consistent when they avoid internal contradictions, most plausible when they accord with a juror's existing knowledge and beliefs about the world, and most complete when they contain all the elements they are expected to have (e.g., motives for the various actors). The third criterion, *uniqueness*, is inversely related to the number of viable stories that exist. In other words, a juror's confidence in any given story is higher when there is only one plausible story that can explain trial events (to the juror's way of thinking), and lower to the extent that more "rival" stories exist. In essence, a story is most convincing or compelling when it is high on coverage, high on coherence, and no plausible rival story can be formulated. Once a juror has settled on a story, its elements are mapped to the legal requirements of the available verdict options and a verdict preference is identified.

The story model has two important implications for jurors' processing of scientific evidence. First, the probative value of a scientific test result will greatly depend on how well it fits into a larger narrative. For example, an incriminating test result might be ignored if there is no other evidence of culpability, or an exonerating test result might be discounted if the rest of the evidence converges neatly into

a story of guilt. Second, individual differences will have considerable influence on how jurors evaluate scientific evidence via their use of existing knowledge and beliefs to fill in the gaps in their emerging stories.

The Director's Cut Model. Devine (2012) offered a multilevel theory of jury decision making that provides a unified account of how both jurors *and juries* make decisions. The multilevel theory consists of two embedded models operating at different levels: the "director's cut" model addresses *juror* thought processes, whereas the "story sampling" model deals with *jury* decision making. Like the story model, the director's cut model assumes individual jurors seek to create a compelling narrative from the evidence that accounts for the facts of the case as they understand them, but it goes further. Based on a systematic review of the empirical literature on jury decision making, the director's cut model also identifies specific variables that affect the construction of an initial mental representation of a trial: pretrial publicity, opening statements of the attorneys, the nature of the case, and characteristics of the defendant and jurors. Incoming trial evidence is assimilated into the initial mental representation and the updated trial representation eventually forms the basis for identifying a *p story* offered by the prosecution or plaintiff, and one or more alternative *d stories* offered (or implied) by the defense.

The director's cut model also diverges from the story model in specifying a different mechanism for evaluating story options—mental models. Mental models are situation-specific cognitive representations constructed on the fly to help individuals make sense of present circumstances, make decisions, and forecast future events (Johnson-Laird 1983). Although largely neglected by jury scholars, they are well established in cognitive psychology and have been used to explain human thought processes in a variety of contexts (Rouse, and Morris 1986; Wyer 2007). A key feature of mental models is that they exist in different forms or *modalities*, including visual and propositional. Visual mental models represent pictorial representations of some object or scene; propositional mental models represent verbal statements in a syllogistic form. Another key feature of mental models is their potential to help decision makers evaluate the likelihood of different outcomes via mental simulation. For instance, a juror could "run" a visual model by applying beliefs about the physical world to a cognitive scenario in order to forecast how quickly or easily a burglar might gain entry into a home.

The director's cut model has several implications for the processing of scientific evidence. First, jurors might use mental models to visualize the handling of scientific evidence (e.g., collection, storage, and testing) and make inferences about the integrity of that process (e.g., the likelihood that a sample was compromised). Second, jurors might use mental models to deal with the statistical aspects of scientific evidence. In particular, test results or expert opinions involving probabilities could be translated into a set of premises that are then used to draw conclusions about the likelihood of error. Some premises in these propositional mental models might be generated automatically based on a juror's beliefs about the competence and trustworthiness of different actors within the criminal justice system (e.g., police, eyewitnesses, crime lab personnel, expert witnesses). Third, mental models will serve as the primary mechanism for evaluating the plausibility

of different stories and choosing among them. Thus, the director's cut model suggests individual differences, mental models, and stories will all play an important role in jurors' processing of scientific evidence.

Exemplar Cueing Theory. Koehler's exemplar cueing theory is unique in being the only one crafted specifically to address how jurors respond to scientific evidence (Koehler 2001; Koehler and Macchi 2004). The theory essentially predicts that the manner in which statistical evidence is presented will influence its impact. Drawing on the well-established availability heuristic identified by Tversky and Kahneman (1973), Koehler (2001) hypothesized some verbal statements used to convey scientific test results will have more impact than others even though they express propositions that are mathematically equivalent. Specifically, the theory predicts jurors will be less likely to convict in trials that feature incriminating scientific evidence to the extent that it is easier to imagine someone other than the defendant who would coincidentally match the crime-scene sample. The likelihood of a coincidental match is commonly referred to as the *random match probability* (RMP).

In the theory's original formulation, Koehler (2001) identified two variables that influence the weight jurors attach to a DNA "match" result: target and frame. *Target* refers to the size of the unit attached to the possibility of error (i.e., a single individual or multiple persons). *Frame* (subsequently relabeled *form*) pertains to the statistical format used to convey the likelihood of a coincidental match (e.g., a frequency, a probability, or an odds ratio). The most intriguing aspect of the theory is a predicted interaction between target and frame. Given a fixed likelihood of coincidental match, jurors are predicted to acquit more often when the statistical value is attached to a group of persons *and* the quantified value is presented in a frequency frame. So, for example, jurors would be much more likely to convict a defendant when told of a 0.0003 probability that a random member of the population would coincidentally match the crime-scene sample than when informed that three of every ten thousand individuals from the population would be expected to match the crime-scene sample by chance.

Initial empirical support for exemplar cueing theory was obtained in a series of experiments (Koehler 2001), but the predicted framing effects basically vanished at very low (i.e., more realistic) RMP levels (e.g., one in one billion as opposed to one in one thousand). Accordingly, Koehler and Macchi (2004) tweaked the original theory by adding a third parameter—*reference class size*. This contextual variable corresponds to the number of persons in the relevant population (e.g., 500 vs. 5 million). Reference class was predicted to interact with target such that providing error rate information referring to multiple persons *and* a relatively large referent class would make it very easy for jurors to imagine coincidental exemplar matches and thus especially likely to acquit. Two empirical experiments have provided support for the revised theory (Koehler and Macchi 2004).

Toward an Integration.

At a glance, the various theories relevant to jurors' processing of scientific evidence may seem to compete, but they can also be viewed as complementary perspectives that differ in terms of their level of analysis. At the most micro level, the connectionist architecture provides an "engine" for the cognitive system and parallel constraint satisfaction offers a criterion for its information processing. Parallel constraint satisfaction addresses how evidence is interpreted dynamically in light of existing "facts," what evidence will be viewed as true, and what evidence will emerge as most salient. Evidentiary cues that achieve a threshold level of activation are more likely to be available for more conscious and/or rational thought processes. Thus, the micro-level process can be viewed as supplying the informational building blocks for mid-level theories such as the story model and the director's cut model that focus on the creation of a compelling narrative that ties together the disparate evidentiary cues. These theories highlight the importance of incorporating scientific evidence into a narrative that takes into account other aspects of the evidence and is congruent with a juror's knowledge and beliefs about the world. Mental models in turn provide a mechanism for evaluating alternative stories, and are particularly well suited for drawing conclusions about different aspects of the scientific evidence. Finally, macro-level theories such as the Bayesian model or the general linear model provide a computational basis for generating and testing predictions based on the lower-level models. Having considered what theory says about jurors' use of scientific evidence, we now turn to what the data can add to our understanding.

Empirical Research on the Impact of Scientific Evidence on Jurors

In this section we review empirical research that addresses three major questions: (1) How well do jurors use scientific evidence? (2) What characteristics of scientific evidence influence juror decision making? and (3) What contextual and dispositional variables affect the impact of scientific evidence?

How Well Do Jurors Use Scientific Evidence?

This is one of the oldest and most frequently examined questions in the empirical literature. It is explicitly evaluative in nature and requires some standard of comparison to be identified. Most often, the decisions of mock jurors have been compared to the predictions of the Bayesian model. In other words, how rationally do jurors use scientific evidence as defined by the dictates of the Bayesian model?

To begin with, some research indicates mock jurors have trouble comprehending scientific evidence (Hans et al. 2011). Many participants have trouble answering basic factual questions about the scientific evidence when knowledge checks are employed. For instance, 34 % of student participants could not calculate the number of persons in the target population who would match the defendant's blood type after an incidence rate for that blood type was provided, and 87 % of venirepersons were unable to correctly paraphrase unopposed expert testimony (Goodman 1992). Further, about 62 % of participants (including judges, judicial officials, and venirepersons) were unable to calculate the number of coincidental matches when supplied with the needed information in another study (McQuiston-Surrett and Saks 2009). Finally, in a study involving venirepersons, mean accuracy for a set of true–false comprehension questions averaged about 70 % prior to deliberation—only 20 % better than chance (Hans et al. 2011). Deliberation does not seem to help much either. In the study by Hans et al. (2011), comprehension improved by only 3 % after deliberation.

Fallacious Conclusions. One comprehension-related issue examined with some regularity is the extent to which mock jurors commit specific kinds of logical reasoning errors. The classic experimental paradigm involves first providing study participants with some basic information about a case and asking them to estimate the probability that a defendant is guilty of the crime. After this, participants are told of a "match" between the defendant and a crime-scene sample for some type of scientific test along with a numeric estimate of the likelihood of a "random" match. Finally, participants are again asked to indicate how likely it is that the defendant is guilty of the crime. Two fallacious conclusions can be inferred from the responses. The *prosecutor's fallacy* occurs when the final probability of the defendant's guilt is equal to one minus the random match probability (i.e., 1-RMP). For instance, the prosecutor's fallacy would be inferred if a study participant reported a 99 % chance of the defendant being guilty after being told of a 1 % chance of a random match. Conversely, the *defense attorney's fallacy* occurs when scientific evidence is treated as essentially irrelevant. This type of error is usually inferred when the final probability of guilt is equal to the initial probability or, alternatively, when it is reported to be equal to one divided by the number of potential coincidental "matches" in the relevant population (i.e., 1/potential matches). Thus, reporting the probability of guilt to be 20 % after learning that five persons were expected to coincidentally match the crime-scene sample would be interpreted as an instance of the defense attorney's fallacy.

In general, mock jurors appear to succumb to these "attorney" fallacies at worrisome levels. In the first research to examine them, Thompson and Schumann (1987) employed a robbery case in which scientific evidence was presented in the form of a match involving hair samples. Study participants read a brief synopsis of the non-scientific evidence and provided an initial probability of guilt, then read the summary of the forensic expert's testimony and provided a final guilt probability. About a quarter of the participants gave final probability-of-guilt estimates consistent with one or both of the two fallacies—13 % for the prosecutor's fallacy and another 13 % for the defense attorney's fallacy. A second experiment involved a

murder case and an initial probability of guilt was provided in the case summary by the investigating detectives. Participants were then exposed to testimony about a blood type match, presented with arguments representing each fallacy, and asked to respond to several questions about each argument. Only 3 % of participants gave responses judged to be consistent with the prosecutor's fallacy, but a whopping 66 % appeared to fall prey to the defense attorney's fallacy in that their guilt probability did not change after receiving the scientific evidence. Less than a quarter of participants (22 %) recognized that both arguments were incorrect.

Subsequent research produced further indication of these reasoning fallacies. Goodman (1992) observed prevalence rates for the prosecutor's fallacy of about 2 % in an experiment involving students and 8 % in a second experiment involving Washington (Seattle) venirepersons. Another study revealed about 4 % of student participants gave responses consistent with the prosecutor's fallacy (Smith et al. 1996). Nance and Morris (2002, 2005) estimated prevalence rates of 6 %, 3 %, and less than 2 % for the prosecutor's fallacy in three large experimental studies involving Illinois venirepersons. They also estimated a 7 % prevalence rate for the defense attorney's fallacy in one of the studies reported in 2002. Finally, consistent with the defense attorney's fallacy, 49 % of participants in an online study agreed with a statement that the scientific evidence (either a DNA or shoeprint match) had little probative value because it could have come from someone other than the defendant (Thompson and Newman 2015).

It is possible the occurrence of fallacious reasoning could be substantially mitigated by the act of deliberating but this does not appear to be the case. Dann et al. (2004) posed true–false questions about mitochondrial DNA (mtDNA) evidence to their Delaware venireperson participants. Before deliberation, 48 % of participants indicated that an item reflecting the prosecutor's fallacy was true (i.e., there is about a 1 % chance that someone else beside the defendant committed the crime), and 39 % did the same for an item consistent with the defense attorney's fallacy (i.e., the mtDNA evidence is completely irrelevant because persons other than the defendant could have been the source). After deliberation, endorsement rates *increased* to 60 % for the item reflecting the prosecutor's fallacy and 40 % for the item reflecting the defense attorney's fallacy. Another study found 20 % of respondents reported a guilt probability consistent with the prosecutor's fallacy before deliberation and 14 % still did so afterwards (Thompson et al. 2013). Deliberation thus does not necessarily reduce the incidence of fallacious thinking.

Taken together, many study participants clearly have difficulty drawing appropriate conclusions from scientific evidence. One potential cause of the variation in the observed prevalence rates for the two "attorney" fallacies is the methodology employed to measure them. In some studies, fallacious thinking has been inferred from change between jurors' initial probability of guilt and their final probability after exposure to scientific evidence; in other studies, fallacious thinking has been assessed via agreement with statements about the evidence that reflect the two fallacies. In addition to these methodological factors, the prevalence of logical reasoning fallacies may also depend on several substantive variables (e.g., the type of case, the type of scientific test, and jurors' ability to reason quantitatively).

Aggregation of Statistical Probabilities. Another task that has created problems for mock jurors in studies involving scientific evidence is the combining of likelihood information for different types of error. Several studies have provided participants with separate probabilities for the likelihood of a coincidental match and a false positive report (i.e., some form of laboratory error), and their results revealed that decision makers have difficulty combining the values in a mathematically appropriate fashion (Britton 1997; Koehler et al. 1995; Nance and Morris 2005; Schklar and Diamond 1999). For instance, Schklar and Diamond (1999) found that 67 % of their mock jurors incorrectly combined RMP and LER values, with the most common mistake (committed by 44 %) appearing to involve some form of averaging of the two values.

Correspondence Between Jurors' Judgments and Predictions of the Bayesian Model. Another line of investigation has compared mock jurors' actual judgments with their expected judgments as forecasted by a theoretical model. This research addresses whether jurors make decisions in a logical (or rational) manner as defined by the underlying theory. Because of its formal, precise, and parsimonious nature, the Bayesian model has usually been the benchmark theory of choice in this work. Studies examining the correspondence between juror judgments and Bayesian predictions often involve a repeated-measures paradigm in which mock jurors provide an initial judgment regarding the probability of guilt based on little or no evidence, and then rate the probative value of evidence that is subsequently introduced. After this, participants are also asked to provide a final probability that the defendant committed the crime. A model-based predicted probability of guilt can then be generated for each individual by converting their ratings of the probative value of the new evidence into a likelihood ratio and multiplying it by their initial probability of guilt. Model-based predicted probabilities can then be compared with actual (reported) final probabilities of guilt and the degree of discrepancy assessed. The greater the departure, the less "rational" the decision process.

The consistent conclusion from research on this topic is that most mock jurors do not use scientific evidence in a logical fashion—at least not according to the Bayesian model (Hastie 1993; Kaye et al. 2007). Specifically—and in sharp contrast to the fears of many legal practitioners—most studies have found jurors to *underutilize* scientific evidence by not giving it as much weight as they indicate that it deserves (Faigman and Baglioni 1988; Goodman 1992; Martire et al. 2013; Nance and Morris 2002, 2005; Schklar and Diamond 1999; Smith et al. 1996: Thompson and Schumann 1987). In other words, mock jurors often fail to adjust their guilt probabilities as much as they should in light of how probative they report the scientific evidence to be. This line of research thus strongly suggests jurors are not awed by scientific evidence and will not convict simply because it is present. Indeed, it indicates they are somewhat skeptical of it.

One particular departure from Bayesian expectations is even rather counterintuitive. Specifically, Martire et al. (2013) found that an average of 13 % of mock jurors across five study conditions reacted to *incriminating* shoeprint match by reporting a *lower* belief in the defendant's guilt—and a whopping 62 % did so in a sixth condition when the incriminating match was verbally characterized as

providing "weak or limited support" for the defendant's footwear being the source of the marks. The phenomenon was replicated in a second experiment in which 67 % of those who learned of an incriminating shoeprint that provided weak/limited support for a match subsequently reported belief change in the direction of innocence. This paradoxical result occurred for an exculpatory comparison as well, with 31 % of those who learned of weak/limited support for a nonmatch result reporting a *greater* belief in guilt. A follow-up study essentially replicated the incongruous finding for an incriminating comparison involving fingerprints (Martire et al. 2014), although there was less indication of a weak evidence effect when the comparison result was displayed using a table or a visual scale.

Other recent research, however, suggests jurors may not always be bad Bayesian thinkers. Specifically, two studies show that when Bayesian predictions are based on a more sophisticated model and/or involve a more sensitive response scale, participants adhere more closely to Bayesian norms and sometimes give even more weight to scientific evidence than it deserves (Thompson et al. 2013; Thompson and Newman 2015). In the first study (Thompson et al. 2013), a logarithmic response scale was developed for measuring guilt probabilities and used in two experiments that featured a rape case involving incriminating DNA evidence. In both experiments, ratings of guilt likelihood were generally in accord with Bayesian estimates and in some conditions the DNA evidence was *overutilized*. In the second study (Thompson and Newman 2015), a rape case was again used along with the logarithmic response scale and a refined Bayesian model that incorporated probabilities for three kinds of errors. DNA evidence was also compared with shoeprint evidence in this study to assess the impact of scientific test type. Mock jurors generally weighted the DNA evidence in a fashion consistent with the predictions of the refined Bayesian model, but still tended to underutilize shoeprint evidence. One caveat associated with both of these studies, however, is that they used a between-subjects design that only allowed for aggregate comparisons across conditions based on estimated distributions of individual belief change.

Overall, the bulk of the research suggests jurors will often—but not always—underutilize scientific evidence relative to Bayesian norms, but the degree of departure may depend on how belief change is measured and/or the type of scientific evidence involved.

What Can Be Done to Help Jurors Use Scientific Evidence More Rationally? Concerns about jurors' understanding of complex evidence emerged in the 1980s and led to the examination of several decision aids that could be implemented at trial to improve their understanding. These aids include: (1) preinstructing jurors before they hear the evidence, (2) allowing jurors to take notes, (3) allowing jurors to ask questions of the witnesses through the judge, (4) providing notebooks or other summaries of the evidence, and (5) allowing jurors to discuss the evidence prior to the end of the trial. Unfortunately, although jurors report liking these aids and legal professionals are increasingly willing to permit them, there is little indication that they substantially improve jurors' understanding of the evidence (Devine 2012).

Relatively little research has examined the efficacy of interventions targeted specifically at improving jurors' understanding of scientific evidence. Two studies have included an instructional intervention where in the expert witness explained how to use the Bayesian model and/or combine separate error likelihood values, but both studies revealed little or no impact on jurors (Faigman and Baglioni 1988; Schklar and Diamond 1999). Another type of aid that has received some attention is the use of a table or chart to help jurors understand Bayesian calculations. Visual aids have been provided to help jurors formulate odds ratios, convert probabilities to odds ratios (or vice versa), or use the Bayesian equation. Research on the impact of these aids has tended to show a modest reduction in divergence from Bayesian estimates and a small increase in the verdict supported by the test result relative to participants who did not receive the aid (Goodman 1992; Nance and Morris 2002, 2005; Smith et al. 1996). One study also assessed the impact of an expert tutorial on DNA presented in either oral (spoken) or multimedia form (Goodman-Delahunty and Hewson 2010). Exposure to the tutorial increased participants' knowledge scores associated with DNA evidence, but had no clear effect on verdicts (although there was a trend for those with lower comprehension scores to convict more often). There were no differences in outcomes as a function of tutorial format either.

Only one study to our knowledge has compared the efficacy of multiple juror decision aids (Dann et al. 2004). This study featured a robbery trial involving a lesser-known form of scientific evidence—mitochondrial DNA (mtDNA). Study participants were recruited from persons who showed up for jury duty at a Delaware courthouse. Four decision aids were examined: note taking, question asking, a checklist to follow in assessing mtDNA evidence, and a notebook containing the expert's slides and information on mtDNA evidence. The design involved a control condition with no aids and five other conditions that allowed/provided the following: (1) note taking only, (2) note taking and question asking, (3) note taking and the checklist, (4) note taking and the juror notebook, and (5) all four aids. The design allowed assessment of the gain in comprehension associated with the individual aids as well as various sets of aids. When question asking was allowed, mock jurors submitted their queries in writing to an on-site expert, who answered when legally appropriate. Jurors filled out a questionnaire, received the materials associated with their study condition, watched a 70-min videotaped trial, indicated a verdict preference, and then deliberated to a group verdict in eight-person juries.

A number of interesting results emerged from this research. First, participating venirepersons generally felt that all of the decision aids were useful. Second, most participants reported using the aids allowed/provided, with 92 % examining the notebook, 88 % taking some notes, and 86 % at least reviewing the checklist. In contrast, only 22 % of participants reported asking a question. Third, the various aids produced relatively modest gains in juror comprehension. Most notably, there were negligible and nonsignificant differences in comprehension between the various experimental conditions and the control condition both before and after deliberation. However, compared with participants who were only allowed to take notes, comprehension scores were somewhat higher for participants in the conditions that received the juror notebook, the checklist, or both. Fourth, the biggest

gain in comprehension occurred when all four innovations were available. Finally, participants who reported using the notebook and/or checklist had higher comprehension scores than those who did not use those innovations—and this remained so even when education level was controlled.

Collectively, the research reviewed in this section indicates jurors have difficulty understanding and using scientific evidence, but it also hints that using various decision aids—especially in combination—can improve jurors' comprehension and performance.

What Characteristics of Scientific Evidence Influence Juror Decision Making?

Research on the different forms of scientific evidence has addressed several questions: (1) Does the presence/absence of some form of scientific evidence affect jurors? (2) Do characteristics of the test itself make any difference to jurors? and (3) Does the impact of a given test outcome depend on the type of test that produced it? Most of the extant research has examined the first two questions and different forms of scientific evidence have rarely been compared in the same study, so the following review will be organized by test type. In particular, we focus on three kinds of scientific evidence that have accrued a fair amount research: the polygraph, DNA, and neuroscientific techniques.

Polygraph Tests. Lie detection via the polygraph was one of the first forms of scientific evidence to be studied by social scientists. The first published study by Koffler (1957) featured 20 students and a repeated-measures design in which participants initially read a brief case and provided a verdict—with all of them acquitting the defendant. They were then told of an additional piece of information—that the defendant had failed a polygraph test (said to have an 85 % accuracy rate)—and asked if this would change their decision, at which point 40 % switched their preferred verdict to guilty. Participants were then asked what their verdict would have been if the polygraph result was 99.5 % accurate, whereupon 85 % said they would have voted to convict. Twenty years later, Carlson et al. (1977) used a similar design but recruited community members to watch one of eight "live" mock trials in an actual courtroom and had them deliberate after the trial. After obtaining individual verdict preferences, participants were told to imagine that one of several witnesses (varying across trials) had failed a polygraph test and asked if that additional evidence would have changed their decision. Only 19 % of those who initially favored a verdict for the side consistent with the polygraph evidence indicated the new information (i.e., unfavorable polygraph result) would have caused them to switch their verdict.

Research on polygraph evidence since these initial studies has focused on the impact of a favorable/unfavorable polygraph test result using more rigorous between-persons designs. The first such experimental investigation featured 12

deliberating juries and manipulation of the presence/absence of polygraph results, as well as the favorability of the result, via a homicide case summary (Markwart and Lynch 1979). Prior to deliberation, 45 % of those who received no polygraph evidence favored guilt; in contrast, 68 % of participants who received the "unfavorable" (pro-prosecution) polygraph result preferred conviction as well as 38 % of those who received a "favorable" (pro-defendant) result. Although there was some spread in the predeliberation verdict preferences, analysis of the taped group deliberations revealed that most juries spent very little time discussing the polygraph results (about 4 % on average).

Subsequent research yielded similar results (Cavoukian and Heselgrave 1980; Myers and Arbuthnot 1997; Spanos et al. 1992–1993). In general, preference for conviction among those exposed to an incriminating polygraph result tends to be only 2–10 % higher than for those who receive no polygraph evidence (Myers and Arbuthnot 1997; Myers et al. 2003, 2006), and incriminating polygraph evidence produced *fewer* convictions than no polygraph evidence in a recent experiment (McCabe et al. 2011). The type of polygraph test does not appear to affect this conclusion either (Myers and Arbuthnot 1997). Evidence of a "passed" polygraph tends to reduce guilty verdicts but does not eliminate them, with 59 % of those who received a pro-defendant test result still convicting in one study (Myers et al. 2006). Several studies have also replicated the dearth of time spent discussing polygraph results (Spanos et al. 1992–1993), with one study revealing only 2 % of deliberation devoted to them (Myers and Arbuthnot 1997). Finally, participants have rated polygraph evidence as one of the least influential forms of evidence presented (Myers and Arbuthnot 1997; Myers et al. 2003, 2006).

To summarize, research on polygraph evidence suggests it has relatively little impact on jurors. Empirical study of the polygraph has tapered off in the last decade, but this is unfortunate in that there is still relatively little research on the topic, what has been done is now rather dated, and most extant research has focused on the simple presence of a polygraph test result (as opposed to other characteristics of the test). Of particular note, participants in studies featuring polygraph evidence have usually been provided with a quantified "accuracy" rate for the test (usually between 80 and 90 %). The provision of these relatively low accuracy rates could be responsible in part for the lack of observed impact, but it is unclear if these accuracy rates are themselves accurate, and little has been done to examine the impact of accuracy rate (essentially error rate) in a systematic manner. In short, it may be premature to conclude polygraph tests have little influence on juror judgments.

DNA Tests. Many people in the United States were introduced to the legal application of DNA testing in the mid-1990s via its salient role in two highly publicized trials involving O.J. Simpson. The gavel-to-gavel national media coverage of these trials—along with a National Research Council report in 1996 calling for more behavioral research on statistical evidence—marks the point at which research on DNA evidence began to appear regularly.

Much of the early work on DNA evidence was conducted by Jonathan Koehler, who was particularly interested in mathematically equivalent but psychologically

different ways that error likelihood could be expressed. Initial studies focused on one particular type of testing error—the possibility of the defendant's DNA matching the crime-scene sample by chance (i.e., coincidentally) when the defendant was not the true source (i.e., RMP). In a number of early experiments, Koehler (1996, 2001) manipulated the *frame* of the RMP (i.e., the statistical metric) along with its *target* (i.e., number of persons associated with the frame). RMP was framed as either a probability (e.g., 0.0001) or a frequency (e.g., one in one thousand), and associated with either one hypothetical target person or multiple target persons from a surrounding community. The combination of a "frequency" frame and "multi-person" target was theorized to provide the easiest cognitive context for imagining multiple "exemplars" who might coincidentally match the crime-scene DNA sample—and thus yield the lowest perceived probative value of the DNA evidence and the fewest convictions.

Several studies by Koehler and his colleagues provided empirical support for the predicted effects of target and frame (Koehler 1996, 2001; Koehler and Macchi 2004). For example, Koehler (1996) observed a conviction rate of 21 % for participants provided with RMP in a frequency format, in contrast to conviction rates of 35 % for those who received RMP as a likelihood ratio, and 41 % for those given RMP information as a posterior odds ratio. In particular, across a number of experiments involving several case types in which the prosecution's evidence was based largely on DNA evidence, various judgmental outcomes (e.g., source probability, guilt probability, conviction) were most favorable to the prosecution when RMP was expressed as a probability for a single target individual (s/p) and most favorable to the defendant when RMP was expressed as a frequency attached to multiple individuals (m/f). Subsequent research with Illinois venirepersons also revealed that RMP format affects conviction rate (Nance and Morris 2002, 2005). In their 2005 study, Nance and Morris observed fewer guilty verdicts when RMP was conveyed in a frequency format (37 %), compared with a likelihood ratio (46 %) or Bayesian chart format (54 %). However, target and frame effects were found to be much weaker when the absolute value (i.e., level) of random match probability was very low (Koehler 2001) or the size of the referent population was large (Koehler and Macchi 2004). Thus, empirical research aligns with the exemplar cueing theory proposition that jurors are less likely to convict when it is easier to imagine coincidental matches.

One issue examined regularly since the mid-1990s is the amount of influence an incriminating DNA test result has on jurors. Research on this topic involves manipulating the presence/absence of a "matching" DNA test result while holding constant the other evidence associated with the case. The presence of an incriminating DNA test result has generally produced conviction rates 20–50 % higher than when no DNA evidence is offered (Dartnall and Goodman-Delahunty 2006; Golding et al. 2000; Nance and Morris 2002, 2005; Smith and Bull 2014), or when an inconclusive DNA test result is presented (Goodman-Delahunty and Hewson 2010; Thompson et al. 2013). For example, in a massive experimental study, Nance and Morris (2005) gathered data from 1520 venirepersons who showed up for jury duty at an Illinois courthouse. Participants in one comparison condition read the

summary of a rape case with no forensic evidence presented; participants in a second comparison condition read about an incriminating DNA test result along with an acknowledgement of the possibility of error ("mere match"). Those in the other nine conditions read a case summary version that included an incriminating DNA result along with a quantified RMP in some format and either unquantified or quantified information regarding laboratory error rate (LER). Nearly half of the participants (46 %) convicted in the nine experimental conditions that received an incriminating test result plus a quantified RMP, whereas only 25 % of participants convicted in the "mere match" condition and 5 % convicted without any DNA evidence. Similarly, conviction rates in another study were 45–66 % higher in several experimental conditions that received an incriminating test result and a quantified LER compared with a control condition that featured an inconclusive DNA test result (Thompson et al. 2013). In short, incriminating DNA test results have a strong, clear effect on mock juror verdict preferences.

Some field research corroborates the intuitive conclusion that incriminating DNA evidence has a strong effect on actual juries. Briody (2004) analyzed 150 homicide cases in Queensland, Australia, half of which involved DNA evidence and half of which did not. The presence of DNA evidence was the best single predictor of conviction in the 92 cases that went to jury trial ($r = 0.41$), being more strongly correlated with jury verdict than either a confession by the defendant ($r = 0.25$) or the presence of incriminating fingerprints ($r = 0.20$). Further, in a multivariate analysis with four other predictors, the odds of a jury convicting were 23 times greater when incriminating DNA evidence was presented. Finally, the authors also used a logistic equation to estimate the conditional probability of conviction in different case circumstances. For one common scenario (i.e., a male defendant, no fingerprint evidence, and the defendant refused a police interview), the estimated probability of conviction was 80 % when an incriminating DNA test was presented but only 15 % in the absence of incriminating DNA evidence. Accordingly, this field research is consistent with the notion that DNA evidence has a strong impact on real juries.

Although the focus has been on incriminating DNA evidence, researchers have also examined the impact of a negative or exculpatory DNA test result (i.e., a "non-match"). In theory, this should be highly beneficial to the defense, but exculpatory DNA evidence does not invariably result in acquittal (Ask et al. 2011; Lieberman et al. 2008; Pozzulo et al. 2009; Simon et al. 2004). A nonmatching DNA test result clearly lowers conviction rates but does not reduce them to zero—sometimes not even close. For example, when the DNA expert in one study opined that the test result was "conclusive beyond a reasonable standard of scientific certainty," an incriminating DNA match yielded a 100 % conviction rate in both a rape case and a murder case. However, replacing the incriminating test result with an exonerating (nonmatch) result still yielded conviction rates of 15 % in the rape case and 36 % in the murder case (Lieberman et al. 2008). In another study, a DNA test result was reported to be either "consistent" or "inconsistent" with a sample provided by the defendant (Pozzulo et al. 2009). Mock jurors who learned of the consistent (i.e., incriminating) DNA test convicted 76 % of the time, but those who

received the inconsistent (i.e., exonerating) result still returned guilty verdicts 25 % of the time. In yet another study, Simon et al. (2004) used a theft case featuring six non-DNA evidence items and varied whether the DNA test incriminated or exonerated the defendant. The non-DNA evidence was evenly balanced, with three items suggesting guilt and three suggesting innocence. Mock jurors who received the incriminating DNA test result convicted 70 % of the time—but so did 32 % of those who got the exonerating test result. Jurors are thus clearly sensitive to the presence of an exculpatory DNA test result, but do not necessarily treat it as conclusive even when the other evidence is equivocal.

In addition to the presence/absence of DNA test results, researchers have examined the *provision* of a quantified random match probability (RMP) as well as the *magnitude* of its value (Britton 1997; Koehler 1996, 2001; Koehler et al. 1995; Koehler and Macchi 2004; Nance and Morris 2005; Schklar and Diamond 1999; Thompson and Newman 2015). Generally speaking, providing a quantified RMP value (as opposed to a "match" result with no acknowledgement of RMP) has had modest and inconsistent effects on probability of guilt and verdict preferences. In one of the first studies to examine RMP parameters, the presence of a quantified RMP value yielded 26–32 % more convictions than RMP-absent conditions in an experiment involving university mock jurors, and 19–36 % more convictions in a replication study involving Texas venirepersons (Koehler et al. 1995). Guilty verdicts were also 21 % more frequent across several conditions in which RMP was reported to be "1 in 40,000" than the RMP-absent condition in a large experiment featuring 1520 Illinois venirepersons (Nance and Morris 2005). Other studies, however, have produced little or no difference in conviction rate as a function of whether RMP was explicitly provided (Britton 1997; Nance and Morris 2002; Schklar and Diamond 1999). Further, RMP magnitude (or a closely related value, the frequency of a matching DNA profile in the population) has generally had small and nonsignificant effects on probability of guilt and verdicts (Britton 1997; Koehler 1996, 2001; Koehler and Macchi 2004). Participants in another study were only sensitive to variation in the absolute value of the RMP when expressed as a frequency as opposed to a likelihood ratio or its verbal equivalent (Thompson and Newman 2015). Thus, research on RMP provision and magnitude does not suggest these test characteristics have strong or consistent effects on jurors in cases centered on DNA evidence.

Research on the influence of LER has generally shown it to have modest impact as well, although results have been mixed (Britton 1997; Koehler et al. 1995; Lieberman et al. 2008; Nance and Morris 2005; Schklar and Diamond 1999; Scurich 2015; Scurich and John 2013; Thompson et al. 2013). Most of the research shows mock jurors to be relatively insensitive to LER provision and/or magnitude. In the first published research on LER, both university students and Texas venirepersons were unaffected by the provision of a quantified LER or its magnitude when provided (Koehler et al. 1995). Similar results were observed in another study in which manipulating LER provision and magnitude produced no clear pattern of differences in conviction rate across four case types (Britton 1997). Providing a quantified LER value further had no significant effects on probability of

guilt or verdict preference in a large study of Illinois venirepersons (Nance and Morris 2005). There was also no main effect of testing lab reliability on conviction rate in another experiment even though the "unreliable" lab was affiliated with the police, run by a former police officer, unaccredited, and the DNA expert's testimony made no mention of using control samples or proficiency testing (Lieberman et al. 2008). Finally, Scurich and John (2013) observed only a small and non-significant 5 % difference in conviction rates when LER was reported to be "1 in 10" as opposed to "1 in 1000."

There is however some empirical support for the impact of LER. In an early study, participants presented with an RMP of "1 in 1 billion" convicted about 12 % more often when LER was quantified at "2 in 100" than when LER was not provided (Schklar and Diamond 1999). Thompson and his colleagues (Thompson et al. 2013) varied the provision and magnitude of *false report probability* (essentially LER) in two experiments using reported levels as low as "zero" (Study 1) or "impossible" (Study 2), up to "1 in 100" (both studies). The conviction rate was 11 % higher in Study 1 when the false report probability was given as "1 in 10,000" rather than "1 in 100," and 15–19 % higher in Study 2 when a quantified LER value was provided as opposed to when participants were told a false report was "impossible." In both studies, there were also substantially more guilty verdicts for the quantified LER conditions than the control condition featuring inconclusive DNA test results. A third study by Scurich (2015) also produced a significant difference in conviction rates as a function of LER magnitude, with mock jurors being three times more likely to convict when the LER was "1 in 1000" versus "1 in 10."

One potential explanation for the inconsistent findings noted above is that the impact of LER information depends on other variables—in particular, the narrative context of the case. Narrative information concerning testing lab reliability was manipulated in a study by Lieberman et al. (2008) and found to interact with the focus of cross-examination. Conviction rates were notably lower in two of four conditions when: (1) cross-examination was focused on the expert and the analyses were performed at a "reliable" lab, *or* (2) cross-examination was focused on the evidence and the analyses were performed at an "unreliable" lab. Similar results were obtained in the study by Scurich (2015) which focused on the narrative context of LER information as well as its magnitude. Reported LER values ("1 in 10" vs. "1 in 1000") were manipulated along with the accompanying verbal description of laboratory procedures (none, sloppy, biased, or sloppy and biased). Participants were three times more likely to convict when they received the lower LER value, and two and a half times more likely to convict when the description of the lab suggested the existence of bias. Thus, the impact of LER information may well depend on the presence and salience of complementary narrative testimony.

Researchers have also begun to examine whether a DNA test result has the same effect as an analogous result from some other form of scientific test. These studies hold constant the result (i.e., an incriminating "match") while varying the type of test that produced it, allowing assessment of the degree to which jurors trust different types of scientific tests. In one of the first such studies, an incriminating DNA match in a rape case produced the same conviction rate as an incriminating

fingerprint match and an 8 % more convictions than a hair match (Lieberman et al. 2008). However, DNA evidence had greater relative impact in a murder case, with a DNA match yielding 17 % more guilty verdicts than a hair match and 33 % more convictions than matching fingerprints. In a second experiment using a more detailed case summary involving sexual assault (Lieberman et al. 2008), a DNA match said to be conclusive "beyond a reasonable standard of scientific certainty" produced significantly higher estimates of guilt probability than matching hair or fingerprints—but not blood type. Clancy and Bull (2015) also compared the impact of scientific test type in a study involving two case types (i.e., murder and rape). There was no impact of test type in the rape case, but the odds of a guilty verdict were 2.6 times higher for a DNA match than a fingerprint match in the murder case. Finally, Thompson and Newman (2015) manipulated the nature, strength, and presentation format of two types of scientific evidence—DNA and shoeprints—in a sexual assault case. They observed a 24 % conviction rate across conditions involving a DNA match, but only a 10 % conviction rate for the conditions involving a shoeprint match. To summarize, jurors appear to treat a declared match based on DNA as more probative than a match involving other forms of scientific evidence.

Neuroscientific Tests. The emerging field of neuroscience is broad and multidisciplinary, and a variety of neuroscientific tests have been developed. Two kinds in particular can be distinguished: neuroimaging and neuropsychological. *Neuroimaging* tests are usually performed in medical settings and involve the use of sophisticated scanning technologies to assess brain structure and/or functioning. Most notably, they yield pictorial images (or "scans") of the brain that are striking for their high resolution and vivid use of color. In contrast, *neuropsychological* tests are usually administered by a clinical psychologist and involve a battery of behavioral tasks (e.g., recognizing a pattern or recalling information). Both kinds of tests are commonly used to determine if an individual (e.g., defendant) has a brain condition that would reliably influence his or her behavior. In criminal trials, the defense may seek to introduce neuroscientific test results to show the defendant lacked the necessary criminal mindset ("mens rea") required for conviction, or as mitigation evidence in the sentencing phase of a capital trial. Neuroscientific tests have slowly gained traction in legal settings over time, but only recently have scholars begun to study their effect on jurors.

The first published empirical study of neuroscientific evidence appeared less than 10 years ago (Gurley and Marcus 2008). In this watershed study, three factors were manipulated in the context of a murder case: defendant psychiatric label (psychotic vs. psychopathic), presence/absence of an MRI scan showing a brain lesion, and the presence/absence of testimony describing a car accident that caused the injury. Four scans were included in the MRI-present condition showing extensive damage to the defendant's prefrontal cortex. Each of the manipulations had a significant effect on verdicts, but perhaps the most notable result of the study was that defendants were 34 % more likely to be found not guilty by reason of insanity when MRI images were presented to corroborate the existence of a brain lesion as opposed to when no neuroimages were presented. These provocative

results regarding the efficacy of neuroscientific evidence led to a number of subsequent studies aimed at isolating some of the potential effects associated with neuroscientific evidence. In particular, jury researchers have examined two broad research questions: (1) Does exposure to neuroscientific test results increase the likelihood of a pro-defense judgment? (2) Does exposure to pictorial images from neuroimaging tests have any incremental impact beyond the test results per se?

The answer to the first question is generally yes—favorable judgments for the defense tend to be more common when jurors are exposed to neuroscientific evidence that indicates a brain lesion. A number of research teams have now addressed this question by contrasting one or more neuroscientific evidence conditions (with or without images) and a no-neuroscience evidence control condition. For example, Schweitzer and Saks (2011) conducted a carefully designed study involving six evidentiary conditions in which the defendant in an assault case claimed insanity. In their focal experimental condition, a neurologist testified that the defendant had frontal lobe brain damage and presented an MRI scan to support the conclusion. Holding constant the expert's conclusion, other experimental conditions varied the provision and nature of an accompanying image presented by the expert (i.e., bar graph or control image) and/or the expert's specialization (i.e., clinical neuropsychologist or clinical psychologist). A control condition was also included wherein the defendant's claim of insanity was based solely on the testimony of family members. Participants in the four conditions that received the neuroscientific evidence collectively rendered significantly more NGRI verdicts (approximately 50 %) than those who read the clinical psychologist's testimony (43 %), as well as those in the control condition who did not hear from a mental health expert (13 %). In short, participants exposed to neuroscientific test results were more likely to choose NGRI than those who did not.

Another study demonstrating the efficacy of neuroscientific evidence was conducted by Schweitzer et al. (2011), who examined several potential influences associated with it. Using large nationally representative samples, they conducted four experiments that varied subsets of the following variables: (1) type of expert who presented evidence of a psychiatric disorder (e.g., neuroscientist, clinical psychologist vs. neurologist), (2) nature of the impairment (e.g., personality disorder vs. structural brain damage), (3) image presence (i.e., yes vs. no), (4) image content (e.g., irrelevant picture vs. fMRI scan), (5) type of neuroimage displayed (e.g.,. MRI vs. fMRI), and (6) image color (i.e., color vs. black-and-white). The four experiments involved different case types (e.g., homicide vs. armed robbery) and focal outcomes (e.g., verdicts vs. sentences). Six conditions were common to all four experiments, including three involving neuroscientific evidence and a control condition with no scientific evidence. This allowed for meta-analytic comparisons to be made using the combined dataset (N = 1374), and these revealed only one significant effect—participants who received neuroimaging-based testimony along with a supporting scan returned 12 % fewer guilty verdicts than those in the control condition who received no scientific evidence. Conviction rates in the other two neuroscientific evidence conditions were not significantly different from the control condition. Participants in this "neuroimaging test + scan" condition also

judged the defendant to be less in control of his actions than did control participants.

The research reviewed thus far indicates jurors are influenced by neuroscientific evidence, but several studies suggest its impact may depend on other variables. For example, in a death penalty case, Greene and Cahill (2012) observed a mitigating effect of neuroscientific evidence indicative of frontal lobe damage, but only when the defendant was said to pose a high risk of being dangerous in the future. Specifically, when the defendant was deemed highly dangerous, participants who learned only of the defendant's psychiatric diagnosis (i.e., psychosis) were almost 22 times more likely to render a death sentence than those who were informed of the diagnosis and given the neuropsychological test results (with or without accompanying MRI and PET scans). When the defendant was said to pose a low risk of future danger, neuroscientific evidence (test results and scans) had no impact on sentence.

Saks and his colleagues examined the impact of neuroscientific evidence in the context of a capital crime and found its impact to depend on the defendant's psychiatric diagnosis (Saks et al. 2014). Large nationally representative samples were recruited and exposed to one of four levels of scientific evidence (clinical, genetic, neurological test results, and neurological results + neuroimage) in two online experiments. The psychiatric condition of the defendant was manipulated as well (i.e., healthy, schizophrenic, psychopathic), along with the defendant's future dangerousness (low vs. high) and the side that introduced the scientific evidence (Study 2 only). In both studies, there were no main effects of scientific evidence type but interactions arose involving evidence type and another variable. In the first experiment, when neuroscientific results were presented without a neuroimage, psychopathic defendants did not differ from healthy or control defendants in terms of sentence, but were given the death penalty 21 % more often than schizophrenic defendants. Conversely, when neuroscientific test results were presented along with a neuroimage, the frequency of death sentences *decreased* for psychopathic defendants but *increased* for schizophrenic defendants, essentially eliminating the difference. In the second experiment, scientific evidence produced a backfire effect against the side that introduced it except in the neuroimage condition. Specifically, presentation of scientific evidence for mitigation purposes by the defense *increased* death sentences, whereas presentation of scientific evidence for aggravation by the prosecution *reduced* death sentences.

Finally, the impact of neuroscientific evidence also varied in another recent study wherein neuroimaging test results (along with an MRI scan) were examined along with evidence of a genetic abnormality and other potential explanations for the defendant's behavior (Applebaum et al. 2015). Three cases were used involving three types of judgments (i.e., prison term, insanity judgment, capital sentence), with neuroimaging evidence presented in two of the studies. The neuroimaging evidence had no effect on the recommended incarceration length for a convicted murderer in the first study, but reduced the likelihood of a death sentence in the third study compared with when the defense attorney claimed the killing was simply impulsive. Thus, the findings of several studies suggest the impact of neuroscience evidence may depend to some extent on the defendant's psychiatric status and the nature of the case/judgment.

In contrast, two studies have examined the impact of neuroscientific evidence in the context of detecting deception, and both revealed an effect on mock jurors (McCabe et al. 2011; West et al. 2014). In the first study (McCabe et al. 2011), fMRI test results (without an image) indicating deception yielded more convictions in a homicide case than a control condition involving no evidence of deception (as well as conditions featuring evidence of deception based on either a polygraph test or a thermal imaging test). However, this difference vanished when the validity of the fMRI was questioned during cross-examination. In the second study (West et al. 2014), a form of neuroscientific evidence based on the measurement of event-related potentials (ERP) was examined in a case involving burglary and assault. In one condition, a police detective testified the defendant had guilty knowledge about the crime based on a behavioral interview. In three other experimental conditions, a neuroscience expert testified about the results of an ERP-based lie detection test that indicated the defendant had guilty knowledge and the type of image accompanying the expert's testimony was manipulated (i.e., no image, a "brain wave" image, and a "brain map" image). Conviction rates were significantly higher in the four conditions featuring expert testimony about the defendant's deceptiveness compared with a no-expert control condition, but there were only trivial differences in the conviction rate among the conditions featuring expert testimony. In other words, the sophisticated ERP-based procedure did not produce more convictions than the police expert.

The other major question noted previously regarding the impact of neuroimaging scans has been driven by concern that jurors will be seduced by colorful high-tech images and invariably make judgments consistent with those images when presented with them. In particular, some prosecutors fear that allowing jurors to see vivid pictures of the defendant's damaged brain will lead jurors to accept defense claims in an uncritical manner. This fear is consistent with several studies wherein presentation of irrelevant neuropsychological information caused some study participants to report a better understanding of the focal behavioral phenomenon and inflate their ratings of the study's scientific quality (Rhodes et al. 2014; Weisberg et al. Weisberg et al. 2008). So are jurors actually influenced by *neuroimages* as opposed to neuroscientific test results?

The visual effect of neuroimages has been addressed in a number of the neuroscientific evidence studies already described. In order to draw conclusions about the incremental value of neuroimaging scans, conviction rates have been compared for two types of conditions: those featuring only verbal explanations of neuroscientific test results and those featuring the same verbal test results plus one or more neuroimages (e.g., an MRI showing a brain lesion). These studies uniformly show little if any effect attributable solely to the presence of a neuroimage (Greene and Cahill 2012; Saks et al. 2014; Schweitzer et al. 2011; Schweitzer and Saks 2011; West et al. 2014). In one of the best studies on this topic, Schweitzer and his colleagues (Schweitzer et al. 2011) conducted four carefully designed experiments to isolate the verbal and visual effects of neuroscientific evidence. Across all four studies, when the effect(s) of neuroscientific evidence was parsed using various comparison conditions, there was no incremental effect associated with presentation of neuroimages over the test results themselves. In particular, it did not matter what

was shown in the images (e.g., a brain scan vs. a graph of brain functioning) or whether the images were presented in color or in black-and-white.

Overall, the growing body of work on neuroscientific evidence suggests neuroscientific test results do yield more pro-defense judgments on occasion, but the exact conditions are not clearly delineated at present. That said, there is no indication that merely showing jurors a pictorial scan of a defendant's brain has any impact beyond verbally informing them of the corresponding neuroscientific test results.

Other Forms of Scientific Evidence. Although scholars have focused thus far on DNA and neuroscientific evidence, research is emerging on several other forms of scientific evidence that are regularly introduced at trial. These include the comparison of fingerprints (Clancy and Bull 2015; Garrett and Mitchell 2013), hair samples (McQuiston-Surrett and Saks 2009), and shoeprints (Koehler 2011; Martire et al. 2013, 2014; Thompson and Newman 2015). In addition, toxicology reports (Jenkins and Schuller 2007) and genetic tests (Applebaum and Scurich 2014; Applebaum et al. 2015; Saks et al. 2014) have also received some attention. Despite the limited research base on these forms, several interesting findings have emerged.

First, the exact wording used to convey a "match" may be less important than whether the possibility of error is acknowledged (Garrett and Mitchell 2013; Koehler 2011; McQuiston-Surrett and Saks 2009). For example, in two experiments involving the comparison of hair samples, two qualitative expressions for an incriminating result (i.e., "match" vs. "similar in all microscopic characteristics") yielded higher estimates of defendant source likelihood than a judgment of indistinguishability plus an RMP estimate in a frequency format that mentioned a large number of potential coincidental matches in the surrounding population (McQuiston-Surrett and Saks 2009). However, the lower perceived likelihood associated with the quantified RMP condition disappeared in the second experiment when the expert also offered the "ultimate" conclusion that the defendant was the source of the crime-scene hair. Another study by Koehler (2011) suggests jurors attend to how forthcoming an expert is about the possibility of different kinds of error. Holding constant the aggregate likelihood of error at "1 in 1000," mock jurors convicted about 20 % less often when the prosecution's shoeprint expert acknowledged three kinds of potential error (i.e., coincidental match, lab mix-up, and examiner misinterpretation) as opposed to when only a lab mix-up and/or a coincidental match was noted. Interestingly, there were no differences on any outcomes between two additional conditions featuring an unqualified "match" result (i.e., individualization) regardless of whether the prosecution's expert was subjected to a rigorous cross-examination addressing all three kinds of potential error.

Another study—one of the first on fingerprints—underscores the relative importance of "match" wording and acknowledging error. Garrett and Mitchell (2013) conducted two experiments involving a burglary trial and found mock jurors to be relatively insensitive to the exact language used to describe a match. In the first experiment, the authors examined 15 different ways of conveying four broad types of conclusion (i.e., simple positive match, bolstered positive match, qualified/inconclusive match, and exclusion). In essence, *all* of the various expressions for a positive match (whether simple or bolstered) produced higher likelihood ratings that the defendant was the robber than inconclusive expressions

and qualified match expressions that acknowledged that someone other than the defendant could have been the source of the crime-scene prints. Participant insensitivity to wording variation within the positive match categories (i.e., simple vs. bolstered) was replicated in the second experiment. The defendant was also rated as more likely to be the culprit when the possibility of error was acknowledged by the expert and then explicitly discounted, as opposed to when error was simply acknowledged to have been possible. In sum, these studies suggest jurors will be largely unaffected by minor wording variation in expressions that essentially convey the defendant matches the crime-scene sample, but sensitive to the acknowledgement of error and the number of types of errors identified as possible.

A second finding that emerges from recent work is that test results providing only modest support for a match can have a paradoxical effect on jurors (Martire et al. 2013, 2014; Thompson and Newman 2015). This phenomenon has been termed the "weak evidence effect." In the first study to note it, Martire et al. (2013) manipulated the evidentiary strength of an incriminating shoeprint comparison (low, moderate, and high) and the format of its presentation (numerical or verbal) in a factorial design. Verbal expressions of evidentiary strength were taken from standards offered by the Association of Forensic Science Providers (2009). Of particular note, the wording for the low evidentiary strength comparison indicated "weak or limited support" for a match. Surprisingly, even though all participants were told the defendant's shoeprints matched the crime-scene marks, 62 % of those in the "low strength-verbal format" condition provided a final rating of the defendant's guilt probability that was *lower* than their a priori probability—indicating they basically treated the weak/limited incriminating evidence as exculpatory. Further, 23 % of participants in this condition switched their preferred verdict from guilty to *not guilty* upon receiving the evidence of an incriminating match.

Additional research has replicated the incongruous effect for both incriminating and exculpatory shoeprint comparisons. In a second experiment described in the 2013 report, Martire and her colleagues found 67 % of the inculpatory-verbal condition participants reported a lower belief in guilt (and 39 % switched their verdicts to not guilty) after being told that the incriminating shoeprint comparison provided only weak/limited support for a match. Similarly, 31 % of those in an exculpatory-verbal condition reported a *stronger* belief in guilt after learning of limited/weak support for a nonmatch (with 18 % switching their verdicts to guilty). Martire et al. (2014) observed similar results in a follow-up study featuring a different form of scientific evidence, with 64 % of those in the "low strength-verbal format" condition downgrading their estimated probability of guilt after learning of an incriminating fingerprint comparison. Finally, consistent with these results, more exculpatory shifts were observed by a different research team in yet another study when a shoeprint comparison was characterized as providing "moderate" support for a match as opposed to "very strong" support (Thompson and Newell 2015). Essentially, verbally characterizing a scientific test as providing tepid support for a particular conclusion may induce belief change in the opposite direction—a "boomerang" effect.

A third finding that emerges from recent work is that jurors do not place much weight on the potential genetic influence on behavior (Applebaum et al. 2015; Applebaum and Scurich 2014; Saks et al. 2014). Several studies have provided some participants with scientific evidence of defendant impulsivity that was attributed to the presence of a genetic abnormality involving the *monoamine oxidase A* (MAOA) allele. Relative to other explanations for impulsive behavior, genetic evidence produced only slightly lower ratings of defendant responsibility and future dangerousness than clinical and neurological evidence in one of two capital cases (Saks et al. 2014), did not affect verdict choice or prison sentence in a homicide case (Applebaum and Scurich 2014), and had no discernible effect on a variety of juror judgments (i.e., length of incarceration, verdict, sentence) across three case types (Applebaum et al. 2015). These underwhelming effects suggest jurors will not be more lenient simply because there is a ready attribution for the defendant's criminal behavior that lies beyond his or her control.

What Contextual and Dispositional Variables Affect the Impact of Scientific Evidence?

In addition to studying the characteristics of scientific tests, jury scholars have begun examining contextual and individual difference variables that might enhance or diminish the impact of scientific evidence. The list of dispositional characteristics that could conceivably affect the impact of forensic science evidence is lengthy and includes demographic variables, attitudes associated with the legal system, and various personality traits. In practice, however, only a few individual difference variables have received more than passing attention from researchers: crime-show TV viewing habits, trust in forensic science evidence, and numeracy. The primary contextual variable of interest thus far has been the strength of the non-forensic evidence associated with the case (e.g., the existence of a confession or the quantity and quality of eyewitness testimony). In this section, we review research on key dispositional and contextual variables that may play an important role in trials involving scientific evidence.

Crime-Show Television Viewing. The first episode of "Crime Scene Investigation" (*CSI*) aired in the autumn of 2000. The program was an instant hit and ran for 15 years, annually ranking as one of the top TV shows and proving so popular that several spin-offs were developed and aired (i.e., *CSI: Miami*, *CSI: New York*, and *CSI: Cyber*). In addition to the *CSI* franchise, a number of other TV shows featuring forensic evidence have achieved multiyear runs, including *Bones*, *Cold Case*, *Forensic Files*, *NCIS*, *NCIS: Los Angeles*, and *Without a Trace*. These shows commonly depict the collection, testing, and interpretation of forensic evidence in a streamlined and uncritically positive manner. Forensic evidence—often DNA—is readily available, easily collected, quickly tested, and unfailingly accurate. So the question arises: Do the TV viewing habits of jurors affect their

processing of scientific evidence? Many people—judges, attorneys, police officers, legal scholars, and journalists—seem to think so (Maeder and Corbett 2015). The basic fear is that repeated exposure to unrealistic, idealized depictions of forensic science evidence will cause jurors to expect the prosecution to present it at every trial. If forensic evidence is presented, jurors will convict; if not, they will acquit. This phenomenon has been referred to as the *CSI effect* (Hayes and Levett 2013). References to it began appearing in the media as early as 2003 (Houck 2006), and it continues to receive a remarkable amount of media attention (Hayes and Levett 2013; Maeder and Corbett 2015).

The considerable fuss devoted to the CSI effect contrasts sharply with the empirical support for its existence. One challenge associated with drawing conclusions from research on the CSI effect is the considerable methodological variation across studies. In particular, researchers have measured exposure to crime-show TV programming in different ways and correlated these various measures with a wide array of attitudinal and behavioral outcomes. Nonetheless, despite this methodological diversity, one finding emerges consistently in the literature—measures of TV crime-show viewing are unrelated or, at best, only weakly related to verdict preferences (Hayes-Smith and Levett 2011; Holmgren and Fordham 2011; Kim et al. 2009; Lieberman et al. 2008; Maeder and Corbett 2015; Mancini 2011; Podlas 2006; Schweitzer and Saks 2007; Shelton et al. 2006).

Several recent studies of the CSI effect provide illustrative results. Mancini (2011) assessed the association between TV crime-show viewing and mock juror judgments in the context of a murder trial that featured ballistics evidence. Participants viewed a 30-min edited video that included footage of a real trial. Exposure to two types of TV crime-show programs—fictional and documentary—were distinguished and measured in terms of both frequency of viewing and perceived realism. All of these crime-show viewing measures were weakly and non-significantly related to dichotomous verdicts as well as a continuous measure of guilt confidence. In another study involving persons who showed up for jury duty and were excused, TV crime-show viewing frequency failed to correlate significantly with verdict preference, but did interact with the level of forensic evidence presented by the prosecution to explain some incremental variance in a multivariate regression model (Hayes-Smith and Levett 2011). Another study that obtained separate measures of TV crime-show viewing frequency and perceived realism revealed small direct and indirect effects for frequency but not realism (Maeder and Corbett 2015), although these weak effects emerged from the post hoc analysis of a saturated path model and could simply reflect capitalization on chance.

One minor exception to the consistent pattern of null findings regarding the CSI effect occurred in a telephone survey where respondents were asked to indicate the likelihood they would convict in a rape or murder case if the prosecution had no scientific evidence (Baskin and Sommers 2010). Individuals who reported watching three or more hours of TV crime shows per week said they would be less likely to convict in the absence of forensic evidence than respondents who watched less crime-show TV programming. A caveat exists for this finding however—the results

were not obtained in the context of a trial and so reflect generalized attitudes about scientific evidence as opposed to their impact in a specific case.

Overall, there is little support for the notion that watching TV crime shows directly affects the verdict preferences of mock jurors, but there is some support for the idea that this form of media exposure affects attitudes and beliefs about forensic science that may indirectly affect juror decision making.

Trust in Forensic Evidence. There are many ways to conceptualize and measure jurors' attitudes and beliefs about scientific evidence, but a common element is the degree of trust that a juror has in its probative value. Trust in forensic evidence likely arises from a number of sources, including background life experiences and transmitted cultural values, and could affect untutored estimates of the likelihood of different kinds of errors as well as how jurors react to trial evidence suggestive of possible errors. Early research on attitudes related to forensic evidence focused on DNA testing. In one of the first studies, a one-item measure of trust in DNA evidence was related to verdict preferences and guilt probability ratings across four types of cases (Britton 1997). Lieberman and his colleagues (Lieberman et al. 2008) found that a four-item measure of "pretrial DNA beliefs" (which they also referred to as a measure of trust) was associated with verdict preferences and explained 11 % of the variance in probability-of-guilt ratings. A later study employing this same measure found it to have a weak and marginally significant direct effect on a continuous measure of guilt confidence in a path analysis involving nine predictors (Maeder and Corbett 2015). Other researchers have developed and validated a 10-item "Forensic Evidence Evaluation Bias" scale (Smith and Bull 2012, 2014). In the 2012 study, the pro-prosecution subscale had a significant indirect effect on guilt probability via the perceived strength of the DNA evidence. The 2014 study yielded the same indirect effect for cases involving sexual assault and robbery. In essence, these studies suggest jurors' trust in forensic evidence is modestly related to their willingness to convict.

Numeracy. Another variable that could influence jurors' reaction to forensic evidence is their facility with numbers—their so-called *numeracy*. Researchers have measured this construct using objective indicators such as courses taken in math and/or stats, as well as subjective measures that tap respondents' self-assessed ability, comfort, or lack of anxiety in working with numbers. Several studies have included one or two isolated items of this type which were then employed as control variables. For example, Kaasa et al. (2007) used a one-item measure that asked participants to indicate their confidence in being able to "draw correct conclusions from numerical data, such as probabilities and frequencies." This was the only significant individual-difference variable to emerge in multiple regression analyses, with respondents who reported more confidence tending to give higher probability-of-guilt ratings and be more likely to render a guilty verdict. Higher scores on this one-item measure of numeracy were also associated with a greater tendency to convict in a study based on a sexual assault case featuring DNA evidence (Thompson et al. 2013).

Multiitem measures of numeracy have also been developed and validated in medical decision making contexts, including an eight-item Subjective Numeracy

Scale (SNS) (Fagerlin et al. 2007; Zikmund-Fisher et al. 2007). This measure has been used in several mock juror studies featuring scientific evidence (Garrett and Mitchell 2013; Martire et al. 2013, 2014). In general, although there is support for its construct validity, subjectively perceived numeracy as measured by the SNS has not correlated strongly with mock juror judgments, failing to achieve statistical significance as a covariate in studies by Martire and her colleagues (Martire et al. 2013, 2014). SNS scores did, however, demonstrate some predictive validity in a study by Garrett and Mitchell (2013), with high-SNS participants tending to report a higher probability that the defendant was the source of crime-scene fingerprints and greater confidence he was the robber.

Finally, an objective measure of numeracy has also been developed and used in juror decision making research (Lipkus et al. 2001). This eight-item test involves statistical questions with correct answers and has been used in at least one study in which the impact of the scientific evidence was found to depend on numeracy (Scurich 2015). Specifically, numeracy interacted with the presence/absence of narrative information regarding testing lab reliability (i.e., the possibility of contamination and/or bias) and the quantified LER value to affect verdict choice. Participants below the median on numeracy were affected by anecdotal narrative information about the lab's reliability, but unaffected by the quantified LER (i.e., "1 in 10" vs. "1 in 1000"). Conversely, those above the median on numeracy were responsive to the explicit LER value provided, but not the anecdotal information about lab reliability.

Case Context. There are good reasons for expecting characteristics of the case context to influence jurors' processing of scientific evidence. According to the story model, jurors attempt to fashion a compelling narrative to account for the evidence at trial. One implication of this cognitive process is that if jurors are presented with strong scientific evidence that is inconsistent with all or even much of the other evidence in the case, the scientific test results might simply be discounted and given no weight. In other words, the story model suggests characteristics of the scientific and non-scientific evidence will interact. Few studies have examined whether aspects of the case context (such as non-forensic evidentiary strength) moderate the impact of scientific test results, but there is some empirical support for the notion.

For example, there is indication that the context of scientific evidence matters to jurors. In an early demonstration, different ways of framing the same probabilistic evidence had a dramatic impact on jurors' verdict preferences in a civil trial (Wells 1992).Participants were reluctant to find a defendant liable when the statistical evidence presented was not specific to the case at hand and did not lend itself readily to a narrative consistent with liability. A subsequent study essentially replicated Wells' initial finding (Niedermeier et al. 1999). This result has been termed the *Wells effect*, and refers to jurors giving greater weight to "contextualized" statistical information as opposed to "naked" statistical data that is less obviously pertinent to the case at hand.

A pair of studies by Nicholas Scurich and his colleagues provides further support for the notion that jurors are reluctant to convict on the basis of "naked" statistical evidence. Two aspects of case context were manipulated in the first study—the

origin of the matching DNA sample and the strength of the non-scientific evidence (Scurich and John 2011). DNA sample origin was manipulated by telling jurors the defendant's sample was obtained either directly from the crime scene or through an automated "trawl" search of an electronic database. Although the characteristics of the DNA test itself were held constant (e.g., a conclusive match), study participants were 14 % less likely to convict when the defendant's sample was identified via the trawl search—presumably because of the added difficulty of constructing a narrative involving culpability from this evidence. A follow-up study was conducted to examine the effect of labeling the electronic database used in a trawl search (Scurich and John 2013). Here, participants were significantly more likely to return a guilty verdict when the defendant's profile was reportedly obtained from a national database for "convicted felons," as opposed to a national "medical" database containing DNA samples obtained from infants who received vaccinations or an unspecified national database.

Other research further supports the conclusion that context matters to jurors. Smith et al. (2011) had mock jurors rate the strength of different versions of DNA evidence and manipulated two characteristics associated with the crime-scene sample: mobility (i.e., transferableness) and investigative relevance (basically a function of the sample's location). Strength of evidence ratings were affected by both factors and, consistent with the story model, all versions of the DNA evidence were rated as stronger when presented in the context of a case vignette that included summaries of the trial evidence offered by the prosecution and defense than when rated in isolation. The strength of the non-scientific evidence has also directly affected verdict choice in several studies that featured DNA evidence. For example, inclusion of a defendant alibi substantially reduced the number of convictions produced by DNA evidence in an early study by Golding et al. (2000). Scurich and John (2011) manipulated the strength of the non-scientific evidence by varying aspects of the defendant's alibi and the prosecution's eyewitness testimony. The strong non-scientific evidence produced a conviction rate double that of the weak non-scientific evidence regardless of the origin of the matching DNA sample. Thompson et al. (2013) varied the non-scientific evidence in a similar fashion, contrasting a strong defendant alibi and an uncertain victim with a weak alibi, a closely matching description of the perpetrator, and a more confident victim. Participants receiving the strong non-scientific evidence convicted 25–45 % more often than those who got the weaker case evidence. Finally, Reiser (2015) varied non-scientific evidence strength along with the presence/absence of a DNA match (Study 1) or the LER associated with a DNA match (Study 2). A marginally significant main effect emerged for non-scientific strength of evidence in the first experiment, whereas the second experiment produced a very clear effect. Most interesting, however, was an interaction whereby perceptions of the quality of the DNA evidence were affected by whether the non-scientific evidence was strong or weak.

Overall, the results of these studies are generally consistent with the story model in showing that juror decisions in trials featuring scientific evidence can be influenced by aspects of the case that affect the coherence of the emergent narrative.

Discussion

Major Empirical Findings

Considerable empirical research now exists on the impact of scientific evidence on jurors, and there is sufficient depth and convergence on a number of topics to offer some observations about what has been learned.

First, jurors generally do not use scientific evidence in a manner consistent with the Bayesian model. Numerous studies demonstrate that predictions derived from the Bayesian model do not correspond well to judgments made by mock jurors, with the latter typically *underutilizing* scientific evidence relative to what the model would forecast (e.g., Goodman 1992; Faigman and Baglioni 1988; Martire et al. 2013, 2014; Schklar and Diamond 1999; Thompson and Schumann 1987). Although jurors occasionally approximate Bayesian predictions (Thompson et al. 2013; Thompson and Newman 2015), these occurrences appear to be exceptions to the general rule. This finding is especially worthy of note in light of the widespread concern that jurors will convict in a reflexive fashion when presented with incriminating scientific evidence.

Second, however, presentation of an incriminating (i.e., matching) scientific test result generally has a strong effect on juror judgments. This "presence" effect has been very robust in studies featuring DNA evidence (Dartnall and Goodman-Delahunty 2006; Golding et al. 2000; Goodman-Delahunty and Hewson 2010; Nance and Morris 2002, 2005; Smith and Bull 2014; Thompson et al. 2013). It has also been observed in studies involving comparisons of different types of scientific evidence, including fingerprints (Garrett and Mitchell 2013), shoeprints (Koehler 2011), bullet lead (Kaasa et al. 2007), and blood type (Goodman 1992). The size of the observed effect has varied somewhat, but the presence of an incriminating DNA test result generally increases convictions by 20–50 % over no scientific evidence or the report of an inconclusive test result.

Third, presentation of an exculpatory (i.e., exonerating) scientific test result does not necessarily yield an acquittal (Ask et al. 2011; Lieberman et al. 2008; Pozzulo et al. 2009; Simon et al. 2004). This finding has been observed repeatedly with DNA evidence, as well as shoeprint evidence (Martire et al. 2013), hair (Lieberman et al. 2008), and fingerprints (Garrett and Mitchell 2013; Lieberman et al. 2008). Assuming an "exclusion" result should be sufficient grounds for a reasonable doubt in most trials, the fact that conviction rates have not dropped to zero in studies featuring exculpatory scientific evidence is troubling. One possible explanation for this phenomenon is that it represents a methodological artifact associated with using simulated trials, arising from one or more mundane causes such as low participant motivation, low salience of the scientific evidence, omission of a "reasonable doubt" instruction, or lack of a deliberation process. Alternatively, the phenomenon may simply reflect participant awareness that errors can occur in both directions—i.e., they can falsely implicate an innocent defendant as well as falsely indicate a guilty individual is not the perpetrator.

Fourth, the impact of a given test result depends on the type of test that produced it. In other words, all scientific tests are not equally probative in the minds of jurors—some types are trusted more than others. Focusing on the extremes, DNA tests are now widely seen as trustworthy by the general public (Lieberman et al. 2008), whereas the polygraph appears to be viewed very skeptically (Myers et al. 2006). Accordingly, test results for some types of scientific evidence will likely affect jurors more than the results of others. A moderating effect of scientific test type could help to explain the inconsistent findings associated with acknowledging the possibility of error and providing an explicit quantitative estimate of its likelihood. Sometimes the possibility of error is not mentioned, sometimes general acknowledgement is made on direct exam with or without provision of a specific quantified value regarding the likelihood of one or more types (i.e., RMP or LER), and sometimes the possibility of error is first acknowledged and then scrutinized during cross-examination. Error-related information has had a relatively large impact on mock jurors in some studies but not others. In general, acknowledging error and/or providing an explicit incidence rate of its occurrence has tended to *increase* pro-prosecution judgments when associated with a "trusted" form of science evidence such as DNA (e.g., Koehler et al. 1995; Nance and Morris 2002, 2005), but have *no impact or decrease* pro-prosecution judgments in studies involving other forms of scientific evidence (e.g., Garrett and Mitchell 2013; Koehler 2011; McQuiston-Surrett and Saks 2009). One factor complicating the assessment of impact associated with type of scientific test is that the attitudes and beliefs of jurors are likely to change over time as lesser-known techniques become more familiar to the general public.

Fifth, the statistical frame used to convey error likelihood (e.g., frequencies vs. probabilities or likelihood ratios) affects juror verdict preferences (e.g., Koehler 1996, 2001; Koehler and Macchi 2004; Nance and Morris 2002, 2005). Most of the research on statistical frame has focused on one particular kind of scientific evidence (i.e., DNA) and one particular kind of error (coincidental match), but mock jurors have been consistently less likely to convict when RMP is provided in a frequency format and reference is made to the existence of multiple individuals in the surrounding community who might coincidentally match the crime-scene sample. It is unclear, however, if these results will generalize to RMP values that are extremely small (as many are nowadays) and/or to forms of scientific evidence other than DNA.

Sixth, the magnitude of quantified error likelihood (i.e., RMP and LER) generally has a modest impact on verdict preferences (Britton 1997; Dartnall and Goodman-Delahunty 2006; Koehler 1996, 2001; Koehler et al. 1995; Nance and Morris 2002, 2005; Schklar and Diamond 1999; Thompson et al. 2013; Thompson and Newman 2015). Statistical significance associated with RMP and LER levels has varied across studies, but the difference in conviction rates between the highest and lowest error-likelihood values provided in a study has tended to range from only 5 to 15 %. Particularly surprising is the underwhelming influence attributable to laboratory error, which almost certainly occurs at a much higher rate than coincidental matches (Koehler et al. 1995). A potential explanation for the

negligible effect of quantified error rate is the lack of complementary information about error in the trial narrative (or perhaps the low salience of such information). In particular, LER values have usually been presented in a brief, unelaborated fashion that might make them easy to overlook, forget, or discount. In one study where LER manipulation produced a notable effect, corresponding anecdotal information about the lab's reliability was also provided (Scurich 2015). In another study, narrative information about an unreliable lab produced a notably lower conviction rate *only* when the prosecution's expert was grilled about lab procedures and conditions on cross-exam (Lieberman et al. 2008). Thus, in keeping with the availability heuristic, numerical error-rate information may well have a stronger effect when accompanied by testimony that establishes a case-relevant scenario conducive to testing error (e.g., an expert's admission that biological evidence was stored improperly, placed in contact with a contaminant, or handled by a racist police officer).

Seventh, the impact of scientific evidence likely depends on contextual variables such as how well it fits into an overall trial narrative and the strength of the other (non-scientific) evidence. Research on "naked" statistical evidence demonstrates that isolated probabilistic information has little effect on jurors compared with equally probative data embedded in a narrative framework (Niedermeier et al. 1999; Wells 1992). Systematic variation of the strength of the non-scientific evidence generally affects verdicts as well (Golding et al. 2000; Reiser 2015; Scurich and John 2011, 2013; Thompson et al. 2013). These findings are consistent with theories such as the story model (Pennington and Hastie 1993) and the director's cut model (Devine 2012), and should serve to guard against any scholarly preoccupation with characteristics of the scientific evidence in isolation.

Eighth, dispositional variables are likely to affect the use of scientific evidence in some cases. Many studies have included one or more measures of participant attitudes, beliefs, or prior knowledge, but there is emerging empirical support for the importance of two particular constructs: (1) *trust in forensic science evidence* (Britton 1997; Lieberman et al. 2008; Smith and Bull 2012, 2014), and (2) *numeracy* (Garrett and Mitchell 2013; Martire et al. 2013, 2014; Scurich 2015). Jurors' ability to work with numbers and their trust in those elements of the legal system associated with the collection, testing, and reporting of scientific evidence are likely to be relevant in many trials. Research has already revealed some effects for these constructs and theory suggests they may interact with characteristics of scientific evidence as well. Other constructs, such as need for cognition, may also play a role (Mancini 2011), particularly when scientific evidence is especially complex. In contrast, TV crime-show viewing habits do *not* appear to have a discernible effect on juror verdict preferences. Put differently, there is very little support for the so-called "CSI effect" (Mancini 2011; Smith and Bull 2012). However, TV crime-show viewing habits might influence juror judgments via mediating variables such as trust in forensic evidence, so there could still be indirect effects of watching crime shows that feature a heavy dose of scientific evidence.

A Conceptual Framework for Scientific Evidence

Drawing together the threads from the preceding discussion, Fig. 1 displays a conceptual framework for understanding the impact of scientific evidence. Our intention is to identify the variables that are most important in cases involving scientific evidence, as well as specify likely relationships among those variables. Consistent with our focus in this chapter, the framework is aimed at trials featuring scientific evidence offered to yield diagnostic conclusions about specific individuals.

In essence, three sets of variables are identified that purportedly determine the impact of scientific evidence on jurors: (1) characteristics of the scientific evidence, (2) characteristics of the non-scientific evidence, and (3) characteristics of the juror. *Characteristics of the scientific evidence* include its type (e.g., DNA, fingerprints, hair), numerical parameters associated with error likelihood (e.g., absolute level and metric of RMP or LER), and verbal parameters associated with the wording used to convey test results (e.g., "match," "exclusion," or phrasing used to describe the level of support for a conclusion). *Characteristics of the non-scientific evidence* include its overall strength and the degree to which it can be formed into a compelling narrative (i.e., story). *Characteristics of the juror* include facility at using numbers (i.e., numeracy), need for cognition, and trust in forensic science. The latter construct is multidimensional and includes beliefs about the probative value

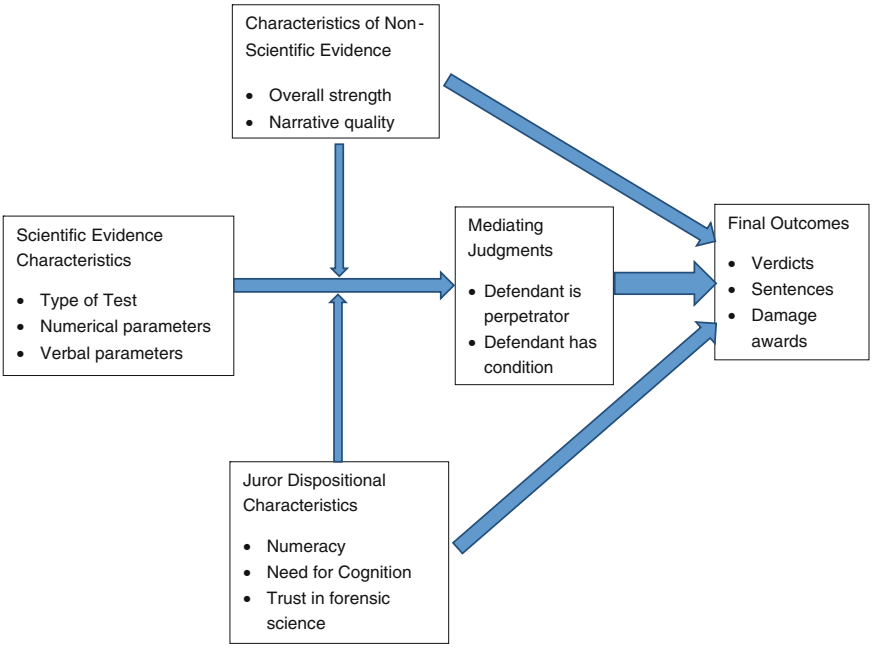

Fig. 1 Conceptual framework for understanding the impact of scientific evidence on jurors

of scientific evidence as well as the integrity and competence of those who handle, test, and present scientific evidence (e.g., the police, evidence technicians, and expert witnesses).

A number of causal relationships are theorized to exist among these variables and depicted with arrows in Fig. 1. In essence, characteristics of the scientific tests, the non-scientific evidence, and the jurors themselves are viewed as the primary determinants of two key juror judgments—whether the defendant is the perpetrator of the crime and/or whether the defendant has a particular condition (e.g., poor impulse control). These judgments in turn are expected to be the major determinants of jurors' preferred verdicts and sentences, although other variables (such as the standard of proof) may be involved as well at this final stage of the decision process.

In addition to the direct effect of scientific test characteristics on juror judgments, characteristics of the non-scientific evidence and the jurors themselves are expected to directly influence decision outcomes. Further, characteristics of the non-scientific evidence and the jurors are viewed as likely to moderate the impact of scientific test characteristics on judgments about the defendant's identity and condition. For example, presentation of a matching DNA test result with low stated probabilities of RMP and LER would ordinarily be expected to have a strong impact on a juror's judgment regarding whether the defendant is the actual perpetrator of a crime. However, if the corroborating non-scientific evidence is exceptionally weak or missing major narrative elements (e.g., a motive for murder), jurors should be more likely to discount the DNA test result. The same dampening effect could occur if a juror had little trust in the results of DNA testing due to knowledge of the errors that can occur in testing labs. Conversely, the impact of a marginal scientific test result might be amplified for jurors with high levels of trust in the legal system or when there is overwhelming non-scientific evidence to corroborate the test result.

Overall, the conceptual framework suggests a number of variables will be involved in determining juror decision outcomes in trials featuring scientific evidence. Moreover, the impact of science evidence characteristics on juror decisions is expected to be mediated by judgments about the defendant, and moderated by contextual and individual-difference variables. Accordingly, two rather straightforward implications of the conceptual framework are that many variables will be needed to fully account for decision outcomes in studies involving scientific evidence, and we should not expect scientific evidence variables to consistently have strong main effects on those outcomes.

Future Research Directions

An important goal of this chapter is to call attention to the need for research on a number of particular questions and topics. In keeping with the framework just presented, one recommendation is for researchers to concurrently examine characteristics of the scientific evidence, the non-scientific evidence, and the decision

makers. For example, the presence/absence of an incriminating scientific test result could be manipulated while varying such things as: (1) the quantified likelihood of error reported, (2) the wording associated with the conclusion, (3) the level of scrutiny afforded to the possibility of error during cross-exam, or (4) the strength of the non-scientific evidence. In addition, measures of participants' trust in forensic science and numeracy could also be obtained. Special attention might be devoted to identifying those conditions when scientific evidence has its greatest and least impact. This would entail a focus on interactions between characteristics of the scientific evidence and other variables associated with the case and the jurors.

A second recommendation is to expand the repertoire of methodologies used to study the impact of scientific evidence. In particular, structured interviews and verbal protocols could be used to increase our understanding of how jurors reason with the evidence. Verbal protocols involve asking participants to "think aloud" as they go about completing some decision task. The resulting narrative is then recorded and coded, with efforts made to identify important variables that emerge and causal connections between them. With respect to juror decision making, verbal protocols could be obtained by asking a representative sample of participants (ideally venirepersons) to talk aloud as they read a written case summary or watch a videotaped trial involving scientific evidence. Structured interviews could be used to obtain similar information from real jurors shortly after they serve on an actual jury. Both approaches would be valuable to use in conjunction with more traditional data-collection strategies, complementing the identification and estimation of relationships using more quantitative approaches.

A third recommendation is for jury scholars to systematically examine the impact of various forms of error. It is probably the rare trial nowadays where the possibility of error and the likelihood of its occurrence do not come up at some point. Ideally, multiple types of error-rate information would be manipulated concurrently. RMP and LER have received a fair amount of attention in studies involving DNA tests, but generally not in studies involving other types of scientific tests. One reason that error information is not manipulated more often in research on scientific evidence is that error rates are often difficult to estimate precisely. Thus, more attention also needs to be devoted to estimating the likelihood of different types of error associated with scientific tests. Other forms of error could be examined as well. One form of test "error" that has not been examined is a "match" that results from using a sample that was planted or tampered with by the police.

A fourth recommendation is for jury scholars to conduct more research on some under-studied forms of scientific evidence that are frequently presented at trial. These include various forms of comparison evidence (e.g., fingerprints, hair, and shoeprints) as well as chemical tests that assess the level of various substances in the body (e.g., alcohol, drugs, and toxins). Moreover, given the ever-increasing understanding of the human genome, research is needed on the impact of presenting jurors with information about specific genes or gene combinations possessed by the defendant. Research on non-DNA genetic evidence has only just begun, with existing work focused primarily on the presence/absence of one particular gene—*monoamine oxidase A* (i.e., Applebaum and Scurich 2014; Applebaum et al. 2015;

Saks et al. 2014). Going forward, carefully controlled experiments could be used to address the impact of different behavioral effects (or claimed effects) associated with various genes, the wording used to express those effects, and the claimed amount of empirical support that underlies them.

A fifth recommendation is for jury scholars to systematically examine decision aids that could be used at trial to assist jurors. Little attention has been devoted thus far to examining aids that could help jurors to process complex scientific evidence. Of particular interest are innovations that could provide scientific information in a more accessible fashion, assess jurors' understanding of it, and help jurors work with statistical probabilities. For example, multimedia tutorials could be created by court-appointed expert witnesses and shown to jurors (e.g., Goodman-Delahunty and Hewson 2010). "Clickers" or other real-time measurement devices could be used to assess jurors' comprehension of the evidence and make sure that it exceeds some threshold of understanding. Decision support software could also be employed to help jurors calculate and combine probabilities in a mathematically appropriate fashion. It will likely be difficult to persuade judges to allow the use of such decision aids at trial, but the existence of empirical data to support claims of their efficacy would certainly help make the case for them.

Finally, jury scholars studying scientific evidence should incorporate group deliberation more often into their research designs. Only a handful of studies thus far have gathered data from interacting groups or focused on actual jury verdicts (e.g., Briody 2004; Kaye et al. 2007; Thompson et al. 2013), but the deliberation process may be especially important in trials that feature scientific evidence. Jurors will likely differ in terms of their trust in scientific evidence and their ability to reason from probabilistic evidence. High levels of within—jury diversity will create situations where jurors have divergent views regarding the probative value of the scientific evidence and group discussion is needed to reconcile them. The story sampling model (Devine 2012) predicts jurors will vary in terms of the amount and nature of their spoken contributions during deliberation as well as the amount of influence they exert on other jurors. This model calls attention to the timing and quality of contributions that concern the scientific test results, and the ability of jurors to incorporate the scientific evidence into a larger narrative. In addition to examining the robustness of effects observed at the individual level, research designs that involve interacting groups allow for the examination of composition effects associated with the distribution of key individual-difference variables within the jury.

Summary

Scientific evidence arises from application of the scientific method. It comes in many forms and this chapter has been focused on those intended to provide diagnostic conclusions about the identity or characteristics of target individuals in legal settings. A number of theories have implications for understanding how jurors

process scientific evidence at trial, and they collectively underscore the importance of cognitive mechanisms such as mental models and narrative stories. Considerable research has now been conducted on various types of scientific evidence and a number of findings have emerged. Mock jurors have trouble understanding scientific evidence and do not use it in a manner consistent with the Bayesian model, although they may perform more "rationally" under some conditions and with the benefit of some trial innovations. The presence of a scientific test result, whether incriminating or exonerating, tends to have a strong effect but neither result is necessarily decisive. Mock jurors are influenced by the acknowledgement that error is possible, the provision of a specific likelihood that it will occur, and the statistical frame used to convey the numeric value. Conversely, they tend to be fairly insensitive to the magnitude of error likelihood values. A number of variables associated with the case and the jurors appear likely to affect juror judgments in trials featuring scientific evidence, and may also serve to moderate the impact of scientific evidence. These characteristics include the strength of the non-scientific evidence as well as jurors' trust in scientific evidence and their ability to work with (and reason from) numbers. Following our review, we offered an integrative conceptual framework and a set of recommendations calling for research that jointly examines characteristics of the scientific evidence, the non-scientific evidence, and the jurors who must evaluate it. Ultimately, we hope this chapter will stimulate research on scientific evidence and focus it on those questions most in need of scholarly attention.

References

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Applebaum, P. S., & Scurich, N. (2014). Impact of behavioral genetic evidence on the adjudication of criminal behavior. *Journal of the American Academy of Psychiatry and the Law, 42*, 91–100.

Applebaum, P. S., Scurich, N., & Raad, R. (2015). Effects of behavioral genetic evidence on perceptions of criminal responsibility and appropriate punishment. *Psychology, Public Policy, and Law, 21*, 134–144. doi:10.1037/law0000039

Ask, K., Reinhard, M.-A., Marksteiner, T., & Granhag, P. A. (2011). Elasticity in evaluations of criminal evidence: Exploring the role of cognitive dissonance. *Legal and Criminological Psychology, 16*, 289–306. doi:10.1348/135532510X510153

Baskin, D. R., & Sommers, I. B. (2010). Crime-show-viewing habits and public attitudes toward forensic evidence: The "CSI effect" revisited. *The Justice System Journal, 31*, 97–113.

Bennett, W. L. (1978). Storytelling in criminal trials: A model of social judgment. *Quarterly Journal of Speech, 64*, 1–22.

Bennett, W. L., & Feldman, M. S. (1981). *Reconstructing reality in the courtroom: Justice and judgment in American culture*. New Brunswick, NJ: Rutgers University Press.

Briody, M. (2004). The effects of DNA evidence on homicide cases in court. *Australian and New Zealand Journal of Criminology, 37*, 231–252. doi:10.1375/acri.37.2.231

Britton, R. L. (1997). *Jurors' sensitivity to variations in the diagnosticity and reliability of DNA typing evidence*. Unpublished doctoral dissertation, University of California, Irvine.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*, 197–253.

Carlson, S. C., Pasano, M. S., & Jannuzzo, J. A. (1977). The effect of lie detector evidence on jury deliberations: An empirical study. *Journal of Police Sciences and Administration, 5*, 148–154.

Cavoukian, A., & Heselgrave, R. J. (1980). The admissibility of polygraph evidence in court: Some empirical findings. *Law and Human Behavior, 4*, 117–131.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology, 39*, 752–766.

Chin, J. M. (2014). Psychological science's replicability crisis and what it means for science in the courtroom. *Psychology, Public Policy, and Law, 20*, 225–238. doi:10.1037/law0000012

Clancy, D., & Bull, R. (2015). The effect on mock-juror decision-making of power-of-speech within eyewitness testimony and types of scientific evidence. *Psychiatry, Psychology and Law, 22*, 425–435. doi:10.1080/13218719.2014.960029

Dann, B. M., Hans, V. P., & Kaye, D. H. (2004). *Testing the effects of selected jury trial innovations on juror comprehension of contested mtDNA evidence: Final technical report*. National Institute of Justice, US Department of Justice. Retrieved from https://www.ncjrs.gov/pdffiles1/nij/grants/211000.pdf

Dartnall, S., & Goodman-Delahunty, J. (2006). Enhancing juror understanding of probabilistic DNA evidence. *Australian Journal of Forensic Sciences, 38*, 85–96. doi:10.1080/00450610609410635

Daubert v. Merrell Dow Pharmaceuticals, Inc. 509 U.S. 579–595 (1993).

Devine, D. J. (2012). *Jury decision making: The state of the science*. New York: New York University Press.

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review, 70*, 192–242.

Einhorn, H., Kleinmuntz, D., & Kleinmuntz, B. (1979). Linear regression and process-tracing models of judgment. *Psychological Review, 86*, 465–485.

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: Development of the subjective numeracy scale. *Medical Decision Making, 27*, 672–680. doi:10.1177/0272989X07304449

Faigman, D. L., & Baglioni, A. J, Jr. (1988). Bayes' theorem in the trial process: Instructing jurors on the value of statistical evidence. *Law and Human Behavior, 12*, 1–17.

Federal Rules of Evidence. (2015). Retrieved from https://www.law.cornell.edu/rules/fre

Frye v. United States, 29 F.1013 (1923).

Garrett, B., & Mitchell, G. (2013). How jurors evaluate fingerprint evidence: The relative importance of match language, method information, and error acknowledgment. *Journal of Empirical Legal Studies, 10*, 484–511. doi:10.1111/jels.12017

General Electric Co., v Joiner, 522 U.S. 136 (1997).

Golding, J. M., Stewart, T. L., Yozwiak, J. A., Djadali, Y., & Sanchez, R. P. (2000). The impact of DNA evidence in a child sexual assault trial. *Child Maltreatment, 5*, 373–383. doi:10.1177/1077559500005004009

Goodman, J. (1992). Jurors' comprehension and assessment of probabilistic evidence. *American Journal of Trial Advocacy, 16*, 361–389.

Goodman-Delahunty, J., & Hewson, L. (2010). Enhancing fairness in DNA jury trials. *Trends and Issues in Crime and Criminal Justice, 392*, 1–6.

Greene, E., & Cahill, B. S. (2012). Effects of neuroimaging evidence on mock juror decision making. *Behavioral Sciences & Law, 30*, 280–296. doi:10.1002/bsl.1993

Gurley, J. R., & Marcus, D. K. (2008). The effects of neuroimaging and brain injury on insanity defenses. *Behavioral Sciences & the Law, 26*, 85–97. doi:10.1002/bsl.797

Hans, V. P., Kaye, D. H., Dann, B. M., Farley, E. J., & Albertson, S. (2011). Science in the jury box: Jurors' comprehension of mitochondrial DNA evidence. *Law and Human Behavior, 35*, 60–71. doi:10.1007/s10979-010-9222-8

Hastie, R. (1993). Introduction. In R. Hastie (Ed.), *Inside the juror: The psychology of juror decision making* (pp. 2–41). New York: Cambridge University Press. doi:10.1017/CBO9780511752896

Hayes, R. M., & Levett, L. M. (2013). Community members' perceptions of the CSI effect. *American Journal of Criminal Justice, 38*, 216–235. doi:10.1007/s12103-012-9166-2

Hayes-Smith, R. M., & Levett, L. M. (2011). Jury's still out: How television and crime show viewing influences jurors' evaluations of evidence. *Applied Psychology in Criminal Justice, 7*, 29–46.

Holmgren, J. A., & Fordham, J. (2011). The CSI effect and the Canadian and the Australian jury. *Journal of Forensic Sciences, 56*, S63–S71. doi:10.1111/j.1556-4029.2010.01621.x

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13*, 295–355.

Houck, M. M. (2006). CSI: Reality. *Scientific American, 295*, 84–89.

Jenkins, G., & Schuller, R. A. (2007). The impact of negative forensic evidence on mock jurors' perceptions of a trial of drug-facilitated sexual assault. *Law and Human Behavior, 31*, 369–380. doi:10.1007/s10979-006-9068-2

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Kaasa, S. O., Peterson, T., Morris, E. K., & Thompson, W. C. (2007). Statistical inference and forensic evidence: Evaluating a bullet lead match. *Law and Human Behavior, 31*, 433–447. doi:10.1007/s10979-006-9074-4

Kaye, D. H., Hans, V. P., Dann, B. M., Farley, E., & Albertson, S. (2007). Statistics in the jury box: How jurors respond to mitochondrial DNA match probabilities. *Journal of Empirical Legal Studies, 4*, 797–834. doi:10.1111/j.1740-1461.2007.00107.x

Kaye, D. H., & Koehler, J. J. (1991). Can jurors understand probabilistic evidence? *Journal of the Royal Statistical Society, 154*, 75–81.

Khoshbin, L. S., & Khoshbin, S. (2007). Imaging the mind, minding the image: An historical introduction to brain imaging and the law. *American Journal of Law and Medicine, 33*, 171–192. doi:10.1177/009885880703300202

Kim, Y. S., Barak, G., & Shelton, D. E. (2009). Examining the CSI-effect in the cases of circumstantial evidence and eyewitness testimony: Multivariate and path analyses. *Journal of Criminal Justice, 37*, 452–460. doi:10.1016/j.jcrimjus.2009.07.005

Koehler, J. J. (1996). On conveying the probative value of DNA evidence: Frequencies, likelihood ratios and error rates. *University of Colorado Law Review, 67*, 859–886.

Koehler, J. J. (2001). When are people persuaded by DNA match statistics? *Law and Human Behavior, 25*, 493–513. doi:10.1023/A:1012892815916

Koehler, J. J. (2011). If the shoe fits they might acquit: The value of forensic science testimony. *Journal of Empirical Legal Studies, 8*, 21–48. doi:10.1111/j.1740-1461.2011.01225.x

Koehler, J. J., Chia, A., & Lindsey, J. S. (1995). The random match probability (RMP) in DNA evidence: Irrelevant and prejudicial? *Jurimetrics Journal, 35*, 201–209.

Koehler, J. J., & Macchi, L. (2004). Thinking about low-probability events: An exemplar cuing theory. *Psychological Science, 15*, 540–546. doi:10.1111/j.0956-7976.2004.00716.x

Koffler, J. (1957). The lie detector: A critical appraisal of the technique as a potential undermining factor in the judicial process. *New York Law Forum, 3*, 123–158.

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

Lieberman, J. D., Carrell, C. A., Miethe, T. D., & Krauss, D. A. (2008). Gold versus platinum: Do jurors recognize the superiority and limitations of DNA evidence compared to other types of forensic evidence? *Psychology, Public Policy, and Law, 14*, 27–62. doi:10.1037/1076-8971.14.1.27

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making, 21*, 37–44.

Maeder, E. M., & Corbett, R. (2015). Beyond frequency: Perceived realism and the CSI effect. *Canadian Journal of Criminology and Criminal Justice, 57*, 83–114. doi:10.3138/cjccj.2013.E44

Mancini, D. E. (2011). The CSI effect reconsidered: Is it moderated by need for cognition? *North American Journal of Psychology, 13*, 155–174.

Markwart, A., & Lynch, B. E. (1979). The effect of polygraph evidence on mock jury decision-making. *Journal of Police Science and Administration, 7*, 324–332.

Martire, K. A., Kemp, R. I., Sayle, M., & Newell, B. R. (2014). On the interpretation of likelihood ratios in forensic science evidence: Presentation formats and the weak evidence effect. *Forensic Science International, 240*, 61–68. doi:10.1016/j.forsciint.2014.04.005

Martire, K. A., Kemp, R. I., Watkins, I., Sayle, M. A., & Newell, B. R. (2013). The expression and interpretation of uncertain forensic science evidence: Verbal equivalence, evidence strength, and the weak evidence effect. *Law and Human Behavior, 37*, 197–207. doi:10.1037/lhb0000027

McCabe, D. P., Castel, A. D., & Rhodes, M. G. (2011). The influence of fMRI lie detection evidence on juror decision-making. *Behavioral Sciences & Law, 29*, 566–577. doi:10.1002/bsl.993

McQuiston-Surrett, D., & Saks, M. J. (2009). The testimony of forensic identification science: What expert witnesses say and what factfinders hear. *Law and Human Behavior, 33*, 436–453. doi:10.1007/s10979-008-9169-1

Myers, B., & Arbuthnot, J. (1997). Polygraph testimony and juror judgments: A comparison of the guilty knowledge test and the control question test. *Journal of Applied Social Psychology, 27*, 1421–1437. doi:10.1111/j.1559-1816.1997.tb01606.x

Myers, B., Latter, R., & Abdollahi-Arena, M. K. (2006). The court of public opinion: Lay perceptions of polygraph testing. *Law and Human Behavior, 30*, 509–523. doi:10.1007/s10979-006-9041-0

Myers, B., Rosol, A., & Boelter, E. (2003). Polygraph evidence: The impact of corroborating evidence on guilt judgments. *Journal of Applied Social Psychology, 33*, 948–962. doi:10.1111/j.1559-1816.2003.tb01933.x

Nance, D. A., & Morris, S. B. (2002). An empirical assessment of presentation formats for trace evidence with a relatively large and quantifiable random match probability. *Jurimetrics, 42*, 403–448.

Nance, D. A., & Morris, S. B. (2005). Juror understanding of DNA evidence: An empirical assessment of presentation formats for trace evidence with a relatively small random-match probability. *Journal of Legal Studies, 34*, 395–444.

National Research Council. (1996). *The evaluation of forensic DNA evidence*. Washington, DC: National Academic Press. Retrieved from http://www.nap.edu/catalog/5141/the-evaluation-of-forensic-dna-evidence

Niedermeier, K. E., Kerr, N. L., & Messé, L. A. (1999). Jurors' use of naked statistical evidence: Exploring bases and implications of the Wells effect. *Journal of Personality and Social Psychology, 76*, 533–542. doi:10.1037/0022-3514.76.4.533

Pennington, N., & Hastie, R. (1981). Juror decision-making models: The generalization gap. *Psychological Bulletin, 89*, 246–287.

Pennington, N., & Hastie, R. (1993). The story model for juror decision making. In R. Hastie (Ed.), *Inside the juror: The psychology of juror decision making* (pp. 192–221). New York: Cambridge University Press.

Petty, R. E., & Cacioppo, J. T. (1986a). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.

Petty, R. E., & Cacioppo, J. T. (1986b). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 123–205). New York: Academic Press.

Podlas, K. (2006). The CSI effect: Exposing the media myth. *Fordham Intellectual Property, Media and Entertainment Law Journal, 16*, 429–465.

Pozzulo, J. D., Lemieux, J. M. T., Wilson, A., Crescini, C., & Girardi, A. (2009). The influence of identification decision and DNA evidence on juror decision making. *Journal of Applied Social Psychology, 39*, 2069–2088. doi:10.1111/j.1559-1816.2009.00516.x

Read, S. J., & Miller, L. C. (1998). *Connectionist models of social reasoning and social behavior*. Mahwah, NJ: Erlbaum.

Reiser, L. A. (2015). *The impact of narrative consistency on jurors' utilization of DNA evidence*. Master's thesis. Retrieved from ProQuest Dissertations and Theses Global. http://search.proquest.com.proxy.ulib.uits.iu.edu/pqdtglobal/docview/1738628885/ED7C83A0F5D447E9PQ/2?accountid=7398

Rhodes, R. E., Rodriguez, F., & Shah, P. (2014). Explaining the alluring influence of neuroscience information on scientific reasoning. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 40*, 1432–1440. doi:10.1037/a0036844

Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin, 100*, 349–363.

Saks, M. J., Schweitzer, N. J., Aharoni, E., & Kiehl, K. A. (2014). The impact of neuroimages in the sentencing phase of capital trials. *Journal of Empirical Legal Studies, 11*, 105–131. doi:10.1111/jels.12036

Schklar, J., & Diamond, S. S. (1999). Juror reactions to DNA evidence: Errors and expectancies. *Law and Human Behavior, 23*, 159–184. doi:10.1023/A:1022368801333

Schweitzer, N. J., & Saks, M. J. (2007). The CSI effect: Popular fiction about forensic science affects the public's expectations about real forensic science. *Jurimetrics, 47*, 357–364.

Schweitzer, N. J., & Saks, M. J. (2011). Neuroimage evidence and the insanity defense. *Behavioral Sciences & the Law, 29*, 592–607. doi:10.1002/bsl.995

Schweitzer, N. J., Saks, M. J., Murphy, E. R., Roskies, A. L., Sinnott-Armstrong, W., & Gaudet, L. M. (2011). Neuroimages as evidence in a mens rea defense: No impact. *Psychology, Public Policy, and Law, 17*, 357–393. doi:10.1037/a0023581

Scurich, N. (2015). The differential effect of numeracy and anecdotes on the perceived fallibility of forensic science. *Psychiatry, Psychology and Law, 22*, 616–623. doi:10.1080/13218719.2014.965293

Scurich, N., & John, R. S. (2011). Trawling genetic databases: When a DNA match is *just* a naked statistic. *Journal of Empirical Legal Studies, 8*, 49–71. doi:10.1111/j.1740-1461.2011.01231.x

Scurich, N., & John, R. S. (2013). Mock jurors' use of error rates in DNA database trawls. *Law and Human Behavior, 37*, 424–431. doi:10.1037/lhb0000046

Shelton, D. E., Kim, Y. S., & Barak, G. (2006). A study of juror expectations and demands concerning scientific evidence: Does the CSI effect exist? *Vanderbilt Journal of Entertainment and Technology Law, 9*, 331–368.

Simon, D. (2012). *In doubt: The psychology of criminal justice processes*. Cambridge, MA: Harvard University Press.

Simon, D., Snow, C. J., & Read, D. J. (2004). The redux of cognitive consistency theories: Evidence judgments by constraint satisfaction. *Journal of Personality and Social Psychology, 86*, 814–837. doi:10.1037/0022-3514.86.6.814

Smith, B. C., Penrod, S. D., Otto, A. L., & Park, R. C. (1996). Jurors' use of probabilistic evidence. *Law and Human Behavior, 20*, 49–82. doi:10.1007/BF01499132

Smith, L. L., & Bull, R. (2012). Identifying and measuring juror pre-trial bias for forensic evidence: Development and validation of the forensic evidence evaluation bias scale. *Psychology, Crime & Law, 18*, 797–815. doi:10.1080/1068316X.2011.561800

Smith, L. L., & Bull, R. (2014). Validation of the factor structure and predictive validity of the Forensic Evidence Evaluation Bias Scale for robbery and sexual assault trial scenarios. *Psychology, Crime & Law, 20*, 450–466. doi:10.1080/1068316X.2013.793340

Smith, L. L., Bull, R., & Holliday, R. (2011). Understanding juror perceptions of forensic evidence: Investigating the impact of case context on perceptions of forensic evidence strength. *Journal of Forensic Sciences, 56*, 409–414. doi:10.1111/j.1556-4029.2010.01671.x

Spanos, N. P., Myers, B., Dubreuil, S. C., & Pawlak, A. E. (1992–1993). The effects of polygraph evidence and eyewitness testimony on the beliefs and decisions of mock jurors. *Imagination, Cognition and Personality*, *12*, 103–113. doi:10.2190/1AB2-3WLX-BFY1-8YCP

Thompson, W. C. (1989). Are juries competent to evaluate statistical evidence? *Law and Contemporary Problems, 52*, 9–41.

Thompson, W. C., Kaasa, S. O., & Peterson, T. (2013). Do jurors give appropriate weight to forensic identification evidence? *Journal of Empirical Legal Studies, 10*, 359–397. doi:10.1111/jels.12013

Thompson, W. C., & Newman, E. J. (2015). Lay understanding of forensic statistics: Evaluation of random match probabilities, likelihood ratios, and verbal equivalents. *Law and Human Behavior, 39*, 332–349. doi:10.1037/lhb0000134

Thompson, W. C., & Schumann, E. L. (1987). Interpretation of statistical evidence in criminal trials: The prosecutors' fallacy and the defense attorney's fallacy. *Law and Human Behavior, 11*, 167–187. doi:10.1007/BF01044641

Tryon, W. W. (2012). A connectionist network approach to psychological science: Core and corollary principles. *Review of General Psychology, 12*, 305–317. doi:10.1037/a0027135

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207–232.

Weisberg, P. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2008). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience, 20*, 470–477.

Wells, G. L. (1992). Naked statistical evidence of liability: Is subjective probability enough? *Journal of Personality and Social Psychology, 62*, 739–752. doi:10.1037/0022-3514.62.5.739

West, M. L., Lawson, V. Z., & Grose-Fifer, J. (2014). The effect of electrophysiological neuroscientific deception detection evidence on juror judgments in a criminal trial. *Basic and Applied Social Psychology, 36*, 133–144. doi:10.1080/01973533.2014.881288

Wyer, R. S, Jr. (2007). Principles of mental representation. In A. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (2nd ed., pp. 285–307). New York: Guilford Press.

Zikmund-Fisher, B. J., Smith, D. M., Ubel, P. A., & Fagerlin, A. (2007). Validation of the subjective numeracy scale: Effects of low numeracy on comprehension of risk communications and utility elicitations. *Medical Decision Making, 27*, 663–671. doi:10.1177/0272989X07303824

The Purpose, Content, and Effects of Expert Testimony on Interrogations and Confessions

Mark Costanzo, Iris Blandón-Gitlin and Deborah Davis

The ability of police to question criminal suspects is an indispensable investigative tool. Without the information and admissions gained through the process of interrogation, many guilty criminals would go free. Unfortunately, the powerful psychological techniques used to elicit true confessions from guilty suspects can sometimes elicit false confessions from innocent suspects. Although it is important to give police the tools they need to solve crimes, it is also important to protect the rights of the accused and to reach just outcomes.

A confession, true or false, is often enough to send a defendant to prison. The power of a confession has been demonstrated through both laboratory and field research. For example, in a study of 125 proven false confessions, Drizin and Leo (2004) found that when suspects falsely confessed, then recanted, pled "not guilty," and proceeded to trial, the conviction rate was 81 %. Research on more than 300 convictions overturned by DNA evidence reveals that about 24 % of wrongful convictions involve false admissions. In many of these wrongful conviction cases, a false confession was the primary or sole evidence against the defendant (www.innocenceproject.org, 2014). Controlled experiments also confirm the powerful impact of confessions. When researchers systematically compared eyewitness, character, and confession evidence, they found that confessions produced the highest conviction rate (Kassin and Neumann 1997). Furthermore, even when mock jurors recognize that a confession was coerced, they are not able to discount it when reaching a verdict (Kassin and Sukel 1997). In sum, a confession is extremely

M. Costanzo (✉)
Claremont McKenna College, Claremont, USA
e-mail: Mark.Costanzo@ClaremontMcKenna.edu

I. Blandón-Gitlin
California State University, Fullerton, Fullerton, USA

D. Davis
University of Nevada, Reno, Reno, USA

B.H. Bornstein and M.K. Miller (eds.), *Advances in Psychology and Law*,
Advances in Psychology and Law 2, DOI 10.1007/978-3-319-43083-6_5

difficult for jurors to discount or ignore, even if that confession is false, even if it is coerced, and even when there is little corroborating evidence.

In this chapter, we argue that the use of expert testimony on interrogation and confessions is one effective means of assisting jurors and judges in their duty to find the truth and reach just verdicts. We review the research literature on police interrogations and false confessions, with an emphasis on the usefulness of this research for providing expert testimony at trial. Because all three of the authors have testified as expert witnesses on confessions, we also offer insights from our experiences in court.

True, False, and Ambiguous Confessions

A confession can be true, false, or ambiguous in veracity. We can often be sure that a criminal confession is true because the information revealed by the confessor is strongly supported by other evidence in the case. For example, there is often medical evidence, or DNA evidence, or fingerprints, or video recordings, or computer records, or multiple, credible eyewitnesses that confirm the account of the crime provided by a suspect in the interrogation room. True confessions disclose accurate details about the crime, often lead to the discovery of new evidence (e.g., where a body is buried), and reveal facts and information only the true perpetrator would know.

We can know for certain that a confession is false when DNA evidence proves that someone else committed the crime, or when it can be shown that it was physically impossible for the confessor to have committed the crime (e.g., the suspect was in a distant location at the time of the crime). False confessions are generally not supported by other reliable evidence in the case. Usually, an account of a crime given by an innocent false confessor will not match the facts of the crime, will contain substantial errors, and will not lead police to new or missing evidence.

A significant problem in distinguishing true from false confessions is created by the process of "contamination." Too often, interrogators leak key details of the crime to the suspect being interrogated (Leo et al. 2009). Prior to or during a long interrogation, police may disclose several critical elements of the crime (e.g., how and where the victim was killed). When a suspect later incorporates these key details into his confession, the result is a contaminated false confession that has the appearance of veracity. Because of the process of contamination, many proven false confessions contain details that only the actual perpetrator should know (Garrett 2010). In the absence of an electronic recording of the entire interrogation, it is usually impossible to assess whether contamination occurred during the interrogation process.

Prior to the advent of DNA identification technology, it was thought that false confessions were extremely rare. Many observers assumed that false confessions would only come from suspects who were mentally impaired or physically abused. We now know that false confessions are far more common than previously supposed. More than 320 wrongful convictions have been uncovered using DNA

evidence, and as mentioned earlier, a surprising number of these convictions have been shown to involve false confessions (www.innocenceproject.org, 2014). The *National Registry of Exonerations* (which includes both DNA and non-DNA-based exonerations) lists 192 cases involving false confessions (www.law.umich.edu/special/exoneration, 2014). As many researchers have noted, these known false confessions are merely "the tip of a much larger iceberg" (Drizin and Leo 2004). The actual number of false confessions is likely to be substantially higher. In part, this is because the biological evidence that permits DNA testing and proves innocence is available in only a small minority (perhaps 10 %) of criminal cases (Scheck et al. 2000).

Some confessions are ambiguous in veracity—they can be proven neither true nor false. If there is a confession, but no other compelling evidence of guilt, it may be impossible to determine with certainty whether the confession is true. Confessions are a powerful form of incriminating evidence and are often sufficient to convict a defendant. Prisoners convicted of crimes may continue to proclaim their innocence from behind bars, but unless strong exonerating evidence is later uncovered, we may never be able to verify whether a confessor is innocent or guilty. Because some wrongful convictions are never discovered, we cannot be sure which names should be added to the list of false confessors.

It is currently impossible to calculate the frequency or rate of false confessions. First, there is the problem of ambiguous confessions that cannot be classified as definitely true or definitely false. There are also methodological obstacles—police departments and state governments do not keep precise records of the number of interrogations conducted annually, nor do they keep detailed records of which confessions were later proven to be true or false. Furthermore, researchers have not been granted access to the records that do exist.

The Admissibility and Purpose of Expert Testimony on Confessions

Like other forms of expert testimony, the admissibility of expert testimony on confessions is governed by the *Frye* standard (*Frye v. United States* 1923), the *Daubert* standard (*Daubert v. Merrell Dow Pharmaceuticals* 1993), and Rule 702 of the Federal Rules of Evidence. Several jurisdictions still rely on the older *Frye* standard, which holds that expert testimony is admissible if it is "generally accepted" in the relevant scientific community. The newer *Daubert* standard suggested several considerations for determining admissibility (e.g., falsifiability, peer review, error rate, and general acceptance) but also emphasized that there is no "definite checklist" for determining the admissibility of scientific evidence. Much has been written about these legal standards and we refer the reader to several fuller discussions of the legal complexities of admissibility standards (e.g., Cutler et al. 2014; Fulero 2004). However, it should be noted that extralegal considerations also

influence which experts are permitted to testify. For example, in an analysis of 1600 written admissibility decisions, Risinger (2000) found a strong bias in favor of allowing prosecution experts to testify and against allowing defense experts to testify.

The legal authority for allowing expert testimony on interrogations and confessions comes from the United States Supreme Court decision in *Crane v. Kentucky* (1986). The Court reviewed a case involving the murder of a liquor store clerk in Louisville, Kentucky. Because of a lack of physical evidence, police were initially unable to identify a suspect. However, a week after the murder, a 16-year-old juvenile named Crane was arrested in connection with the robbery of a local service station. During the course of a long interrogation, Crane confessed to several robberies and the shooting of a police officer. When asked about the liquor store murder, he initially denied involvement, but eventually confessed to that crime as well. Many of the details offered by Crane contradicted other evidence in the case. For, example, although the murder occurred at approximately 10:40 p.m., Crane said it happened during the day. Although no money was missing from the cash register, Crane described taking the money from the cash register. He later recanted his confession and claimed that it had been the result of coercion.

The defense attorney sought to introduce evidence about the coercive circumstances surrounding the confession. Mr. Crane had been held in a small, windowless room, interrogators denied his repeated requests to telephone his mother, he was interrogated for several hours, and a total of six police officers participated in the interrogation. The trial judge first ruled that the confession was "voluntary" and then ruled that information challenging the credibility of the confession could not be presented to the jury. Crane was found guilty and sentenced to 40 years in prison.

Although the Kentucky Supreme Court affirmed the conviction and ruled that the trial judge did not err by excluding expert testimony about "the physical and psychological environment" that led to the confession, the U.S. Supreme Court later reversed and remanded the case. The Court held that, "… certain interrogation techniques, either in isolation, or as applied to the unique characteristics of a particular suspect, are so offensive to a civilized system of justice that they must be condemned under the Due Process Clause of the Fourteenth Amendment." Further, the Court held that,

> …a defendant's case may stand or fall on his ability to convince the jury that the manner in which the confession was obtained casts doubt on its credibility…. stripped of the power to describe to the jury the circumstances that prompted his confession, the defendant is effectively disabled from answering the one question every rational juror needs answered: If the defendant is innocent, why did he previously admit his guilt?

Ten years later, in one of the first cases after Crane to involve expert testimony on confessions (*United States v. Hall* 1996, 1997), the Seventh Circuit Court of Appeals vacated the kidnapping and murder conviction of Larry Hall because the trial court had improperly excluded the testimony of two confessions experts. The first, Dr. Richard Ofshe, was a professor of sociology at the University of

California, Berkeley and a nationally recognized expert on coercive persuasion and false confessions. Dr. Ofshe "would have testified about the fact that experts in his field agree that false confessions exist, that individuals can be coerced into giving false confessions, and that certain indicia can be identified to show when they are likely to occur" (*United States v. Hall*, 1996, pp. 8–9). The second expert, a psychiatrist, "would have testified about Hall's susceptibility to various interrogation techniques and his propensity to give a false confession." The trial judge ruled that such testimony would "invade the prerogative of the jury" and "add nothing to what the jury would know from common experience." In vacating the conviction and remanding the case, the Seventh Circuit found that expert testimony may have been critical for Hall's defense and should have been allowed (*United States v. Hall* 1996). The federal district court subsequently held that "the science of social psychology, and specifically the field involving the use of coercion in interrogations, is sufficiently developed in its methods to constitute a reliable body of specialized knowledge under Rule 702" (United States v. *Hall* 1997, pp. 22–23). However, despite these promising early rulings, there has been considerable variability in how the courts have responded to expert testimony on confessions.

No blanket statement can be made about the admissibility of expert testimony because the decision to admit or exclude testimony is made by individual judges. Admissibility decisions vary between jurisdictions and between judges within jurisdictions. Given that the scientific reliability of research on interrogations and confessions has generally been acknowledged by the courts, the decision to admit or exclude expert testimony usually hinges on a judgment about whether or not such testimony will be helpful to jurors. In considering the nebulous issue of "helpfulness," judges may decide that expert testimony is relevant and that it will assist jurors in their job of understanding the evidence and rendering a fair, fully informed verdict. Alternatively, a judge may decide that an expert will not be allowed to testify because the testimony might confuse jurors or be too influential in the decision-making process of jurors. Perhaps the most common reason for excluding expert testimony is a judgment that what the expert has to say is already part of the common knowledge of ordinary jurors (Cutler et al. 2014).

If a disputed confession is introduced at trial, and the defense argues that the confession is false, jurors will naturally wonder why an innocent person might confess to a crime he did not commit. The proposition that a suspect might falsely confess to a serious crime is strongly counterintuitive (unless that suspect is physically abused by his interrogators or is mentally impaired). The purpose of expert testimony at trial is to help jurors understand the phenomena of false confessions by providing an overview of research on interrogation and confessions. Such testimony is intended to assist the jury in making a fully informed decision about how much weight to place on the defendant's confession. More specifically, expert witnesses can educate the jury by: discussing research documenting the reality of false confessions; explaining how and why particular interrogation techniques can sometimes induce innocent suspects to confess; and identifying factors that increase the risk of a false confession. When an audio or video recording of the interrogation exists, the expert will usually be permitted to review

that recording and to offer an analysis of the interrogation techniques and the nature of the admissions. Here, as in other areas, responsible experts refrain from providing "ultimate opinion" testimony. That is, they do not opine about whether the defendant's confession is true or false (American Psychological Association 2013).

Although expert testimony on confessions is usually offered at trial, it may also be heard by a judge at a pretrial suppression hearing (at which time the judge decides which evidence to allow at trial and which evidence to exclude). Based on that hearing, the judge may reach the conclusion that the confession was coerced and will not be admitted into evidence at trial. The use of expert testimony in pretrial suppression hearings and at jury trials should exert an educative and corrective effect on the future behavior of police and prosecutors (Costanzo and Leo 2008). By exposing flaws, expert testimony in disputed confession cases may indirectly deter misbehavior and improve police and prosecutorial screening practices. It should help to reduce the use of psychologically coercive interrogation methods and decrease the number of false confessions elicited and introduced into evidence at trial. As a consequence, fewer innocent defendants will be wrongfully convicted because of false confessions.

The Content of Expert Testimony on Confessions

To help jurors understand why some suspects falsely confess, experts must try to succinctly review the large body of research on the social–psychological dynamics of interrogations and confessions. The expert's review will often include descriptions of how power and deception are used during the interrogation process, the effect of interrogation techniques on the decision-making process of the suspect, and the factors that increase vulnerability to false confessions. When relevant to the case under consideration, expert testimony may also include descriptions of one or more suspect vulnerabilities (youth, cognitive impairment, mental illness, and cultural/language barriers) that increase the risk of false admissions. Each of these areas is reviewed below.

Before summarizing the research foundation for expert testimony on confessions, it is important to emphasize that such testimony occurs in the rarified, rule-bound setting of the courtroom. The critical feature of expert testimony is that it is presented in response to a series of questions posed by attorneys. Instead of presenting a thorough, well-structured, uninterrupted overview of relevant research for the purpose of assisting the jury, experts must try to convey all relevant knowledge in the form of brief answers to specific questions. Experts cannot answer questions that were not (but should have been) asked. Consequently, the content of testimony is usually far more limited than most experts would like. And, because questioning of the expert occurs in the context of an adversarial system, many questions may be dismissive or even hostile. If, as is usually the case, the expert has been retained by the defense, cross-examination by the prosecutor will be designed to challenge and discredit the expert, the basis of the expert's testimony, and any conclusions reached by the expert.

The Power Asymmetry Between Interrogator and Suspect

Interrogation is an inherently psychological process. During most interrogations, police use power, persuasion, and deception to extract a confession from a reluctant suspect. Although suspects may also attempt to persuade or deceive, they are at an enormous disadvantage. Their main challenge is to withstand the intense, psychologically taxing process of interrogation without making an admission of guilt.

For more than 60 years, social scientists have studied the use of power in social relationships. Researchers have described power as:

> The ability of one individual in a relationship (the influence agent) to exert influence on another person (the target of influence) so that the influence agent obtains the specific outcomes he or she wants in a given situation while being able to resist influence attempts by the target (Simpson et al. 2015).

In the interrogation room, the police officer (the agent) exerts influence on the suspect (the target) with the goal of obtaining a confession while resisting any attempts by the suspect to deny or remain silent.

Most models of interpersonal influence emphasize that the ability to exert influence rests on several "bases of power" (French and Raven 1959; Keltner et al. 2003; Raven et al. 1998). These bases include *reward* power—the ability to provide positive reinforcement; *coercive* power—the ability to punish; *legitimate* power—influence that flows from holding a socially sanctioned role or job; *referent* power—influence that occurs when a target identifies with and admires the influence agent; *expert* power—possession of special knowledge or access that might benefit the target; and *informational* power—specific, useful information that can only be obtained if the target cooperates.

To understand the power of the interrogation process, it is important to recognize that police officers bring nearly every form of power with them into the interrogation room. First, because of their important role in maintaining social order, police enjoy considerable legitimate power. The have the right to make demands of citizens and to expect compliance with those demands. Second, most citizens are taught to trust and rely on police. This assumption of trustworthiness is another source of influence. The belief that police are trustworthy makes it difficult for suspects to challenge lies told by police officers during an interrogation. Of course, although a majority of Americans express trust and confidence in the police, many minority suspects are deeply suspicious of the police (Newport 2014). Third, police have more coercive power than virtually any other members of society. No one else is legally permitted to arrest their fellow citizens or to physically subdue, handcuff, or place suspects in jail. Fourth, police have the power to reward suspects by promising or implying that a suspect will receive more lenient treatment if he or she chooses to cooperate. Finally, police can exploit their power as experts on the legal system. They possess insider knowledge gained through training and experience. Because police are representatives of and experts on the legal system, their assertions about how a suspect will be treated by the legal system must be taken seriously.

Police enter the interrogation room already possessing tremendous power over the suspect. Aside from military or prison settings, there are few places where the power disparity is so great. Even before the interrogation process begins, and before police officers deploy the potent interrogation tactics at their disposal, the suspect is already at an extreme power disadvantage.

The Foundational Lie of the Interrogation Process

The interrogation process can only be understood in the context of our adversarial legal system. As one prominent scholar has noted, "Police interrogators view themselves as and act like agents of the prosecution, and thus the suspect's adversary, not as neutral fact investigators" (Leo 2008, p. 120). In the American system, police serve the interests of prosecutors, and the goal of an interrogation is to elicit an admission of guilt that will facilitate prosecution. However, in the interrogation room, police labor to create the illusion that they can and will function as the suspect's allies in the criminal justice system. Indeed, the foundational lie told by police to crime suspects during the interrogation is that "we're going to help you improve your situation." Put differently, police misrepresent their role as allies instead of adversaries, and suggest that they can and will help the suspect minimize the negative consequences of his crime. Without this foundational lie, the interrogation tactics described below could not be nearly so effective.

As the interrogation begins, police often misrepresent the interrogation as an "opportunity" for the suspect to help himself. Specifically, the suspect is offered an opportunity to "tell your side of the story." This ploy diverts attention from the true purpose of interrogation: to elicit incriminating statements. It also suggests that talking to the police may be the only way for the suspect to help himself—by "telling his side of the story" he might be able to convince the police that he was not involved in the crime, or that he was minimally involved, or that his involvement was justified. The police may also imply that the suspect's version of events can only become part of the official record if he speaks to police; and that unless he tells his story, they will (reluctantly) have no choice but to proceed with the case using only the incriminating evidence they have already collected. In effect, police want the suspect to believe that he is being offered a precious, time-limited opportunity to clear his name and to shape the version of events that will be presented to a prosecutor, judge, or jury. Alternatively, if the suspect refuses to talk, the opportunity will vanish.

During the interrogation, police officers may imply that they will use their discretionary power to help the suspect and influence his case. They can offer to serve as his allies in the legal system by writing up a favorable report, by influencing the prosecutor before charges are filed, by talking to the judge, or by testifying at trial in a way that depicts the suspect in a sympathetic light (e.g., he had understandable reasons for committing the crime and he is now remorseful). Sometimes interrogators explicitly offer to assist a distressed suspect by arranging

for psychological counseling or help from social services. The catch, of course, is that these benefits can only be delivered if the suspect admits to committing a crime (Leo et al. 2008).

Police Interrogation Techniques

Everyday, police conduct interviews with criminal suspects and witnesses. The goal of these interviews is to gather information about crimes and to identify likely perpetrators. What distinguishes an interrogation from an interview is the goal of an interrogation: to induce a confession. In an interrogation, police have already reached the conclusion that the suspect is guilty of the crime. Consequently, interrogations, unlike interviews, tend to be confrontational and emotionally intense.

When criminal suspects are held in custody for the purpose of interrogation, police must inform them of their *Miranda* rights: the right to remain silent, the right to appointed counsel, and a warning that anything they say can be used against them in court (*Miranda v. Arizona* 1966). These warnings, originally intended to blunt police use of coercion, have had much less impact than is commonly assumed. A large majority of suspects (roughly 80 %) waive their rights and submit to interrogation (Leo 2001). Police use a variety of strategies to encourage suspects to waive their rights, including reading the *Miranda* warnings in a rushed, confusing, or perfunctory manner (Domanico et al. 2012). The warnings are presented as a meaningless bureaucratic ritual. By minimizing the significance of the warnings, police disguise the adversarial nature of interrogation (Leo 2008; Scherr and Madon 2013).

By far, the most influential of all modern interrogation methods is the Reid technique (Inbau et al. 2013). According to Reid and Associates, Inc, more than half a million interrogators have been trained to use the technique since 1974 (www.reid.com, 2015). The tactics advocated by Reid have become so pervasive that even interrogators who have never received Reid training are still likely to use Reid tactics. These tactics have been incorporated into many police department training manuals and have spread through communication between police interrogators.

Behavioral Detection of Deception. The underlying principles of the technique are described below. However, it is important to note that the Reid technique includes training on how to detect lying by observing and decoding the verbal and nonverbal behavior of the suspect. Unfortunately, the behavioral cues interrogators are taught to detect are flawed indicators of deception (Bond and DePaulo 2006). Scores of studies have examined whether people can detect when others are lying or telling the truth. The general finding is that both untrained observers and people with relevant professional training (e.g., police, psychiatrists, polygraph examiners) perform only slightly better than chance (Bond and DePaulo 2006, 2008). Whereas chance accuracy is 50 %, average accuracy is about 54 %.

Interestingly, training to improve accuracy does not reliably improve the ability to detect lies, but it does have effects: after training, people are more confident about their judgments, and they list more reasons for their judgments (Costanzo 1992; Kassin and Fong 1999). For example, Kassin et al. (2005) had prison inmates give true and false videotaped confessions to crimes they did or did not commit. Next, police detectives and college students judged the truthfulness of these statements. College students performed better than chance at distinguishing true from false confessions, but police detectives did not. However, detectives were significantly more confident about the accuracy of their judgments, and showed a bias toward judging false confessions as true. The tendency to infer guilt increased with job experience and interrogation training.

Police estimates of their own skill at detecting lies are at odds with the available research. In a large survey, police estimated that they could distinguish between truthful and deceptive statements from suspects at about a 77 % rate of accuracy (Kassin et al. 2007). This overconfidence can have powerful consequences in the interrogation room. Once a police officer concludes that a suspect's behavior is indicative of lying (instead of the nervousness and anxiety created by the interrogation), the psychological pressure applied to the suspect is likely to escalate. The ambiguous nonverbal behavior of the suspect is likely to be interpreted as further evidence of guilt, as are his verbal denials. Confronted with what they perceive as a clearly guilty suspect who is stubbornly maintaining his innocence, police may redouble their efforts to elicit a confession. The list of proven false confessions includes several cases in which innocent suspects were subjected to coercive interrogation techniques based on little more than intuition or a misreading of a suspect's behavior as suspicious or deceptive (Innocence Project 2015).

Misconceptions about lie detection ability also extend to jurors (Costanzo et al. 2010). In a large-scale study of jurors, 53 % believed that police interrogators are better than ordinary people at identifying lies (only 25 % disagreed), and 60 % believed that interrogators' ability to detect lies improves with experience (only 17 % disagreed). At trial, jurors' misplaced confidence in the lie detection abilities of police may cause them to give too much weight to confident but mistaken police judgments about the deceptiveness of a defendant.

Underlying the Reid approach (which formally consists of nine steps) are four basic techniques: control, social isolation, certainty of guilt, and minimization of culpability. Each of these techniques is described below.

Control. In the interrogation room, both the physical environment and the nature of the interaction are controlled by the police. Suspects are typically interrogated in a small, sterile, windowless room furnished with only a small table and two or three chairs. Usually the suspect is seated with his back near the rear of the room, with the police officer (or officers) seated between him and the door. There is little space and no distractions. The temperature may be uncomfortably cold or hot. Although the room is the "home turf" of the police officer, it is a strange, anxiety-provoking environment for the suspect. The character of the interrogation room has been carefully considered—Interrogation manuals include recommendations on even

small aspects of the physical environment (e.g., size of room, type of chairs, furniture arrangement).

The social interactions that occur during the interrogation are also heavily controlled by the police officers. The direction, content, pacing, and length of the conversation are all under the management of the interrogator. Interrogation training includes guidance about every aspect of interrogator behavior—when to stand, when to sit, when to move physically closer to the suspect, key phrases to use, and counterarguments to employ during the interrogation process. Both the physical setting and the nature of the interaction communicate to the suspect that he has lost control of the situation.

Social Isolation. Whenever possible, suspects are interrogated alone. Social isolation deprives the suspect of emotional support and prevents others from offering information that might contradict the account of the crime being presented by the police. The presence of a friend or ally in the room would bolster the suspect's resistance to persuasion and create challenges to the interrogator's version of events. This social isolation creates anxiety and a desire to escape from the interrogation room. Together, the combination of control and social isolation enables interrogators to shape the suspect's perception of his situation. Alone in a carefully constructed, emotionally distressing environment, it is impossible for a suspect to evaluate the interrogator's claims about the crime, the evidence, and the likely consequences of admission or denial. Control and social isolation lay the foundation for the tactics that follow.

Certainty of Guilt. When a suspect is confronted with an accusation that he committed the crime, an innocent suspect (and many guilty suspects) will respond with denials. Interrogators are trained to challenge, discount, and dismiss all such denials. The message from interrogators is: "denial is futile; you did it, we know you did it, and the evidence against you is indisputable." Perhaps the most potent tool that interrogators have at their disposal is the ability to lie about the existence of incriminating evidence. Police are permitted to confront suspects with seemingly objective and irrefutable evidence of guilt—whether or not such evidence actually exists (*Collins v. Brierly* 1974; *People v. McRae* 1978). Police may pretend to have eyewitnesses, fingerprints, DNA, or surveillance videos that clearly establish guilt. Police may even falsely inform the suspect that he failed a polygraph test. After describing the "evidence" against the suspect, interrogators typically argue that any reasonable judge or jury would easily find him guilty. These repeated accusations and claims of incriminating evidence sometimes convince even innocent suspects that their best option is to falsely admit guilt.

As part of the certainty of guilt tactic, interrogators will often tell the suspect that the purpose of questioning is not to discuss *whether* he 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 focus from denial to explanation.

Minimization of Culpability. The use of minimizing scenarios works in tandem with the certainty of guilt tactic. After convincing the suspect that he will very

likely be convicted by a judge or jury, interrogators try to clear the path for an admission by offering exculpatory, face-saving justifications for the crime. For example, it might be suggested to a murder suspect that the victim was killed by accident or in self-defense. It might be suggested to a robbery suspect that money was stolen to pay for an important family expense. It might be suggested to a rape suspect that sex was consensual and initiated by the victim. The crime might be portrayed as a "mistake" or "accident" or a "spur of the moment" misjudgment. Minimizing scenarios help to entice an admission by shifting blame to others, by redefining the circumstances that caused the act, or by redefining the act itself. Such scenarios legally, morally, or psychologically justify the crime, and are designed to make a suspect more comfortable with admitting guilt.

Although police are prohibited from making explicit promises of leniency, the use of minimizing scenarios clearly implies that anyone making a judgment about the suspect's behavior (e.g., a prosecutor, a judge, or a jury) will be likely to treat the suspect with leniency. Minimization of culpability is intended to suggest that the offense may not be that serious and that the suspect will be charged with a less serious crime and receive a shorter prison sentence (or no prison sentence at all) if he just accepts responsibility, admits guilt, and expresses remorse (Leo et al. 2009). Conversely, continued denial will result in a more serious charge and a longer prison sentence.

Reshaping the Suspect's Decision Calculus

It is extremely difficult for jurors who have not experienced a high-pressure interrogation to understand how an innocent person might decide to confess to a crime he did not commit. Such a decision seemingly runs counter to both logic and self-interest. Indeed, only about 7 % of jurors believe that, if interrogated by the police, they might falsely confess to a crime they did not commit (Costanzo et al. 2010). This finding is consistent with a large body of social–psychological research indicating that people underestimate the extent to which their own behavior might be shaped by strong situational pressures (Zimbardo 2007). What jurors and judges sometimes fail to recognize is that the psychological assault of interrogation is intended to reshape the suspect's understanding of his decision options and the consequences of making an admission.

Unless the suspect is mentally impaired, there is no immediately obvious, easy-to-understand explanation for a false confession. Because this is so, part of the difficult job of the expert witness is to explain why an innocent suspect might falsely confess. As the naïve decision model depicted in Fig. 1 illustrates, the key factor that should figure into a suspect's decision-making process is the strongly negative potential consequences of confessing: a "guilty" verdict and possibly a prison sentence. As Fig. 1 also shows, confessing appears to offer no clear benefit that might counterbalance the huge apparent costs of confessing. From the commonsense perspective of most jurors evaluating the veracity of a confession, this

Fig. 1 Naive Model of Suspect Decision-Making

large cost/little benefit calculation should prevent all but the most psychologically vulnerable suspects from making false admissions.

From the vantage point of outsiders looking into the interrogation room, it is extremely difficult to imagine how a suspect could reasonably weigh the alternatives and come to the conclusion that offering a false confession is in his or her own self-interest. Without a fuller appreciation of the dynamics of the interrogation process, jurors are likely to reach the conclusion that a confession is dispositive of guilt. This fuller appreciation is sometimes provided by expert testimony.

So, why do not the heavy costs associated with a confession deter all innocent suspects from making false confessions? A partial answer is that, as explained later, some suspects suffer from individual vulnerabilities (e.g., youth, mental illness, mental disability) that render them especially susceptible to the pressures of interrogation. However, because the majority of false confessions have come from mentally normal adults, the fuller answer is that the process of interrogation distorts and reshapes the decision-making process of the suspect. A suspect may enter the interrogation room confident that he will never make an admission. However, the interrogation process is designed to restructure the subjective cost/benefit analysis of the suspect. A successful interrogator will convince the suspect that making an admission is the best available means of improving his otherwise bleak prospects (Leo et al. 2007). As illustrated in Fig. 2, the decision-making process induced by interrogation removes the option of denying involvement in a crime. Through use of the certainty of guilt tactic and false-evidence ploys, interrogators shift the burden of proof—the suspect is now presumed guilty.

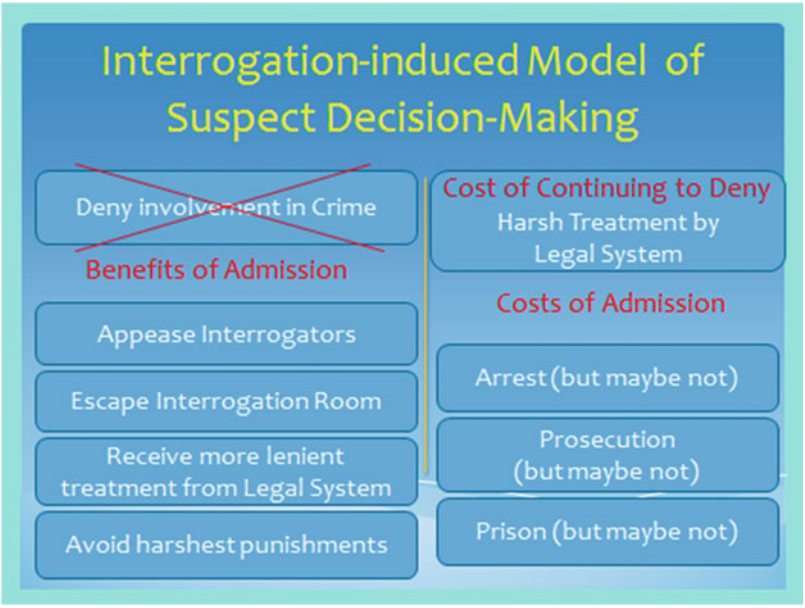

Fig. 2 Interrogation-Induced Model of Suspect Decision-Making

Over the course of the interrogation, the suspect is likely to become convinced not only that continued denials are ineffective, but that continuing to deny guilt will lead to especially harsh treatment by the legal system. For example, police may tell the suspect that if he fails to make an admission, he will be seen by prosecutors, judges, and jurors as a cold, remorseless criminal. He will be easily convicted and receive the maximum prison sentence. In contrast, police may suggest than an admission will enable them to "help" the suspect by softening how he is treated by the legal system. A person who "stands up" and "takes responsibility" for his "mistake" will be viewed more positively. In effect, the interrogation process introduces a clear benefit of admission—more lenient treatment from the legal system. It also introduces a clear cost of continuing denial—harsher treatment from the legal system.

Decision researchers often advise that the quality of decisions can be improved by full consideration of costs and benefits, generating multiple alternatives, reducing stress, taking time to consider alternatives, and consulting with others (Heath and Heath 2013). Interrogations are intentionally designed to violate such guidelines. The process of interrogation creates stress, restricts alternatives, imposes time pressure, and removes the ability to gather information or to receive emotional support from others. First, the stressful, emotionally taxing process of interrogation weakens the suspect's capacity to think clearly and dispassionately evaluate alternatives (Costanzo 2004; Davis and Leo 2012b; Davis and O'Donohue 2004). Second, alone in the interrogation room with one or more interrogators, the suspect

is intentionally deprived of the social support that might strengthen resistance to persuasion. Third, access to information is restricted. In the interrogation room, it is impossible for suspects to evaluate claims about the evidence against them or claims about how their case will be viewed by others in the legal system.

As shown in Fig. 2, the interrogation process forces suspects to reassess the perceived costs and benefits of confessing. Because the option of continuing to deny involvement in the crime seems to make things worse, the alternative of admitting involvement (for a justifiable reason) may appear to be the only viable course of action. Indeed, an admission is presented as the only option that will improve the suspect's treatment in the legal system. Because the decision to confess is made under conditions of extreme stress, relief from the stress of interrogation also comes to be perceived as a clear benefit of confessing. An innocent suspect worn down by the pressures of interrogation may come to see an admission as the only means of appeasing a relentless interrogator and escaping the interrogation room. Indeed, in several proven false confession cases, people have made false admissions to end the interrogation and escape from the interrogation room, believing that they would be able to straighten out everything later (Innocence Project 2014).

Most theories of decision-making assume that people rationally weigh the value of alternatives and the subjective probabilities of various consequences that may arise from choosing each alternative. However, the process of interrogation is designed to prevent rational, dispassionate decision-making. Through tactics that rely on deception, manipulation, and coercion, the interrogator gradually restructures the psychological context within which the suspect makes the decision about whether or not to confess. Suspects become more focused on the immediate goal of ending the interrogation and become less able to appreciate the long-term consequences of making an admission (Guyll et al. 2013). The short-term benefit of ending the interrogation may become so salient that it outweighs the potential long-term costs of a confession.

Vulnerability and Resistance to False Confession

There appear to be five basic pathways to false confession. Suspects confess to crimes (truly or falsely) to (1) escape the intolerable pressures of interrogation, (2) because they become convinced that confessing will lead to better outcomes than denial, and/or (3) because they are unwilling or unable to resist compliance with the demands of the interrogator. In addition, some suspects may confess, in part, (4) because they are less resistant to prosecution and incarceration, (as is more likely for some minorities, for example, see Villalobos and Davis, in press). More rarely, innocent suspects may be (5) led to doubt their own memories and become convinced they actually did commit the crime (Gudjonsson et al. 2014; Kassin 2007). Resistance to these pathways to confession is rooted in the personal resources available to the suspect in the interrogation room, accessible relevant knowledge, self-efficacy, and the valuation of potential outcomes.

Personal Resources Available to the Suspect. Intact self-regulatory resources are crucial for resistance to the first, second, and fifth pathways to confession (Davis and Leo 2012a, b; Follette et al. 2007). This includes self-regulation of emotion, cognition, and behavior. For example, to resist the first pathway to confession (to escape an intolerable situation), the suspect must be able to regulate emotions to minimize the experience of distress and must override impulses to satisfy the immediate need for escape in favor of actions promoting long-term good. To avoid making an incriminating statement, it is crucial for the suspect to control his attention and focus on the long-term goal of avoiding incarceration.

To resist becoming convinced that confession will optimize the suspect's long-term outcomes, self-regulation is again crucial. The suspect must regulate emotions to facilitate cognition. The suspect must focus attention to maintain the priority of his own goals and avoid diversion to those of the interrogator, and focus on relevant versus irrelevant information. He must be able to access relevant information from memory, to analyze incoming information, to critically evaluate options, to plan responses to the interrogator, and to choose among alternative responses to interrogator demands. Likewise, proper cognitive functioning and access to memory are crucial if the suspect is to resist the fifth pathway of becoming convinced that he did commit the crime. Even if the suspect does possess adequate relevant knowledge and cognitive skills, inadequate self-regulatory resources can interfere with use of these skills during interrogation (see Davis and Leo 2012a, b for reviews).

Accessible Relevant Knowledge. To arrive at a truly knowing, intelligent, and optimal decision as to how to react to interrogator demands, the suspect must possess a wealth of relevant knowledge and be able to critically and accurately analyze the claims of the interrogator. Ideally, this knowledge should include the right to refuse interrogation. But, if the suspect is interrogated, he will be confronted with a host of misinformation from the interrogator. This misinformation will range from falsification of the interrogator's attitudes toward the suspect and goals for the interrogation, to specific claims about the nature and import of evidence against the suspect, likely reactions of others to denial versus confession, and legal outcomes including prosecution, sentencing, and incarceration (see Davis 2010 for a review of deceptions inherent to interrogation). To successfully resist such deceptions, the suspect would benefit from an accurate understanding of the law and workings of the legal system (and thus his real options and likely outcomes of confession vs. denial), the realities of various forms of evidence (e.g., What is DNA? From where can it be collected and for how long after it is deposited? What can degrade it?), and the resources available outside the interrogation to dispute the allegations. The suspect would also need the intellect to use such information effectively.

Self-efficacy. The process of interrogation is intended to shift the suspect's attention from establishing innocence to minimizing the consequences of guilt. Interrogators work to undermine the suspect's sense of self-efficacy and to replace it with a sense of hopelessness. A number of suspect vulnerabilities can magnify the sense of hopelessness.

Some vulnerabilities derive from lack of personal and financial resources. Those without supportive social and professional networks and without personal financial resources, for example, may feel they have less ability to mount legal and public resistance to the allegations. Others may lack self-efficacy due to the evidentiary situation they face, or believe they face. Interrogators attempt to characterize the inculpatory information against the suspect as definitive. For some suspects, this is particularly effective because of the very nature of the accusations. Davis and Leo (in press), for example, have suggested that claims such as "he-said, she-said" sexual abuse/assault allegations undermine suspects' self-efficacy for defense in that no evidence exists or is likely to be found to prove the allegations false. Thus, the suspect may be at increased risk of false confession simply because of the nature of the allegations against him. This is particularly true when stereotypes associate his social group (such as stepfather) with the crime in question (such as child sexual abuse).

As suggested by the preceding example, self-efficacy may generally be undermined by "stereotype threat" (discussed more fully below). To the extent the suspect's social category is stereotypically associated with the crime in question or with crime generally, the suspect may feel reduced self-efficacy in convincing others of his innocence (see Davis and Leo in press). Finally, some suspects may suffer reduced self-efficacy, in that they more generally lack confidence in their ability to communicate and/or to persuade or influence others (or at least to persuade those in authority).

Valuation of Expected Outcomes. Suspects are affected by the expected outcomes of confession versus denial and by the valuation of those outcomes (e.g., Kassin et al. 2010). These outcomes can be immediate (such as distress, displeasing the interrogator and others), or long-term (such as prosecution, imprisonment, and impact on reputation, family, or career). Vulnerability to confession increases when a suspect expects relatively more negative outcomes of denial and/or relatively fewer negative outcomes of confession. Interrogation tactics, of course, are designed to facilitate these perceptions. However, some suspects are more vulnerable to one or both. For example, among the negative outcomes of denial are the potential displeasure, anger, violence, or other retaliation of the interrogator. Suspects may be unwilling to resist the interrogator when such outcomes are viewed as more likely (as may be true for minority suspects or others who fear or distrust law enforcement), or when negative outcomes are viewed as more aversive (such as low IQ suspects, those who are particularly conflict avoidant or high in social desirability concerns) (Villalobos and Davis, in press). Such suspects can be more vulnerable through the third pathway of compliance.

A second class of influences on the valuation of outcomes involves suspects' consideration of how a confession will affect others. The suspect may value protection of another person over his own outcomes (which is among the prominent causes of false confessions, e.g., Sigurdsson and Gudjonsson 1996), or may fear reprisals if he fails to confess to protect the responsible party (e.g., in gang-related crimes).

A third class of influences involves "temporal discounting," the suspects' valuation of immediate versus long-term outcomes. Young suspects, for example, tend to be particularly sensitive to immediate concerns such as getting out of the interrogation, and tend to give less consideration or weight to long-term consequences such as prosecution and imprisonment (see discussion of juvenile suspects below).

Finally, some suspects may be more vulnerable to confession because they evaluate potential prosecution/incarceration less negatively. Villalobos and Davis (in press) review evidence, for example, that some minority suspects are more likely to expect to be incarcerated at some point in their lives, and are less averse to that possibility (even viewing it as inevitable or a mark of status). Populations such as the homeless or others without social or professional ties to lose if incarcerated may likewise value incarceration less negatively.

Interrogation-Related Regulatory Decline

The process of interrogation is designed to impair the self-regulatory abilities of the suspect. Over the course of the interrogation, the suspect is likely to experience a decline in the ability to control his attention, thoughts, emotions, and behavior. In addition, the circumstances surrounding the interrogation (e.g., emotional distress, sleep deprivation) may weaken the suspect's motivation to maintain his innocence.

Self-regulation. Davis and Leo (2012a, b) coined the term "interrogation-related regulatory decline" (IRRD) to refer to failures of self-regulation linked to circumstances surrounding interrogation (such as crime related distress, drug use, etc.), or to the interrogation itself (such as length, aversive tactics, fatigue, etc.). Such failures of self-regulation affect emotion regulation, impulse control, suggestibility, and confession. Many such factors are subtle, may not be recognized by investigators and attorneys, and may not be presented to juries at trial (Davis and Leo 2012c).

The suspect may enter the interrogation already severely compromised by crime related stressors (such as death of friends or family), sleep deprivation, intoxication, or illness—all of which are likely to impair self-regulation. Distress, fatigue and glucose depletion increase as the interrogation progresses and are exacerbated when suspects are not provided with refreshments (or are too stressed to ask for or consume them). Distress is further promoted by accusations, isolation from social support, the uncomfortable interrogation environment, aversive tactics, length of the interrogation, and unavailability of nicotine, drugs, refreshments or toilet breaks. Moreover, subjective reactions of the suspect can also be stressful: such as the experience of stereotype threat, feelings of hopelessness and fear, or the emotions provoked by threats to one's social identity and self-esteem.

While substantial evidence links self-regulation failure with lowered resistance to influence, experimental research has just begun to link failures of self-regulation

to behavior during interrogation. Several studies have shown, for example, that comprehension of and memory for *Miranda* warnings is affected by stress. Two studies showed that the mere fact of an accusation of wrongdoing can impair memory for and understanding of the warnings. Rogers et al. (2011) found that participants who were accused of a theft felt more stress and demonstrated worse recall and comprehension of *Miranda* warnings than participants not accused of the theft. Similarly, Scherr and Madon (2012) found that participants accused of cheating (vs. those not accused) felt greater stress and scored significantly lower on measures of *Miranda* comprehension. Scherr and Madon (2013) asked students accused of misconduct to waive their rights to have a student advocate present when questioned. Students were less likely to comprehend the waiver when accused of a more serious offense, presumably as the result of increased stress associated with more serious accusations. Finally, Davis et al. (2013) showed that Black and Hispanic (but not White) participants remembered *Miranda* warnings more poorly when anticipating an interview about criminal behaviors than about thrill-seeking activities. This impairment was presumably the result of increased cognitive load and demands for emotion regulation imposed by the "stereotype threat" experienced by minority participants (for whom stereotypes associate their race with crime).

Not surprisingly, the experience of interrogation has been shown to increase physiological stress responses. Among guilty versus innocent students accused of cheating, the innocent at first experienced less physiological distress (likely due to initial confidence in refuting the accusation). But the innocent who continued to resist confession ultimately exhibited greater sympathetic nervous system activation. This finding suggests that innocents who resist confession will experience greater depletion of resources over time and suffer increasing vulnerability to confession as a result (Guyll et al. 2013).

Other studies have linked stressors to interrogation reasoning and behavior. Davis et al. (2013), for example, tested the notion that regulatory decline will affect the ability to remember and use exonerating information to evaluate one's options in interrogation. Participants were asked to read a murder case in which an innocent man was accused of murder. The strength of evidence varied. After completing a relatively easy or difficult (depleting) working memory task, participants read a transcript of the innocent suspect's interrogation, and answered questions regarding the strength of evidence against the suspect, and rated the likelihood he would be convicted despite his innocence and other case-related perceptions. Those who completed the difficult task rated the strength of evidence and the likelihood of conviction more similarly across high and low evidence strength conditions than those who completed the easier task, and recalled less of the potentially exonerating evidence. The authors interpreted the results as indicating participants were less able to access memory for the evidence against the suspect when self-regulatory resources were more depleted.

Madon and colleagues (2013) provided evidence that stress, manipulated via the length of an interview, increased the tendency to make more admissions and to focus on immediate consequences (relief from the pressures of interrogation) rather

than long-term consequences (punishment). Using similar methodology, Scherr et al. (2014) showed that when cognitive and regulatory resources are naturally weaker, participants made more admissions to minor criminal/unethical behaviors. That is, larks (those who arise and go to sleep early) made more admissions when interviewed at "off peak" times (7:30 p.m.); whereas owls (those who arise and go to sleep late) make more admissions when interviewed early in the day (7:30 a.m.). Moreover, both larks and owls made more admissions later in the interview, as their fatigue increased.

The foregoing studies clearly demonstrate that stress and self-regulation affect thinking and behavior in interrogation-related contexts. However, much remains to be done to illustrate the many sources of self-regulation failure inherent to interrogation, as well as the many specific effects on impulse control, reasoning, and decision-making in interrogation (see Davis and Leo 2012b).

Acute Motivational States. Distress can also promote confession through its effects on the valuation of outcomes of confession versus denial. That is, when sufficiently distressed, one may cease to care. False confessor Jerry Lee Hobbs, for example, described his reactions as follows: "I found my daughter…She didn't even have eyes in her head. I was already broken. They didn't have to break me" (Martin 2011). Like many whose children or loved ones are killed, suspects may feel so profoundly distressed or depressed that they do not care, in the moment, what happens to them. They may also feel guilt for the death, even though they were not criminally liable. For example, a child may die through an injury the parent failed to prevent, but did not cause. That parent may feel some sort of punishment (e.g., incarceration) is deserved.

Though some of the factors affecting acute vulnerability to influence may be noted in police reports, many will likely not (see Davis and Leo 2012c). Thus, it is important for attorneys and interrogation experts to investigate the condition of the suspect going into interrogation and to identify the emotional state of the suspect at the time of the interrogation. Sleep deprivation, fatigue, lack of nourishment, intoxication, drug withdrawal, and other factors could impair self-regulation or affect acute motivations.

Suspect Vulnerabilities that Increase the Risk of False Confession

There are several types of suspect vulnerabilities that raise the risk of false confessions. Risk factors internal to the suspect include youth, mental impairment, and a variety of temporary impairments (e.g., sleep deprivation, being under the influence of drugs or alcohol, grief, hunger) that render suspects more vulnerable to police persuasion and manipulation. We devote special coverage to a type of suspect vulnerability that has received little attention in the research literature—being a foreigner or being a member of a minority group.

Juveniles. There is a consensus among scientists, mental health practitioners, and legal scholars that juveniles are especially vulnerable to interrogation tactics (Kassin et al. 2010), and laboratory research shows that adolescents are more likely to falsely confess to wrongdoing than are adults (Redlich and Goodman 2003). In a large sample of proven false confessions, 32 % of confessors were under the age of 18, and 63 % were under the age of 25 (Drizin and Leo 2004). Gross et al. (2005) evaluated more than three hundred cases of exonerated individuals during a 15-year period in the U.S., and reported a greater number of juvenile false confessors than adult false confessors. Surprisingly, 75 % of juvenile false confessors were 12–15 years old when they admitted to homicide or rape. Despite such findings, police are generally allowed to question juveniles in the same manner as adults (Feld 2013), and very little coverage of juvenile vulnerability is included in police interrogation training (Inbau et al. 2013). In fact, the coverage of potential risk for juvenile suspects in the Inbau et al. manual is surprisingly light (see p. 149), compared to the more in-depth coverage on developing special "themes" to exploit the weaknesses of young suspects (see p. 250).

Intuitively, jurors and judges may not have great difficulty making the link between interrogation pressures and the likelihood of false confession in the juvenile population (see Najdowski and Bottoms 2012). However, it may be difficult to keep the vulnerability of a juvenile in mind if the accused adolescent is a member of a gang, or has a history of getting in trouble with the law, or if the nature of the crime is especially heinous. It may therefore be important for an expert to highlight key cognitive and social mechanisms that cause adolescents to more easily succumb to interrogation pressures. Characteristics of juveniles such as deficient understanding of legal rights, immature impulse control, oversensitivity to rewards, the tendency to discount long-term consequences, and susceptibility to social influence may lead to false admissions, especially if manipulative or high-pressure tactics are used by interrogators. Juveniles are in fact subjected to high-pressure interrogation tactics. Malloy et al. (2014) reported that in an incarcerated population, the majority of the juveniles indicated experiencing high-pressure techniques during interrogations, including the use of force, explicit threats, and refusals to provide protection when requested (e.g., a lawyer, a parent, or a phone call). Content analyses of actual interrogations of juvenile suspects also reveal the use of many high-pressure tactics (e.g., repeated accusations, certainty of guilt, and holding suspects alone in interrogation rooms) (Cleary 2014; Feld 2013).

Research shows that compared to adults, juveniles misunderstand their *Miranda* rights (Frumkin et al. 2012), and may without comprehension, waive their rights and allow police to question them without advocates (see Rogers 2011). Indeed, research suggests that juveniles wave their rights at a higher rate than adults (Feld 2013; also see Cleary 2014). Rogers et al. (2014) studied a sample of legally involved juveniles (13–17-year-olds held at a probation facility) to determine the relationship between maturity and understanding of the *Miranda* warnings. Regardless of maturity level (low, medium, or high), results showed that about 80 % of the youth exhibited significant erroneous beliefs about their rights (e.g., family would have to pay for appointed counsel) that could compromise their

decisions about *Miranda* waivers. Moreover, on immediate recall tests, all maturity groups failed to recall about half of the warnings they were just given. It is possible that the vocabulary used (e.g., "appointed") and concepts included (e.g., "indigent") in some *Miranda* warnings exceed the comprehension ability of even the most mature adolescents (Rogers et al. 2014).

Adolescence is a time when impulsive behavior and discounting of risks in decision-making is more pronounced. Shulman and Cauffman (2013) contend that compared to adults, adolescent decision-making is biased toward immediate gratification due to an oversensitivity to rewards and insensitivity to costs. They argue that this might be mediated by asynchronous brain development of the socioemotional and cognitive control systems. Specifically, they suggest that because adolescents' mental control system is immature, implicit uncontrollable emotional forces may influence adolescents more than adults in their initial risk appraisals. Shulman and Cauffman (2013) asked a large sample of adolescents and adults to imagine engaging in scenarios involving risky behaviors (e.g., getting into a car with an intoxicated driver, having unprotected sex, and law-breaking activities) and to decide on the costs and benefits of such behaviors, the seriousness of the consequences, and likelihood of negative outcomes. The relevant results showed that adolescents were more reward biased in their risk appraisal (focusing more on benefits) than were adults. In an interrogation context, an adolescent, more easily than an adult, may focus on the immediate benefits of ending the interrogation by confessing (even if falsely), rather than focusing on the long-term negative consequences of such actions.

Juveniles tend to be more suggestible than adults, often showing a pattern of "going along" when faced with outside pressures, including adults in positions of authority. They also tend to be highly susceptible to peer influence and place a great deal of importance on peer relationships (e.g., Gardner and Steinberg 2005). According to Malloy et al. (2014), the desire to protect a peer may be especially powerful for adolescents and can lead to false confessions. To explore this hypothesis, Malloy et al. (2014) surveyed a large group of 14–17-year-old incarcerated males about their experiences during interrogations. Falsely confessing to serious crimes was reported to be common (one-third of the sample), and half of the false confessions were made to protect someone, either a friend or a family member. The authors suggested that juveniles falsely confessed primarily to protect an in-group member from getting a harsh sentence. Additionally, it was common for false confessions made under duress to be retracted but not for false confessions made to protect someone.

Mental Impairment. A leading expert on police interrogations and confessions, Gisli Gudjonsson (2003, 2014), refers to characteristics such as developmental disorders, mental disorders, personality disorders, low or borderline intelligence, and abnormally high suggestibility and compliance as "mental vulnerabilities." These vulnerabilities can render suspects susceptible to providing false information. Even though mental vulnerabilities are commonly recognized as risk factors, confessions from mentally impaired suspects may be easily believed by the trier of

fact, regardless of veracity (see Henkel 2008). This underscores the importance of engaging a forensically trained clinician to evaluate the individual suspected of having a mental vulnerability. Police routinely interrogate persons with mental vulnerabilities (see Gudjonsson 1993), and this group is believed to be overrepresented in detained populations (Redlich 2004; Rogers 2011). Furthermore, in a recent study of inmates in a U.S. county jail with serious mental illness, Redlich et al. (2011) found that false confessors reported experiencing more frequent and longer questioning, and more pressure during interrogations than true confessors. This suggests that mentally ill suspects are likely to be subjected to the same intense interrogation tactics as nonvulnerable populations.

Analyses of wrongful convictions show that a significant percentage of false confessors can be classified as "mentally retarded" or "mentally ill" (Gross et al. 2005). Furthermore, there appears to be inadequate police training and sensitivity in recognizing mental impairment during interrogations (Redlich 2004). An analysis of 75 cases involving false confessors revealed that low or borderline intelligence of suspects was undetected or ignored by police interrogators (Perske 2011). In the Drizin and Leo sample, 22 % of the false confessors could be classified as mentally retarded, and 10 % had been diagnosed with some form of mental illness. We do not know how many false confessors have some form of undiagnosed impairment.

Mentally retarded individuals tend to be more acquiescent, more likely to agree with illogical questions, and are more easily confused by the tactics of interrogators (Finlay and Lyons 2002). Like mentally retarded suspects, mentally ill suspects may suffer from an inability to foresee the long-term consequences of admissions made in the interrogation room. Suspects suffering from a psychotic disorder may experience delusions that prevent them from assessing the reality of their own memories. This impairment makes mentally ill suspects more susceptible to internalizing and believing information presented in the form of minimizing scenarios or false-evidence ploys.

Moreover, individuals with mental impairment sometimes show difficulty communicating, understanding the language in the questions asked, and remembering critical information that may exonerate them or be helpful in making decisions (Gudjonsson 2003). Deficits in communication, memory, and understanding can lead suspects to be less assertive and more reliant on the interrogator for information. Many mental health or personality disorders are also associated with deficits in emotion regulation and impulse control (e.g., Follette et al. 2007). Consequently, people suffering from these disorders are more susceptible to distress during interrogation, and to confessing simply to escape the distressing situation.

Foreigners and Minorities. Suspects who are not members of the dominant culture may be at a distinct disadvantage in the interrogation room. Culture, as concisely defined by Hofstede (2011), is "the collective programing of the mind that distinguishes the members of one group or category of people from others" (p. 3). Matsumoto (1997) specifies that culture is "the set of attitudes, values, beliefs, and behaviors shared by a group of people…" (p. 4). As these definitions suggest, members of a specific culture share core beliefs, values and attitudes that influence

their behaviors in ways that distinguish them from other cultures. The influence of cultural differences on interrogations and confessions has received relatively little attention from researchers (see Leo et al. 2009; Villalobos and Davis, in press, for exceptions).

There are at least four reasons why cultural differences might render foreigners or minorities more vulnerable in the interrogation room. First, foreigners are likely to lack knowledge about the dominant culture's legal system, interrogation rights, immigration laws pertinent to suspects, and the process of police investigation. Second, limited understanding of the English language may increase the risk of offering false incriminating information. Third, suspects from a collectivistic or high power distance culture may be less likely to challenge authority, more fearful of police, more open to accepting police officers' scenario of events, and more likely to comply with the suggestion to confess. Fourth, the phenomenon of stereotype threat may be at play when the suspect belongs to a marginalized group. Each of these four factors—or a combination of them—may explain the enhanced risk of false confession for foreign-born or minority suspects. Because of these influences, the stress of being interrogated may be exacerbated when the suspect is not part of the mainstream culture or is a member of a marginalized group. Nevertheless, in their important manual on criminal interrogations and in their popular 4-day training, Reid and Associates hardly mention cultural influences on suspects' behavior (Inbau et al. 2013). This is surprising given that police routinely interrogate suspects who have emigrated from other countries.

Limited knowledge of the legal system and language barriers. Suspects from foreign countries may be more vulnerable to interrogation tactics for a variety of reasons, including concerns about their immigrant status (if they are not naturalized citizens), confusion about the intricacies of the criminal justice system, and language barriers that may prevent them from fully participating in their own advocacy. For example, the complex legal language in the *Miranda* warnings can leave foreign-born suspects, especially those with limited English language skills, in jeopardy of no legal protection. Even translated *Miranda* warnings (e.g., to Spanish) can be flawed or incomplete and may lead to serious impairments in understanding (Rogers et al. 2009). Convoluted presentation of the interrogation rights and complex language can be mentally challenging, which can not only increase cognitive load and acute stress but also induce meaningless waivers.

Language barriers also create vulnerability in the interrogation room, a place where the questioner is trained to fully control the conversation. In the U.S., during important legal proceedings (e.g., in court) suspects with Limited English Proficiency (LEP) are often provided with certified interpreters. These interpreters are accredited to assist with communication needs and have a duty to translate everything as accurately as possible. However, in the interrogation room, translation is usually in the form of a bilingual detective acting as an interpreter between the suspect and a monolingual interrogating detective. Alternatively, if the interrogating detective is bilingual, he or she will conduct the questioning in the suspects' native language.

Berk-Seligson (2009) analyzed cases involving interrogations of LEP Latino suspects by detectives with limited language skills. She concluded that the lack of quality communicative assistance during interrogations constitutes "language coercion" and renders many confessions unreliable. During interrogations, detectives manipulate the communication exchange by talking most of the time and by asking leading and confirmation-seeking questions with the goal of eliciting incriminating statements. In this high-pressure environment, suspects with LEP experience then added anxieties of being unable to fully understand or clearly communicate with the interrogator. Berk-Seligson (2009) found that interrogators acting in the role of interpreters often misrepresent themselves as impartial helpers. Indeed, outside the interrogation room, interpreters do act as impartial translators or even as allies to people accused of crimes. However, in the interrogation room, this is not the case. Officers acting as interpreters may not accurately translate the exchanges during interrogations and, at times, misconstrue what the suspect is saying (Berk-Seligson 2009). In many cases, interpreters act primarily as interrogators.

One of the authors of this chapter (IBG) has analyzed half a dozen interrogations of LEP suspects and noted the same translation issues raised by Berk-Seligson (2009). In only one case did the officer acting as an interpreter accurately translate both questions and answers. Even in this case, the interpreter added her own questions or interpretations. However, in the other interrogations, one or more concerns surfaced because of misinterpretation or misrepresentation by the person doing the questioning.

In one case, the immigrant LEP suspect had been in the U.S. for 2 years, and even though his native language skills suggested higher education, he did not understand the *Miranda* warnings when read in Spanish. The bilingual interrogator who was somewhat proficient in Spanish began by telling the LEP suspect that he had rights in this country and that once he agreed to them they could talk. This positive spin on the rights suggested that the suspect needed *to agree*, not to exercise his rights. Once the detective began to read the *Miranda* warning, the suspect appeared confused, and he told her that he did not understand clearly. She attempted to explain, but confused the suspect even more by, for example, suggesting that he had a right to an attorney in court (but not in the interrogation). It is likely that the detective was influenced by the competing needs to apply tactics that would elicit consent, but to also explain in a way that made the consent seem voluntary and therefore legal. In the process, she miscommunicated what it means to have those rights. These lost-in-translation aspects of the interrogation, which are important for even native speakers of English, are even more critical for LEP suspects.

Like other vulnerable groups (e.g., juveniles and suspects with mental impairments), LEP suspects may be more likely to be compliant or suggestible when they do not understand the language being used in legal contexts (see Lamb et al. 2011; Ridley et al. 2013). LEP suspects might also experience increased cognitive load and stress from trying to communicate in broken English or from struggling to understand communication in a context that is not only designed to manipulate, but has the added component of being unclear.

Cultural Heritage and Experience. Most cultures include a strong pressure to comply with authority, especially the police, but the degree to which this is programed into people's consciousness and obeyed by individuals varies as a function of two important cultural dimensions. According to Hofstede's (1980) model of cultures, societies can be conceptually located along interrelated dimensions, which are relevant to interrogation contexts: *individualistic* to *collectivistic* and *low* to *high power distance* dimensions. The individualistic–collectivistic dimension refers to the degree to which members of a particular society are integrated into groups. In collectivistic cultures (e.g., parts of Asia and Latin America), people have a strong sense of belonging and loyalty to cohesive in-groups beyond the immediate family. In contrast, in individualistic societies (e.g., United States), people have a sense of independence, individual identity, and individual freedom (Hofstede 2010). Among the differentiating characteristics between collectivistic versus individualistic societies are: a "we" versus "I" consciousness, a desire to maintain harmony versus speaking ones' mind, and feelings of shame versus guilt in response to norm transgression (Hofstede 2011).

The dimension of power distance refers to the extent to which people in a culture accept and expect that power is distributed unequally. Power distance is measured from the subordinate's point of view. In high power distance cultures (e.g., many countries in Asia, Latin America, Africa, and Eastern Europe), people accept unequal distribution of power and strict hierarchies within institutions as a part of life, perceiving their legitimacy as irrelevant (Hofstede 2011). In contrast, people in low power distance societies (such as Germanic and English-speaking Western countries) believe that power should be legitimate and subject to criteria of merit. Research indicates that characteristics which differentiate high from low power distance societies include: teaching of obedience versus teaching equality, respect for and fear of elders versus neither respect for nor fear of elders, and an expectation of being told what to do versus an expectation of being consulted (Hofstede 2011).

In collectivistic cultures, communication is more indirect and contextual than in individualistic cultures (Adair and Brett 2004). This suggests that a suspect from a collectivistic culture may be more acutely aware of and sensitive to the implied messages of threat and leniency conveyed by an interrogator. Moreover, individuals from collectivistic cultures are likely to be concerned with status and power differences (Adair and Brett 2004). This may explain why studies on conformity show that collectivistic countries tend to have higher levels of conformity than individualistic countries (Bond and Smith 1996). During interrogations, Beune et al. (2010) found that "behavior that accuses, threatens, or warns of particular consequences for the suspect's family and/or friends" (p. 911) was more influential in eliciting information from suspects belonging to collectivist cultures than from suspects belonging to individualistic cultures.

Experience with and knowledge of the legal system in one's own country may also distort perceptions of the interrogation process in the United States (Leo et al. 2008). A foreign suspect may come from a culture where the legal protections for suspects are weak or where law enforcement is corrupt and physical abuse of suspects is common. For example, in Japan, the detention period before one is officially charged can be up to

3 weeks, during which time suspects may go through multiple interrogations by police and prosecutors (Oi 2013). In China, use of torture to extract confessions was not outlawed until 1996, and the practice has not been eradicated, probably because penalties for torture remain relatively minor (Watt 2014). According to a 2014 world report by Human Rights Watch, Mexico's war on drugs has led to the continued practice of police torture to force confessions from suspects even though constitutional law outlawed such practices in 1991. And, in parts of Africa and the Middle East, torture is still used (Human Rights Watch 2014). Fear of physical abuse at the hands of police can create great anxiety in suspects before the process of interrogation even begins and can lead some suspects to offer incriminating admissions to prevent the abuse they fear is imminent (Leo et al. 2008).

Stereotype Threat. Stereotype threat refers to "being at risk of confirming, as self-characteristic, a negative stereotype about one's group" (Steele and Aronson 1995, p. 797). Stereotype threat applies to any stigmatized group in a context in which the stereotype about their group is salient. In threatening contexts, the individual can experience apprehension or excessive anxiety that negatively affects their behavior and ironically may confirm the stereotype. In a review of the stereotype threat literature, Najdowski (2011) argued that stereotype threat will be activated for suspects who are members of minority groups that are stigmatized with respect to criminality. For example, for a variety of reasons, including negative stereotypes that depict Blacks as criminals, African-Americans are more likely than Whites to be suspected of crimes by the police. If African-American suspects are aware of this stereotype, they will experience increased stress and mental load. Signs of stress and attempted control of behaviors may be wrongly perceived as signs of guilt.

Najdowski (2013) reported preliminary findings about the relationship between race/ethnicity and stereotype threat in the context of police investigations. One study examined whether African-Americans would experience stereotype threat in a realistic police encounter during a criminal investigation. The results showed that compared to Whites, African-Americans experienced stereotype threat and this led them to appear more nervous than Whites. In similar research, Najdowski et al. (2012) assessed whether Hispanics would experience stereotype threat when interacting with the police. In Study 1, Hispanics, more than Whites, reported that they would be concerned about being judged unfairly by police due to stereotypes. In Study 2, imagining a scenario of a police encounter led Hispanics more than Whites to agree that they would feel stereotype threat.

Once interrogated, populations subject to stereotype threat may experience additional sources of vulnerability to confession. Davis and Leo (2012a), for example, reviewed evidence that stereotype threat is depleting, causes self-regulation failure, and impairs impulse control, cognition, and decision-making. As noted earlier, Davis et al. (2013) provided an initial demonstration that stereotype threat impaired memory for Miranda warnings among African-Americans and Hispanics. The authors further suggested that stereotype threat can increase hopelessness in interrogation (due to suspects' perceptions that

they will be presumed guilty), and render the suspect more vulnerable to tactics designed to convince them that confessing is in their best interests. Moreover, suspects experiencing stereotype threat (and the threat posed by the belief they will be presumed guilty) may be more susceptible to minimizing scenarios offered by interrogators (see Davis and Leo, in press; Villalobos and Davis, in press). Such scenarios imply that the stigmatized social identity suggested by the stereotype (e.g., child-molesting stepfather) is not true, and offer a more positive identity (e.g., one who committed an accidental or one-time offense).

Juror Knowledge and the Impact of Expert Testimony

Judges' decisions about whether to allow expert testimony at trial are largely determined by assumptions about jurors. As discussed earlier, judges have substantial discretion in deciding whether an expert witness will be permitted to testify. The decision to exclude such testimony is often based on nothing more than judges' untested assumptions about what jurors believe and how jurors might be influenced by expert testimony. Similarly, defense attorneys usually argue for (and prosecutors usually argue against) allowing expert testimony without the benefit of data on what potential jurors are likely to know and believe. Fortunately, there is now a growing body of survey research that enables us to draw some conclusions about the "common knowledge" of jurors. Before reviewing this research, it is important to consider why jurors tend to believe confessions, particularly in the absence of expert testimony about interrogations and confessions.

Jurors' Bias to Believe Confessions

Jurors are predisposed believe confessions for several reasons. First, research on attributional biases suggests that observers are likely to infer that a confession is valid. There is a huge literature in social psychology demonstrating that when analyzing another person's behavior, we tend to underestimate the power of the situational forces acting on that person (Ross and Nisbett 1991; Zimbardo 2007). This tendency means that when evaluating a confession, we tend to discount the pressures of the interrogation process and attribute a suspect's confession to guilt. An interesting program of research by Daniel Lassiter and his colleagues has demonstrated that this tendency to discount situational pressures is exacerbated by the camera angle used to record most confessions. Specifically, a camera angle showing only the confessor leads viewers to rate the confession as more voluntary and less coerced than a camera angle showing both the interrogator and the suspect (Lassiter 2010).

Second, sometimes jurors do not see interrogations in their entirety. Instead, they may see a "recap"—a preselected portion of a video recording of a confession. In

some cases, interrogators do not even turn on the video camera until after the suspect has confessed and his confession is clear and well rehearsed. Consequently, what a jury might see is only the end product of a lengthy, intense process of interrogation. What is seen on video may have been orchestrated and rehearsed. The coercive process that culminated in a confession may not be visible to jurors.

Third, when trying to understand the behavior of others, observers generally assume that people are probably behaving in their own self-interest, in a way that leads to benefits and avoids costs. Similarly, jurors who are asked to evaluate the credibility of a confession may begin with an assumption of rationality. It is clearly contrary to self-interest to falsely confess to a crime. Unless a suspect is mentally impaired or has been subjected to physical abuse, observers who have never experienced a coercive interrogation are likely to have great difficulty understanding why anyone would offer a false confession. As discussed earlier, part of the job of an expert witness is to explain to jurors how the interrogation process is used to reconfigure the cost/benefit analysis of the suspect.

Fourth, the completeness and coherence of a confession is likely to make it convincing to the jurors who must evaluate its authenticity. Evidence in support of this proposition comes from research on the "story model" of juror decision-making (Pennington and Hastie 1993). According to this model, jurors evaluate whether a story is plausible on the basis of its coverage and coherence. 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), evidence is organized by jurors into story form. The critical elements of a coherent story are provided in many false confessions (Appleby et al. 2013; Garrett 2008). False confessions often contain vivid, specific details about a crime. These particulars may simply be a product of how the confessor imagines that the crime might have transpired, or the specifics may have been gleaned from news accounts of the crime, or the details may have been provided to the suspect by police—intentionally or unintentionally—over the course of the interrogation process. In addition, false confessors often provide a convincing motive for the crime. Motive is an essential element of a compelling story (McKee 1997), and a statement of a motive (e.g., greed, lust, jealousy, revenge) lends credibility to the confession. The believability of a confession is also enhanced by some expression of remorse. Remorse may be especially compelling because it may intuitively appear not to be a response to coercion, but a spontaneous emotional response. Like details and motive, an apology or expression of regret may lend a tone of realism and credibility to a false confession.

Finally, verbal acknowledgements by the suspect increase the credibility of a confession. In response to prompts from interrogators, suspects often acknowledge that they were treated in a fair, nonthreatening, nonhostile manner. Suspects may also verbally acknowledge that their statements were offered voluntarily. Interrogators are taught to elicit these acknowledgements because they increase the believability of the confession and undermine any arguments by the suspect that the confession was coerced. Such acknowledgements also help to ensure that the confession will be admissible at trial.

Jurors' Beliefs

Several survey studies have assessed juror knowledge and beliefs about interrogations and confessions. These studies have used diverse and realistic samples, including: surrogate jurors selected by a trial consulting firm to match the demographic characteristics of jury pools in seven U.S. cities (Costanzo et al. 2010), actual jurors waiting to be called in a California courthouse (Blandón-Gitlin et al. 2011), a jury-eligible community sample in Connecticut (Henkel et al. 2008), a diverse group of jury-eligible adults responding to an Internet questionnaire (Chojnacki et al. 2008), and jury-eligible college students (Leo and Liu 2009). Although the surveys differ in the specific questions asked of respondents, several findings are consistent across studies. These findings reveal significant juror misconceptions about interrogations and false confessions.

Although respondents were aware that false confessions can occur, they believed that false confessions are unlikely to be elicited from innocent suspects. Huge majorities of respondents (92 % in the Costanzo et al. sample and 87 % in the Henkel et al. sample) believed that they personally would never falsely confess to a crime if interrogated by police. Additionally, even though people correctly recognize coercion in some police tactics—such as threats and the use of deceptive tactics (e.g., false-evidence ploys)—they are unable to recognize as coercive other tactics known by scientists to increase the likelihood of false confessions (e.g., minimizing scenarios; Blandón-Gitlin et al. 2011; Leo and Liu 2009). Importantly, while people believe that true confessions from guilty suspects can be elicited using coercive tactics, they generally do not believe these same tactics elicit false confessions from innocent suspects (Blandón-Gitlin et al. 2011; Leo and Liu 2009). The surveys also reveal a lack of knowledge about the existence, frequency, and legality of many popular police tactics (e.g., cutting off denials, deception, downplaying the seriousness of the crime; Chojnacki et al. 2008; Henkel et al. 2008).

Potential jurors are also misinformed about the difficulty of detecting deception. A majority of participants believed that police interrogators are better than ordinary people at identifying lies (even from just body language; Chojnacki et al. 2008); and about 60 % believed that interrogators' ability to detect lies improves with experience (Costanzo et al. 2010). Interestingly, this belief is consistent with the beliefs of police officers (see earlier discussion). When potential jurors were asked about their own ability to detect false confessions, about 40 % believed that they would be able to discriminate a true from a false confession by viewing a video recording of the interrogation (Costanzo et al. 2010). As discussed earlier, these beliefs are at odds with empirical findings.

With respect to the usefulness of expert testimony, perhaps the most relevant finding is that an overwhelming majority of participants indicated receptiveness to expert testimony. Roughly three-quarters (74 % in the Costanzo et al. sample and 80 % in the Chojnacki et al. sample) indicated that it would be helpful for jurors to hear an expert witness testify about interrogation techniques and about why a defendant might falsely confess to a crime. Apparently, jurors believe that they

have something useful to learn from expert testimony, and that the information provided by an expert is not already part of their "common knowledge." If jurors believe a particular type of testimony would assist them with the difficult job of understanding the evidence and reaching a fair verdict, perhaps judges should be predisposed to allow such testimony. Of course, judges may believe they are better arbiters of what jurors would find helpful than are jurors themselves.

Jurors' Responses to Expert Testimony

Although there is little specific research on how jurors respond to expert testimony on confessions, research from other domains indicates that jurors use expert testimony responsibly and do not overweight that testimony when reaching a verdict decision (Cutler et al. 1990; Vidmar and Hans 2007). We would expect testimony on confessions to conform to this same general pattern.

Encouragingly, expert testimony on situational factors in interrogations does seem to have a small but educational influence on jurors. Woody and Forrest (2009) created a mock trial in which the crucial evidence against the defendant was a confession. The interrogation that produced the confession either included false-evidence ploys or did not, and expert testimony about confessions was either provided or was not. As in other studies, interrogations with false-evidence ploys were recognized as more coercive and deceptive than those without, but the manipulation did not lead to fewer guilty verdicts. However, the presence of expert testimony did have a small, marginal effect in reducing convictions. Similarly, Blandón-Gitlin et al. (2011, Study 2) showed a small but significant decrease in conviction rates after mock jurors were exposed to expert testimony in a disputed confession case. Participants who changed their verdicts from pre- to postpresentation of expert testimony rated tactics used in the interrogation as more coercive. They also rated all aspects of the expert testimony as more influential in their decision-making than participants who did not change their verdicts. As with other forms of expert testimony (e.g., eyewitness expert testimony; Cutler 2009), the goal is to sensitize jurors rather than to merely increase juror skepticism. That is, expert testimony should help jurors recognize the factors that increase the risk of a false confession. Conviction rates should be reduced only if significant risk factors are present in the case being decided.

There is very little empirical research on the influence of clinical expert testimony on jurors' evaluation and decision-making in confession cases. However, there is anecdotal evidence suggesting that informing juries and judges about the increased risk for false admissions among members of special populations who experience coercive interrogations can be effective (e.g., Gudjonsson 1999, 2014; Thomas 1995). There is also experimental research showing that mock jurors are significantly more likely to discount a coerced confession when they are informed that the suspect suffers from a medical condition (e.g., a serious heart problem) that influenced his confession (Henkel 2008). Jurors also seem willing to discount the

confession of a juvenile if they perceive the confession to be coerced (Najdowski and Bottoms 2012).

Effective expert testimony can help jurors contextualize a confession by linking it to the interrogation process and the individual vulnerabilities of the suspect. Further, testimony by an expert can focus the attention of jurors on factors that research indicates might increase the probability of a false confession. That focus should improve the quality of juror decisions.

Future Research

There is now a substantial body of research on interrogation tactics and the psychology of false confessions. This research has accumulated over several decades and has employed experimental, observational, archival, and survey research. Every year, new research is added to this solid scientific foundation. Despite this impressive research base, there are several areas that could benefit from further attention from researchers.

As noted above, one area in need of further exploration is the impact of expert testimony on jury decision-making. The few studies that have been published on this topic are encouraging in that expert testimony appears to have the desired educational effect on jurors. However, we know little about which aspects of expert testimony jurors find helpful and persuasive. We also do not know the extent to which expert testimony on confessions informs jurors. Can they distinguish between reliable and unreliable confessions as a function of expert testimony? In other words, can they transfer the knowledge gained from expert testimony into a careful evaluation of the confession evidence? It would be useful to know which components of an expert's testimony jurors find confusing or clarifying. Experimental research could be used to determine which types of information help jurors understand confessions, but it would also be illuminating to conduct postverdict interviews with actual jurors who have heard expert testimony about interrogations and confessions. Such interviews would provide valuable information about how jurors use expert testimony in the context of a complex, lengthy trial. We also know little about how the process of jury deliberation affects the interpretation of expert testimony. It is unclear whether the process of jury deliberation shifts the beliefs of individual jurors. Uncertain jurors might be especially persuadable during the process of group deliberation. Postverdict interviews with actual jurors would enable us to gain a better understanding of whether the information communicated via expert testimony played a prominent role during deliberation.

Judges are an understudied group. Probably because of the difficulty of gaining access to large groups of judges, we do not have a thorough understanding of how judges decide to admit or exclude expert testimony on confessions. We also know little about why judges sometimes limit such testimony—for example, allowing an expert only to offer general testimony without referring to the interrogation or confession in the case being tried. Both survey research (asking judges a series of questions about expert testimony) and experimental research (systematically

varying case scenarios presented to judges) could provide valuable insights into the decision-making process of judges.

As noted earlier, more research is needed on the role of culture in the interrogation room. Although the large body of research on cultural differences suggests that culture plays a significant role in the interrogation process, we need research directly exploring this link. That research must investigate not only the impact of cultural values and experiences on confessions, but also the impact of language barriers. Limited language proficiency and/or indirect communication can be sources of confusion and increase vulnerability to false confessions.

Other questions deserving of further investigation involve the influence of case type (e.g., murder vs. theft), defendant characteristics (e.g., gang affiliated vs. not), and type of expert testimony (focused on the vulnerability of the individual suspect vs. focused on the impact of the interrogation process). It may be that expert testimony is clearly effective for some combinations of cases or defendants, but less effective for other combinations. It would also be useful to know if safeguards other than expert testimony can help fact finders in their evaluation of confession evidence. Such safeguards might include jury instructions emphasizing how to evaluate confession evidence, or the use of effective closing arguments that highlight aspects of a confession that make it less reliable.

Conclusion

The expert witness who offers testimony on interrogations and confessions has the difficult job of answering a deceptively simple question: "Why would someone confess to a crime he or she did not commit?" To provide an answer, the expert must be able to clearly synthesize the large and complex body of research described above. When done well, expert testimony can help jurors and judges reach a fair, fully informed verdict. By educating the jury about the existence of false confessions, the psychological dynamics of interrogations, and the risk factors for false confessions, expert testimony at trial should reduce the number of confession-based wrongful convictions. Without the educational and sensitizing effect of expert testimony, judges and jurors may simply accept a confession at face value even though it may be coerced and false.

References

Adair, W. L., & Brett, J. M. (2004). Culture and negotiation processes. In J. M. Gelfand & J. M. Brett (Eds.), *The handbook of negotiation and culture* (pp. 158–176). Palo Alto, CA: Stanford University Press.

Appleby, S. C., Hasel, L. E., & Kassin, S. M. (2013). Police-induced confessions: An empirical analysis of their content and impact. *Psychology, Crime & Law, 19*(2), 111–128. doi:10.1080/1068316X.2011.613389

American Psychological Association. (2013). Specialty guidelines for forensic psychology. *American Psychologist, 68*, 7–19. doi:10.1037/a0029889

Berk-Seligson, S. (2009). *Coerced confessions: The discourse of bilingual police interrogations*. New York: Mouton de Gruyter.

Beune, K., Giebels, E., & Taylor, P. J. (2010). Patterns of interaction in police interviews: The role of cultural dependency. *Criminal Justice and Behavior, 37*(8), 904–925. doi:10.1177/0093854810369623

Blandon-Gitlin, I., Sperry, K., & Leo, R. (2011). Jurors believe interrogation tactics are not likely to elicit false confessions: Will expert witness testimony inform them otherwise? *Psychology, Crime & Law, 17*(3), 239–260. doi:10.1080/10683160903113699

Bond, C. F, Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*(3), 214–234.

Bond, C. F, Jr., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin, 134*(4), 477–492. doi:10.1037/0033-2909.134.4.477

Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin, 119*(1), 111–137. doi:10.1037/0033-2909.119.1.111

Chojnacki, D., Cicchini, M., & White, L. T. (2008). An empirical basis for the admission of expert testimony on false confessions. *Arizona State Law Journal, 40*(1), 1–45.

Cleary, H. M. D. (2014). Police interviewing and interrogation of juvenile suspects: A descriptive examination of actual cases. *Law and Human Behavior, 38*(3), 271–282. doi:10.1037/lhb0000070

Collins v. Brierly. (1974). 492 F.2d 735 (3d Cir.1974).

Costanzo, M. (2004). *Psychology applied to law*. Belmont, CA: Wadsworth.

Costanzo, M., & Leo, R. (2008). Research findings and expert testimony on police interrogations and confessions to crimes. In M. Costanzo, D. Krauss, & K. Pezdek (Eds.), *Expert psychological testimony for the courts* (pp. 69–98). Mahwah, NJ: Erlbaum.

Costanzo, M. (1992). Training students to decode verbal and nonverbal behavior: Effects on confidence and performance. *Journal of Educational Psychology, 84*, 303–313.

Costanzo, M., & Peterson, J. (1994). Attorney persuasion in the capital penalty phase: A content analysis of closing arguments. *Journal of Social Issues, 50*(2), 125–147. doi:10.1111/j.1540-4560.1994.tb02413.x

Costanzo, M., Shaked-Schroer, N., & Vinson, K. (2010). Juror beliefs about police interrogations, false confessions, and expert testimony. *Journal of Empirical Legal Studies, 7*(2), 231–247. doi:10.1111/j.1740-1461.2010.01177.x

Crane v. Kentucky. (1986). 476 U.S. 683.

Cutler, B. L. (Ed.). (2009). *Expert testimony on the psychology of eyewitness identification*. New York: Oxford University Press.

Cutler, B., Findley, K., & Loney, D. (2014). Expert testimony on interrogation and false confession. *UMKC Law Review, 82*(3), 1–36.

Cutler, B. R., Penrod, S. & Dexter, H. R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior, 14*, 185–191.

Daubert v. Merrell Dow Pharmaceuticals, Inc. (1993). 113 S. Ct. 2786, 15–16, 21, 58.

Davis, D. (2010). Lies, damned lies, and the path from police interrogation to wrongful conviction. In M. H. Gonzales, C. Tavris, & J. Aronson (Eds.), *The scientist and the humanist: A festschrift in honor of Elliot Aronson* (pp. 211–247). New York: Psychology Press.

Davis, D., & Leo, R. A. (2012a). Acute suggestibility in police interrogation: Self-regulation failure as a primary mechanism of vulnerability. In A. Ridley, F. Gabbert, & D. J. La Rooy (Eds.), *Suggestibility in legal contexts: Psychological research and forensic implications* (pp. 171–197). New York: Wiley.

Davis, D., & Leo, R. A. (2012b). Interrogation-related regulatory decline: Ego depletion, failures of self-regulation, and the decision to confess. *Psychology, Public Policy and Law, 18*(4), 673–704. doi:10.1037/a0027367

Davis, D., & Leo, R. A. (2012c). To walk in their shoes: The problem of missing, misrepresented and misunderstood context in judging criminal confessions. *New England Law Review, 46*(4), 737–767.

Davis, D., & O'Donohue, W. T. (2004). The road to perdition: "Extreme influence" tactics in the interrogation room. In W. T. O'Donohue & E. Levensky (Eds.), *Handbook of forensic psychology* (pp. 897–996). New York: Elsevier Academic Press.

Davis, D., Williams, M. J., & Villalobos, J. G. (2013, March). *Interrogation-related regulatory decline: The roles of prior effort and stereotype threat*. Paper presented at the 2013 American Psychology-Law Society conference. Portland, OR.

Domanico, A. J., Cicchini, M. D., & White, L. T. (2012). Overcoming Miranda: A content analysis of the Miranda portion of police interrogations. *Idaho Law Review, 49*, 1–22.

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

Feld, B. C. (2013). Real interrogation: What actually happens when cops question kids. *Law & Society Review, 47*(1), 1–36. doi:10.1111/lasr.12000

Follette, W. C., Davis, D., & Leo, R. A. (2007). Mental health status and vulnerability to interrogative influence. *Criminal Justice, 22*(3), 42–49.

French, J. R. P, Jr., & Raven, B. H. (1959). The bases of social power. In D. Cartwright (Ed.), *Studies in social power* (pp. 150–167). Ann Arbor: University of Michigan Press.

Frumkin, I. B., Lally, S. J., & Sexton, J. E. (2012). The Grisso tests for assessing understanding and appreciation of Miranda warnings with a forensic sample. *Behavioral Sciences & The Law, 30*(6), 673–692. doi:10.1002/bsl.2018

Frye v. United States. (1923). 293 F. 1013, 1014 (D.C. Cir. 1923).

Finlay, W. M. L., & Lyons, E. (2002). Acquiescence in interviews with people who have mental retardation. *Mental Retardation, 40*(1), 14–29. doi:10.1352/0047-6765

Fulero, S. (2004). Expert psychological testimony on the psychology of interrogations and confessions. In G.D. Lassiter (Ed.), *Interrogations, Confessions , and Entrapment* (pp. 247–263). New York: Kluwer Press.

Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: An experimental study. *Developmental Psychology, 41*(4), 625–635. doi:10.1037/0012-1649.41.4.625

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

Garrett, B. L. (2010). The substance of false confessions. *Stanford Law Review, 62*, 1051–1119.

Gross, S. R., Jacoby, K., Matheson, D. J., Montgomery, N., & Patel, S. (2005). Exonerations in the United States, 1989 through 2003. *Journal of Criminal Law and Criminology, 95*, 523–553.

Gudjonsson, G. H. (1999). *The psychology of interrogations, confessions and testimony*. Chichester, England: Wiley.

Gudjonsson, G. H. (1993). Confession evidence, psychological vulnerability and expert testimony. *Journal of Community & Applied Social Psychology, 3*(2), 117–129. doi:10.1002/casp.2450030205

Gudjonsson, G. H. (2014). Mental vulnerabilities and false confession. In M. St-Yves (Ed.), *Investigative interviewing: The essentials* (pp. 191–222). Toronto: Carswell.

Gudjonsson, G. H. (2003). *The science of interrogations and confessions: A handbook*. Chichester, England: Wiley.

Gudjonsson, G. H., Sigurdsson, J. F., Sigurdardottir, A. S., Ssteinthorsson, H., & Sigurdardottir, V. M. (2014). The role of memory distrust in cases of internalized false confession. *Applied Cognitive Psychology, 28*(3), 336–348.

Guyll, M., Madon, S., Yang, Y., Lannin, D. G., Scherr, K., & Greathouse, S. (2013). Innocence and resisting confession during interrogation: Effects on physiologic activity. *Law and Human Behavior, 37*(5), 366–375. doi:10.1037/lhb0000044

Heath, C., & Heath, D. (2013). *Decisive*. New York: Crown.

Henkel, L. A. (2008). Jurors' reactions to recanted confessions: Do the defendant's personal and dispositional characteristics play a role? *Psychology, Crime & Law, 14*(6), 565–578. doi:10.1080/10683160801995247

Henkel, L. A., Coffman, K. A. J., & Dailey, E. M. (2008). A survey of people's attitudes and beliefs about false confessions. *Behavioral Sciences & The Law, 26*(5), 555–584. doi:10.1002/bsl.826

Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. *Online Readings in Psychology and Culture*, *2*(1). Retrieved from http://dx.doi.org/10.9707/2307-0919.1014

Hofstede, G. (2010). The GLOBE debate: Back to relevance. *Journal of International Business Studies, 41*, 1339–1346. doi:10.1057/jibs.2010.31

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: Sage.

Huntley, J. E., & 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. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2013). *Criminal interrogation and confessions* (5th ed.). Burlington, MA: Jones & Bartlett Learning.

Kassin, S. M. (2007). Internalized false confessions. In M. P. Toglia, J. D. Read, D. R. Ross, & R. C. L. Lindsay (Eds.), *Handbook of eyewitness memory: Memory for events* (Vol. 1, pp. 169–187). Mahwah, NJ: Erlbaum.

Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions: Risk factors and recommendations. *Law and Human Behavior, 34* (1), 3–38. doi:10.1007/s10979-009-9188-6

Kassin, S. M., & Fong, C. T. (1999). I'm Innocent!: Effects of training on judgments of truth and deception in the interrogation room. *Law and Human Behavior, 23*(5), 499–516. doi:10.1023/A:1022330011811

Kassin, S. M., Leo, R. A., Meissner, C. A., Richman, K. D., Colwell, L. H., Leach, A., & La Fon, D. (2007). Police interviewing and interrogation: A self-report survey of police practices and beliefs. *Law and Human Behavior, 31*, 381–400.

Kassin, S. M., Meissner, C. A., & Norwick, R. J. (2005). I'd know a false confession if I saw one: A comparative study of college students and police investigators. *Law and Human Behavior, 29*(2), 211–227. doi:10.1007/s10979-005-2416-9

Kassin, S. M., & Neumann, K. (1997). On the power of confession evidence: An experimental test of the fundamental difference hypothesis. *Law and Human Behavior, 21*(5), 469–484. doi:10.1023/a:1024871622490

Kassin, S. M., & Sukel, H. (1997). Coerced confessions and the jury: An experimental test of the "harmless error" rule. *Law and Human Behavior, 21*(1), 27–46. doi:10.1023/a:1024814009769

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review, 110*(2), 265–284. doi:10.1037/0033-295x.110.2.265

Lamb, M. E., Malloy, L. C., & La Rooy, D. J. (2011). Setting realistic expectations: Developmental characteristics, capacities, and limitations. In M. E. Lamb, D. J. La Rooy, L. C. Malloy, & C. Katz (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 15–48). Hoboken: Wiley-Blackwell.

Lassiter, G. D. (2010). Psychological science and sound public policy: Video recording custodial interrogations. *American Psychologist, 65*(8), 768–779.

Leo, R. A. (2008). *Police interrogation and American justice*. Cambridge: Harvard University Press.

Leo, R. A. (2001). Questioning the relevance of Miranda in the twenty-first century. *Michigan Law Review, 99*(5), 1000–1029. doi:10.2307/1290525

Leo, R. A., Costanzo, M., & Shaked-Schroer, N. (2009). Psychological and cultural aspects of interrogation and false confessions: Using research to inform legal decision making. In D. Krauss & J. Lieberman (Eds.), *Psychological expertise in court: Psychology in the courtroom* (Vol. 2, pp. 25–56). London: Ashgate.

Leo, R. A., & Liu, B. (2009). What do potential jurors know about police interrogation techniques and false confessions? *Behavioral Sciences & The Law, 27*(3), 381–399. doi:10.1002/bsl.872

Madon, S., Yang, Y., Smalarz, L., Guyll, M., & Scherr, K. C. (2013). How factors present during the immediate interrogation situation produce short-sighted confession decisions. *Law and Human Behavior, 37*(1), 60–74. doi:10.1037/lhb0000011

Malloy, L. C., Shulman, E. P., & Cauffman, E. (2014). Interrogations, confessions, and guilty pleas among serious adolescent offenders. *Law and Human Behavior, 38*(2), 181–193. doi:10.1037/lhb0000065

Martin, A. (2011, November 25). The prosecution's case against DNA. *New York Times*.

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

McKee, R. (1997). *Story*. New York: Harper-Collins.

Miranda v. Arizona. (1966). 384 U.S. 436.

Najdowski, C. J. (2011). Stereotype threat in criminal interrogations: Why innocent black suspects are at risk for confessing falsely. *Psychology, Public Policy and Law, 17*(4), 562–591. doi:10.1037/a0023741

Najdowski, C. J. (2013, March). *Stereotype threat in police encounters: Why African Americans are targeted as criminal suspects*. Presented at the annual meeting of the American Psychology-Law Society, Portland, Oregon.

Najdowski, C. J., & Bottoms, B. L. (2012). Understanding jurors' judgments in cases involving juvenile defendants: Effects of confession evidence and intellectual disability. *Psychology, Public Policy and Law, 18*(2), 297–337. doi:10.1037/a0025786

Najdowski, C. J., Bottoms, B. L., Goff, P. A., & Spanton, J. (2012, March). *Do hispanics experience stereotype threat in police encounters?* Presentation at the annual meeting of the American Psychology-Law Society, San Juan, Puerto Rico.

Newport, F. (2014). Black and white attitudes toward police. http://www.gallup.com/poll/175088/gallup-review-black-white-attitudes-toward-police.aspx

Oi, M. (2013). Japan crime: Why to people falsely confess. BBC News. http://www.bbc.com/news/magazine-20810572

Olsen-Fulero, L., & Fulero, S. M. (1997). Commonsense rape judgments: An empathy-complexity theory of rape juror story making. *Psychology, Public Policy, and Law, 3*, 402–427.

Pennington, N., & Hastie, R. (1993). The story model for juror decision-making. In R. Hastie (Ed.), *Inside the juror: The psychology of juror decision-making* (pp. 192–222). Cambridge: Cambridge University Press.

People v. McRae. (1978). 23 Crim. L. Rep. (BNA) 2507 (N.Y. Sup. Ct. 1978).

Perske, R. (2011). Perske's list: False confessions from 75 persons with intellectual disability. *Intellectual and Developmental Disabilities, 49*(5), 365–373. doi:10.1352/1934-9556-49.5.365

Raven, B. H., Schwarzwald, J., & Koslowsky, M. (1998). Conceptualizing and measuring a power/interaction model of interpersonal influence. *Journal of Applied Social Psychology, 28*(4), 307–332. doi:10.1111/j.1559-1816.1998.tb01708.x

Redlich, A. D. (2004). Mental illness, police interrogations, and the potential for false confession. *Psychiatric Services, 55*, 19–21. doi:10.1176/appi.ps.55.1.19

Redlich, A. D., & Goodman, G. S. (2003). Taking responsibility for an act not committed: The influence of age and suggestibility. *Law and Human Behavior, 27*(2), 141–156. doi:10.1023/a:1022543012851

Redlich, A. D., Kulish, R., & Steadman, H. J. (2011). Comparing true and false confessions among persons with serious mental illness. *Psychology, Public Policy and Law, 17*(3), 394–418. doi:10.1037/a0022918

Ridley, A. M., Gabbert, F., & La Rooy, D. J. (2013). Suggestibility in legal contexts: What do we know? In A. M. Ridley, F. Gabbert, & D. J. La Rooy (Eds.), *Suggestibility in legal contexts: Psychological research and forensic implications* (pp. 217–228). Hoboken: Wiley Blackwell.

Risinger, M. D. (2000). Navigating expert reliability: Are criminal standards of certainty being left on the dock? *Albany Law Review, 64*(1), 99–152.

Rogers, R. (2011). Getting it wrong about Miranda rights: False beliefs, impaired reasoning, and professional neglect. *American Psychologist, 66*(8), 728–736.

Rogers, R., Correa, A. A., Hazelwood, L. L., Shuman, D. W., Hoersting, R. C., & Blackwood, H. L. (2009). Spanish translations of Miranda warnings and the totality of the circumstances. *Law and Human Behavior, 33*(1), 61–69. doi:10.1007/s10979-008-9129-9

Rogers, R., Gillard, N. D., Wooley, C. N., & Fiduccia, C. E. (2011). Decrements in Miranda abilities: An investigation of situational effects via a mock-crime paradigm. *Law and Human Behavior, 35*(5), 392–401. doi:10.1007/s10979-010-9248-y

Rogers, R., Steadham, J. A., Fiduccia, C. E., Drogin, E. Y., & Robinson, E. V. (2014). Mired in Miranda misconceptions: A study of legally involved juveniles at different levels of psychosocial maturity. *Behavioral Sciences & The Law, 32*(1), 104–120. doi:10.1002/bsl.2099

Ross, L. D., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. New York: McGraw-Hill.

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

Scherr, K.C. & Madon, S. (2012). You have the right to understand: The deleterious effect of stress on suspects' ability to comprehend Miranda.*Law and Human Behavior, 36*, 275–282.

Scherr, K. C., & Madon, S. (2013). Go ahead and sign: An experimental examination of Miranda waivers and comprehension. *Law and Human Behavior, 37*(3), 208–218. doi:10.1037/lhb0000026

Scherr, K. C., Miller, J. C., & Kassin, S. M. (2014). Midnight confession: The effect of chronotype asynchrony on admissions of wrongdoing. *Basic and Applied Social Psychology, 6*(4), 321–328. doi:10.1080/01973533.2014.917974

Shulman, E. P., & Cauffman, E. (2013). Reward-biased risk appraisal and its relation to juvenile versus adult crime. *Law and Human Behavior, 37*(6), 412–423. doi:10.1037/lhb0000033

Sigurdsson, J.F., & Gudjonsson, G.H. (1996). The psychological characteristics of 'false confessors'. A study among Icelandic prison inmates and juvenile offenders. *Personality and Individual Differences, 20*(3), 321–329.

Simpson, J. A., Farrell, A. K., Oriña, M. M., & Rothman, A. J. (2015). Power and social influence in relationships. In M. Mikulincer, P. R. Shaver, J. A. Simpson, J. F. Dovidio (Eds.), *APA handbook of personality and social psychology: Interpersonal relations* (Vol. 3, pp. 393–420). Washington, DC: American Psychological Association. doi:10.1037/14344-015

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test-performance of African-Americans. *Journal of Personality and Social Psychology, 69*(5), 797–811. doi:10.1037/0022-3514.69.5.797

Thomas, T. N. (1995). Polygraphy and coerced-compliant false confession: 'Serviceman E' redevivus. *Science & Justice, 35*, 133–139. doi:10.1016/S1355-0306(95)72646-6

United States v. Hall. (1996). 93 F.3d 1337 (7th Cir. 1996).

United States v. Hall. (1997). 974 F. Supp. 1198 (1997).

Villalobos, J. G. & Davis, D. (in press). Interrogation and the minority suspect: Pathways to true and false confession. In M. K. Miller & B. H. Bornstein (Eds.) *Advances in psychology and law* (Vol. I).

Vidmar, N., & Hans, V. (2007). *American juries: The verdict*. New Haven, CT: Yale University Press.

Watt, L. (2014). *China reveals cases of torture in police detention*. Associated Press.

Woody, W. D., & Forrest, K. D. (2009). Effects of false-evidence ploys and expert testimony on jurors' verdicts, recommended sentences, and perceptions of confession evidence. *Behavioral Sciences & The Law, 27*(3), 333–360. doi:10.1002/bsl.865

Zimbardo, P. (2007). *The Lucifer effect*. New York: Random House.

Psychology and the Federal Rules of Evidence

Amy Kleynhans and Brian H. Bornstein

The primary function of the Federal Rules of Evidence (FRE), which came into effect on June 1, 1975, was to "enact into law the will and intent of the Supreme Court and its Advisory Committee" and to "operate as guidance for the exercise of discretion" within the judicial system (Weissenberger 1992, p. 1307). The FRE govern all cases in federal courts, be they criminal or civil. State courts have generally either adopted the FRE or have enacted their own similar state rules. This chapter will concentrate on the FRE, with the understanding that individual state rules might differ slightly across jurisdictions. Many of the rules are based on assumptions about the ways in which those involved in the legal process (e.g., judges, jurors, litigants, or attorneys) engage in decision making. Others were created as reflections of Congress' policy decisions and do not necessarily take into account the impact such rules could have on legal actors.

Very little psychological research has addressed the impact and assumptions of the FRE. What research has been conducted tends to focus mainly on specific aspects of the rules. For example, the majority of the research has focused on Rule 702, which addresses testimony by expert witnesses (see Blau 1998; Dror et al. 2013; Faust and Ziskin 1988; Kassin and Wrightsman 1985). This is not surprising, as expert testimony is an area that is relevant and applicable to psychologists in a number of capacities—both those who study areas related to law (e.g., eyewitness testimony, interrogation practices) and those who perform psychological evaluations for the courts. These research and practicing psychologists could be called into the courtroom to testify, and therefore, it is important to understand how best to do so. However, 702 has little to do with behavioral assumptions of the law of evidence.

A. Kleynhans (✉)
401 N. 44th St., Apt. 1108, Lincoln, NE 68503, USA
e-mail: ajkleynhans@gmail.com

B.H. Bornstein
Department of Psychology, UNL, 335 Burnett Hall, Lincoln, NE 68588-0308, USA
e-mail: bbornstein2@unl.edu

B.H. Bornstein and M.K. Miller (eds.), *Advances in Psychology and Law*,
Advances in Psychology and Law 2, DOI 10.1007/978-3-319-43083-6_6

This chapter will examine the psychological implications for jurors and verdict outcomes of the FRE. It will begin by examining some of the previous research related to the FRE and then move onto discussing specific rules that are contrary to or do not seem to take into account psychological literature. Then, we will examine rules that seem to rely on and are supported by psychological research. Finally, we will look to some of the rules and areas of evidence law in which psychological research is lacking.

The Rules and the Current Body of Literature: 401 Relevance and 403 Prejudice

Before any piece of evidence may be admitted into the courtroom, it must pass Rule 401, which governs the standard of general admissibility for all evidence in the courtroom. This rule states that only relevant evidence is admissible. Evidence is relevant if it has any tendency to make a fact more or less probable than it would be without the evidence and if this fact is of consequence in determining the action (FRE 2004). However, whereas irrelevant evidence is always inadmissible, relevant evidence is not necessarily always admissible. One bar to admissibility is Federal Rule of Evidence 403, which states that the court may exclude relevant evidence if its probative value is substantially outweighed by a danger of unfair prejudice. The Notes of the Advisory Committee on the Proposed Rules define unfair prejudice to mean an undue tendency to suggest a decision would be made on an improper basis, commonly, though not necessarily, an emotional one. This reflects that Congress is not unaware of the impact that emotions can have on members of the jury. Criminal defendants have a constitutional right to be considered innocent until proven guilty beyond a reasonable doubt and a right to an impartial jury and fair trial. This means that jurors should be finders of fact who are not unduly swayed by the emotion-laden nature of testimony or other such evidence, but are influenced only by the strength of the case at hand. Thus, Congress drafted the FRE in an attempt to protect the rights of parties and in recognition that jurors might rely on their emotions to make quick bias-driven decisions, compromising their impartiality, and harming the parties. Rule 403 provides courts with a discretionary tool to exclude evidence that would otherwise be admissible under the rules but could evoke such emotion in the jury and lead to an unfairly prejudicial outcome (Gold 1984).

Much of the research in psychology and law that examines whether a specific argument or piece of evidence would influence the verdict outcome ultimately examines whether that item is unduly prejudicial. For example, studies show that a defendant's religion might affect verdict decisions (see Bornstein and Miller 2009; Johnson 1985). To apply this research to the courtroom, lawyers would need to rely on a 403 argument to bar any evidence of religion if it could negatively affect their clients.

Issues arising under 403 are brought to the judge for a determination of whether the evidence would promote or undermine accurate fact-finding by the jury (Gold

1984). This means that the judge must make decisions that require an understanding of human decision making. This clearly raises the question of whether judges have the requisite knowledge to make such decisions. Judges typically are not psychologists and will most likely rely on their own experiences and reactions to the type of evidence in question. A study by Teitelbaum et al. (1983) examined the validity of the premise that judges are accurately able to determine whether a piece of evidence will ultimately be prejudicial to the jury. Using lawyers as stand-ins for judges, the researchers compared lawyer ratings to those of community members to determine perceptions of how prejudicial specific items of evidence would be to a defendant's case. Overall, the community ratings of prejudice were almost always significantly lower than those of the attorney group. Aside from the obvious problem of using lawyers to represent judges (the researchers found defense lawyers rated items more prejudicial than did plaintiffs' lawyers), this result requires some consideration. Even if the lawyer results do correspond with judicial determinations, then this indicates that judges are likely to exclude evidence that members of the jury would not find unduly prejudicial. This overinclusivity may be good, if one believes that it is better to err on the side of caution and keep out items that could prejudice a jury. However, it also means that juries are not seeing or hearing evidence that could be crucial in arriving at an accurate decision. It is also important to note that this study relied on self-report of beliefs that an item is prejudicial and, as will be discussed later, jurors may not always be aware that their verdict is being influenced by a prejudicial piece of evidence. Judges might be aware of this natural tendency and correct for it by ruling more stringently and keeping potentially prejudicial evidence out. Further research is clearly required in this area.

One example of a hotly contested 403 issue arose in the trial of Casey Anthony, a woman accused of murdering her two-year-old daughter Caylee (*State of Florida v. Anthony* 2011). During the trial, the prosecution entered into evidence a video, depicting an image of Casey and Caylee together, which then morphed Caylee's face into a skull (Black 2011; see Fig. 1). The defense objected, but the judge deemed the tape highly relevant, which inspired many legal scholars to question how this could be relevant and not prejudicial under both a 401 and 403 argument. The picture of Caylee and Casey could be admitted as relevant to provide background information and set the context of a mother and her daughter. The image of

Fig. 1 Caylee Anthony image presented to the jury

Caylee's skull could also be relevant to show a specific pattern of wounds or disintegration. However, the introduction of the skull superimposed over the child's picture arguably has no relevance in and of itself and even if it did, it is potentially prejudicial to the defendant. The prosecution seems to have introduced the image, seen below, purely in an attempt to influence the jury's emotion. Casey Anthony was ultimately acquitted, but had she been found guilty, she would likely have had a strong argument for a new trial on appeal.

Courts often deem graphic photographs to be relevant and admissible, as gruesome crimes often beget gruesome photographs in the course of the investigation (Bright and Goodman-Delahunty 2004). To the prosecution, gruesome pictures might be important evidence to point to the defendant's guilt, whereas the defense may view that same evidence as inflammatory, prejudicial, and violating the defendant's right to a fair and impartial trial (Douglas et al. 1997). Multiple psychological studies have examined the impact of gruesome images on verdict outcomes, suggesting that emotional reactions to such horrifying evidentiary details could inhibit logical, rational processing, and draw jurors' attention away from evidence with more probative value (Bright and Goodman-Delahunty 2006). This research finds that such vivid pictures are more persuasive than pallid evidence and could invoke moral outrage that leads viewers to demand accountability and responsibility (Bell and Loftus 1985; Thornton et al. 1991).

Douglas et al. (1997) examined the prejudicial effect that graphic images might have on the verdicts of mock jurors. Participants were assigned to one of three conditions: Color, Black and White, or Control. Those in the control group only read a graphic description of the victim's body in the medical examiner's testimony that contained all the same information that could be viewed in the photographs. In the Color and Black and White conditions, participants read the same description but also viewed three explicit photographs of an actual homicide victim, either in color or in black and white. In addition, participants completed questionnaires measuring their emotional reactions to viewing the evidence. Participants exposed to the victim photographs were almost twice as likely to find the accused guilty as participants in the control condition. The two photograph conditions did not differ from one another. Finally, the more that participants reported feeling sad, vengeful, outraged, shocked, and anxious in response to the photographs, the more likely participants were to find the defendant guilty. This result means that, while gruesome photographs might be relevant to the case, defense attorneys should rely on Rule 403 to argue that they are unduly prejudicial to their client's case. Despite being presented with the same factual information, the simple presence of photographs invoked such emotion in the mock jurors that the defendant was far more likely to be found guilty. Potentially the most important result of Douglas et al.'s study was that when they asked participants whether they felt they acted in a fair and unbiased manner and whether the autopsy information, be it photographs or testimony, affected their verdict, there was no difference across the conditions. This means that, despite a clear prejudicial impact of the photographs on the verdicts, participants were unaware of this influence. Participants believed they were deciding based only on the facts. Not all studies of the phenomenon have found a

prejudicial effect of gruesome photographic evidence, however (for review, see Bornstein and Greene 2017; Bornstein and Nemeth 1999).

Rules Conflicting with Psychology Research

Although the majority of the FRE are based on assumptions about human behavior and decision making, some of these assumptions do not reflect rigorous scientific research. This section will examine two of these problematic rules, focusing on their assumptions and the conflicting research that suggests that the reasons for their implementation may be flawed.

404: Character Witness Evidence

The first rule we will discuss is Rule 404, which pertains to character evidence. This rule prohibits the use of "evidence of a person's character or character trait…to prove that on a particular occasion the person acted in accordance with the character or trait" (FRE 404(a)(1); 2011). This is also known as the propensity argument, because it may prejudice the trier of fact into believing that somebody is of such a type of character that he or she is likely to act in accordance with that trait at all times and therefore is more likely to have committed the act in question. In addition, the assumption is that character evidence, if admitted, may subtly permit the trier of fact to reward a defendant for good behavior or punish him for bad behavior, despite the actual facts.

Despite this prohibition, there are exceptions that allow character evidence to be admitted. The most common exception applies to defendants in criminal cases. Rule 404(a)(2)(A) allows defendants to offer evidence of their own character traits related to the alleged crime, and only then is the prosecutor allowed to offer rebuttal evidence. If a defendant chooses to introduce evidence of his character, he may generally do so only by reputation or opinion evidence (FRE 405(a); 2011). This means that the defendant will need to bring in a character witness who will provide the jury or judge with background information about the defendant (Hunt and Budesheim 2004). When presented, however, this form of evidence may not include specific examples but only general impressions of the defendant's character. Once introduced, the prosecutor may cross-examine the character witness by asking him or her about relevant specific instances. For example, if in a trial for assault, the witness for the defendant testifies that her opinion of the defendant is that he is peaceful and calm, the prosecutor may ask the witness on cross-examination about whether the witness knows about the last five bar fights that the defendant instigated. This would likely be allowed because it tests the credibility of the witness. If she did not know about those bar fights, then her testimony might not be reliable or trustworthy. Jurors are not supposed to use this evidence of specific acts to infer the

defendant's guilt (in accordance with the 404 propensity rule), but only to evaluate the witness's credibility. In addition, the prosecution may now call its own witness to testify that the defendant has a reputation of being violent and angry. If the defendant had not called the original witness, the prosecutor would have been unable to do so. Whether or not the defense chooses to introduce the defendant's character into evidence is therefore an important consideration because it may open the door to contrary evidence and specific negative examples that would otherwise have been excluded. Thus, psychological research on the jury's perception of character evidence is crucial for the defense to know so that they can decide if introducing the evidence is a risk worth taking.

Hunt and Budesheim (2004) stated that admissibility of such evidence is predicated on three assumptions: (1) That the defendant's personality characteristics are relevant to determining whether or not he committed the act in question; (2) that jurors actually use the character evidence when evaluating a case against the defendant; and (3) that cross-examination or rebuttal evidence will reduce the credibility of the character witness without negatively biasing impressions of the defendant.

For the first assumption, personality psychologists have examined whether people do behave consistently. In general, personality traits are inconsistent and cannot accurately predict behavior across different situations (Mischel 1969; Peterson 1968; Vernon 1964).

For the second assumption, some research has shown that descriptions of personality traits do not necessarily influence impressions of the target and instead might be more likely to be used to form an impression of the person providing the description (Wyer et al. 1994). In addition, it is important to evaluate the combination of the first and second assumptions. When jurors are presented with such evidence, do their interpretations reflect that people do not generally behave consistently across different situations? Wissler and Saks (1985) investigated how prior conviction evidence, which we will discuss in more detail in the next section, influenced mock jurors' assessment of credibility and guilt. They manipulated the type of crime for which the defendant had received a prior conviction. They found that mock jurors were significantly more likely to convict when the prior conviction was the same as the present charge than when the previous conviction was for a dissimilar crime. Therefore, jurors assumed that because the defendant's behavior related to the current charge was consistent with his previous behavior, he had a propensity to commit that specific type of crime. Therefore jurors may actually use character evidence, but do so in a manner inconsistent with research that people act inconsistently.

For the third assumption listed above, research suggests that on balance, the prosecution might have an unfair advantage due to their ability to cross-examine with specific examples, whereas the defendant may use only general information (Miller and Burgoon 1982). The lack of detail that the defense witness provides could lead jurors to believe that the witness does not have detailed, credible information about the defendant, particularly as it is unlikely that jurors are aware that the defense is not allowed to use specific examples.

Hunt and Budesheim (2004) conducted a study to examine the impact of character evidence (hereinafter referred to as "CE") on jurors. They manipulated the type of CE presented by the defense: no CE, general CE, or specific, positive CE. They also manipulated the type of CE brought by the prosecution: no CE, specific bad acts cross-examination, or a rebuttal reputation/opinion witness. The results indicated that the character witness was viewed more favorably when presenting specific and detailed evidence than when presenting general evidence. Further, jurors rated the witness more negatively when the prosecutor introduced rebuttal CE. In addition, when participants were exposed to witnesses who provided positive CE, it did not reduce conviction judgments, but when they saw witnesses who were cross-examined about specific bad acts, they were significantly more likely to find the defendant guilty. This means the defense is always at a disadvantage when they introduce character evidence because positive CE did not reduce guilty verdicts, but specific bad acts CE did increase guilty verdicts.

In sum, the second of Rule 404's assumptions holds true in that jurors are using the defendant's personality characteristics in evaluating the case against the defendant. However the first and third assumptions do not hold true. For the first, as people generally do not behave consistently across different situations, character evidence may not be relevant to determinations of guilt. For the third, jurors are not just using the prosecution's rebuttal or cross-examination evidence only to reduce the witness's credibility; they are also using it as inadmissible propensity evidence and to infer the defendant's guilt. This suggests that defendants need to be particularly cautious about introducing CE because once they do, 404 allows the prosecution to introduce evidence that can harm them more than their initial CE helps their case.

608 and 609: Prior Convictions

Rule 404 has another exception that allows the admission of character evidence for witnesses. Rule 607 states that "any party…may attack the witness's credibility." These attacks, however, are limited under Rule 608 only to evidence related to the witness's truthfulness or untruthfulness. This is one of several methods for impeaching or discrediting the witness. Credibility cannot be attacked unless the person testifies as a witness, and evidence of truthfulness can only be introduced after the witness's credibility has been attacked. Whether or not the witness's testimony is honest and factual is crucial for the jury to discern, and therefore propensity evidence is admissible to help the jury determine whether the testimony could potentially be a lie. Specific examples are allowed during cross-examination if they are probative of truthfulness.

One approach to using this type of impeachment is through prior conviction evidence, which is governed by Rule 609. Conviction evidence is "significant… because it stands as proof of the commission of the underlying criminal act," because the witness was found beyond a reasonable doubt to have committed the act (Advisory Committee Notes 1972). When the defendant is the witness, this

might seem to contradict Rule 404(b)(1), in which evidence of prior bad acts is not admissible. This evidence is prohibited because the jury may punish the defendant for his previous behavior, for simply being a "bad" person, or they may assume that someone who has behaved in a specific way in the past is much more likely to do so in the future. In addition, possessing a prior criminal record may cause jurors to think that, since the defendant already has a criminal record, another conviction would not be as serious as if he had no prior convictions. Therefore jurors might require less evidence to be convinced that the defendant is guilty beyond a reasonable doubt (Cleary 1984). However, there are two ways in which prior conviction evidence would be deemed admissible. The first is under Rule 404(b)(2), which states that in a criminal case, evidence of a crime, wrong, or other act "may be admissible for another purpose [than character], such as proving motive, opportunity, intent, preparation, plan, knowledge, identity, absence of mistake, or lack of accident." For example, if the victim was killed by a gunshot and the defendant denied knowing how to use a gun, a prior conviction for armed robbery in which the defendant used a gun would be admissible to show knowledge.

Rule 609 allows some prior conviction evidence to be introduced for witness impeachment purposes. This means that as long as the conviction is not admissible for a purpose other than character, as laid out above in 404(b)(2), then any of the defendant's prior convictions will be inadmissible unless he chooses to testify. In enacting Rule 609, the legislators do appear to have taken into account psychological evidence of the prejudicial propensity effect that prior conviction evidence could have on a defendant. The Advisory Committee Notes for Rule 609 (1972) specifically state that the "rule incorporates certain basic safeguards, in terms applicable to all witnesses but of particular significance to an accused who elects to testify." For all witnesses other than the defendant, the conviction *must* be admitted when offered by the prosecution for any crime punishable by death or imprisonment for more than 1 year, regardless of the actual punishment received, subject to the prejudicial balancing test of 403. For criminal defendants, however, as opposed to all other witnesses including civil defendants, the rule applies a more stringent test. Whereas the 403 standard states that the evidence may be excluded if the probative value is substantially outweighed by prejudice, 609 states that prior conviction evidence must be admitted for criminal defendants who testify only if the probative value outweighs its prejudicial effect. Thus, the usual standard is reversed for these defendants who testify, such that prior conviction evidence is assumed to be highly prejudicial and is *not* admissible unless the probative value is so high and so crucial to the case that it cannot conceivably be excluded. This suggests that the law recognizes that prior conviction evidence is automatically dangerous to the defendant. However, if the conviction involved proving or admitting a dishonest act or false statement, such as for perjury, the conviction is automatically admissible regardless of whether the witness is the defendant. In these cases, the conviction is determined to be highly probative of a character for truthfulness and must therefore be admitted.

When admissible as impeachment evidence under Rule 609, the jury is expected to use the prior conviction only to determine whether the witness, potentially the

defendant, has a propensity to lie on the stand. However, it is impermissible for the jury to use the prior conviction to infer that the defendant has a propensity to commit crimes (Tanford and Cox 1988). Because the law does take into account the prejudicial effect this could have, the additional safeguard of limiting instructions is usually given to the jury. The instructions will explicitly state to the jury that they are to use the evidence only to assess credibility and not to assign blame.

There is a very large body of research on limiting instructions that this chapter will not examine in much detail, but it is relevant to discuss in relation to impeachment via prior conviction. Limiting instructions have generally been found to be ineffective in cases involving both inadmissible evidence and prior convictions (Hans and Doob 1976; Steblay et al. 2006; Thompson et al. 1981; Wissler and Saks 1985). There is some evidence for a backfire effect in which the instructions only make the inadmissible evidence more salient by focusing the jury's attention on what they are intended to ignore, thereby enhancing the prejudicial effect (Pickel 1995; Wolf and Montgomery 1977). Prior conviction evidence differs from inadmissible evidence, as the jury does not need to disregard it entirely, but use it for credibility purposes only and not propensity. Greene and Dodge (1995) found that mock jurors were more likely to convict and found the defendant less credible when prior conviction evidence was used for impeachment purposes, regardless of whether limiting instructions were provided. The authors surmised that the mock jurors may not have understood the instructions, which is a common determination in jury instruction research (see, e.g., Alvarez et al. 2016; Bornstein and Greene 2017; Dumas 2007; Lieberman 2009).

Tanford and Cox (1988) found that prior conviction impeachment evidence appeared to influence jury decision-making in legally impermissible ways. They manipulated prior conviction evidence and character evidence in a civil negligence trial, and their results indicated that prior convictions did not decrease credibility perceptions (the legally permissible and intended impact of such evidence), but they did increase perceptions of propensity (the legally impermissible use). They also found that evidence of a dishonest character lowered credibility judgments of the defendant witness as intended by the rules, but it also led to increased propensity judgments. Overall, propensity inferences increased the likelihood of liability verdicts. This means that both character and prior conviction evidence produced inappropriate propensity bias, and the latter did not even succeed in its intended purpose of impeaching the defendant's credibility.

In the same series of studies, Tanford and Cox (1988) also manipulated the use of limiting instructions, as well as whether their participants reached their verdicts alone or through group deliberations. They found that without the instructions, credibility judgments were not affected by the prior conviction. However, propensity ratings were higher when prior conviction evidence was provided, regardless of whether instructions were given. Therefore, providing limiting instructions in this study may have helped participants to use the evidence for its intended purpose of lowering credibility, but it did not deter them from using it for making propensity judgments. Participants who deliberated in groups showed even more bias than when they reached a verdict alone. This could be due to the impact

of group polarization, a phenomenon in which preexisting tendencies and beliefs are enhanced through group discussion (Arima 2013; Iyengar and Westwood 2014; Myers and Bishop 1971; Myers and Lamm 1976).

The results of these studies indicate that through the implementation of Rule 609 and its safeguards, Congress does seem to have an awareness of the possibly prejudicial impact of prior conviction evidence. As such, stricter guidelines may be required for admittance when the defendant is a witness. In the current form, the defense needs to be aware of the potential damage that can ensue from allowing the defendant to testify, as it opens the door for the prosecution to submit to the jury prior conviction evidence that would otherwise be excluded. As it stands, the stricter 609 balancing standard means that it is important that judges keep such evidence out through weighing the probative value against the presumed prejudice. However, as we know from the discussion earlier in this chapter, judges might not always accurately estimate the prejudicial value of the evidence. They may also overestimate the jury's ability to follow instructions appropriately and limit their application of the evidence.

Rules Corresponding to Psychology Research

Whereas the previous section discussed areas in which psychology has revealed problems with the assumptions made by the FRE, some of the rules do seem to be accurate in their assumptions about how humans behave. These FRE rules therefore serve as adequate safeguards for issues that could arise without the FRE protections. In this section, we will discuss some of these rules and how they correspond with psychology research.

406 Habit

Rule 406 applies to areas of habit and routine practice. The rule states that "evidence of a person's habit or an organization's routine practice may be admitted to prove that on a particular occasion the person or organization acted in accordance with the habit or routine practice." The Advisory Committee Notes (1972) acknowledge that while habit and character do appear to be similar, habit "in modern usage, both lay and psychological, is more specific. It describes one's regular response to a repeated specific situation." Character is noted as more akin to the sum of many individual habits that represent an overall tendency, whereas habit is discussed almost as conditioned responses. The specific behavior for a habit must be reflexive and automatic, such that anything appearing conscious and volitional is not likely to be admitted under Rule 406. In essence, the rule allows the trier of fact to predict behavior from actions in the past. For example, if a question of fact in a negligence case involving a car accident is whether or not the defendant used his

indicator before turning, the defense may introduce evidence, either through opinions or specific examples, that the defendant routinely used his blinker prior to turning. The jury may use this information to infer that because the defendant had a habit of using his blinker in the past, he was more likely to have used it on the day in question. This rule is important because the defendant himself may be unlikely to answer the question specifically of whether he did in fact use his indicator on the day of the accident, in the same way that many of us often arrive home after work and realize we were driving mindlessly and not paying conscious attention to the route that we drove home. We can still assume we followed the same route that we always take based on habit.

In this sense, while the previously discussed research finding that personality traits may differ across unique situations, this rule assumes that behavior is far more likely to be consistent in those individually unique situations (Mischel 1969; Peterson 1968; Vernon 1964). Research bolsters this assumption. Psychological theorizing about habit and its relation to consciousness goes back at least as far as James (1890). Psychology describes habitual behavior as "a goal-directed type of automaticity…[that is] instigated (by certain triggering stimuli) in the presence of a specific goal" (Aarts et al. 1998, p. 1358).

Aarts and Dijksterhuis (2000) suggested that once habits are established, priming the act then automatically evokes the habitual response. They recruited participants who all owned bicycles but differed in the frequencies with which they used them. In the first phase, participants in the goal priming condition read sentences describing five different travel goals, such as going shopping at the mall. Those in the control condition did not read these sentences and instead skipped straight to the second phase, in which all participants were asked to indicate as quickly as possible whether a specific presented mode of transportation was a realistic means of transport to reach a specific presented location. In the third phase, the researchers measured the participants' habit strength by asking how frequently they used their bicycles. The dependent variable in the study was the response latency across the five target location-transportation pairs in phase two. The results indicated a two-way interaction of habit strength and goal priming. When participants who were habitual bicycle users were not primed with travel goals, their response times did not differ from nonhabitual participants' response latencies. Conversely, participants in the goal priming condition showed significantly faster responses when they were habitual bicycle users than when they were not.

This indicates that habitual behaviors are automatic when a specific goal, such as traveling to the mall, is activated. Rule 406 corresponds with this research by assuming that if the party can provide evidence that a specific behavior in response to a stimulus (e.g., signaling when turning) has occurred continually in the past, then it is more likely that it did occur at the time in question despite the lack of actual evidence of its occurrence on the specific day.

407 Subsequent Remedial Measures, 408 Compromise Offers and Negotiations, 409 Offers to Pay Medical Expenses, and 411 Liability Insurance

Rule 407 refers to any potential measures that a party took subsequent to an earlier injury or harm that would have made that incident less likely to occur. For example, the plaintiff slips and falls in the defendant's store and the defendant consequently installs nonslip flooring. This evidence is inadmissible under 407 as proof that the defendant was actually negligent, though it could be admitted for another purpose. For example, if the defendant denies that he owned the store, the flooring change could be admitted to show that he does have ownership and control. Rule 408 mandates that any settlement and negotiation offers from one party to another, or any statements made in pursuance thereof, are inadmissible as evidence of the validity or amount of the disputed claim. Similarly, Rule 409 prohibits evidence of offers to pay medical expenses to prove liability for the injury.

All three rules are based on related public policy reasoning. For subsequent remedial measures in 407, proponents suggest that without the rule, people, and organizations would be deterred from taking safety precautions in response to accidents because otherwise plaintiffs would be able to use this information to bias the jury. It "also seems unfair…for a defendant to be penalized at a trial for taking the socially desirable action of decreasing risks" (Best 2009, p. 21). The purpose of Rule 408 is to encourage settlements. Without it, parties involved might be afraid to talk about disputes for fear that anything they say may be later used against them in court (Best 2009). Similarly to the reasoning behind Rule 407, Rule 409 was adopted so that those who offer payment of medical expenses for humanitarian reasons will not be subsequently punished for their compassionate offer (Best 2009; Mueller and Kirkpatrick 2013).

Rule 411 prohibits disclosure of any information that the party was or was not insured against liability. The Advisory Committee Notes (1972) state that "the knowledge of the presence or absence of liability insurance would induce juries to decide cases on improper grounds." Knowledge of a defendant's insurance could lead juries to be less concerned with reaching an erroneous decision and to find the defendant liable regardless of evidence of wrongdoing, since the defendant would not be paying for it personally. Alternatively, if the jury is informed that the *defendant* is uninsured, they may be hesitant to impose a large damage award that could potentially bankrupt the defendant. Additionally, if the jury is aware that the *plaintiff* has insurance, they may want to assign lower damages to avoid double recovery. All of these rules work to "blindfold" the jury and keep away information that lacks probative value and could bias their decisions (Greene et al. 2008).

These rules seem to assume two factors: that people want to behave in prosocial ways, and that jurors interpret these behaviors as indications of culpability. A number of studies have examined how people behave when they transgress against others; results indicate that participants tend to engage spontaneously in reconciliatory behavior such as confession, concession, and apology (Gonzales

et al. 1990, 1992; Ohbuchi et al. 1989). In addition, when participants believed they had committed a transgression, such as knocking over a confederate or breaking expensive equipment, they were more likely to attempt to compensate the victim (Konecni 1972; Regan 1971).

However, the question that remains is: How do jurors interpret this behavior? Do they automatically infer that these actions indicate guilt? There is some research that suggests that jurors perceive defendant apologies at the time of the incident as an admission of guilt, but that expressions of remorse in general may result in more favorable perceptions (Bornstein et al. 2002). For example, a study by Robbennolt (2013) showed that participants viewed offenders who offered apologies during settlement negotiations as more moral and as accepting of more responsibility than offenders who did not apologize. They also experienced a lesser desire to punish the offender. This indicates that expressions of remorse may cause jurors to be more likely to find the defendant guilty or liable, but to punish them less severely.

Thus, as stated above, Congress likely enacted these rules out of concern that potential defendants would be discouraged from acting prosocially for fear that jurors may interpret such behavior as inferring guilt. The purpose of these rules therefore is to encourage people to behave prosocially, and if an incident occurs, regardless of fault, to take steps to remedy the injury and prevent it from occurring again in the future. In other words, these rules acknowledge that people might engage in these behaviors when they feel guilty, but for those that are not guilty, Congress does not want to prevent them from acting simply because they could be punished later for conciliatory actions.

Some critics of these rules are concerned that "blindfolding" the jury will only work if jurors do not discuss these topics among themselves during deliberation and do not make their own potentially inaccurate assumptions regardless of the lack of information provided. Greene et al. (2008) tested the assumptions underlying 408's settlement offers and 411's insurance status. Based on a study by Diamond and Vidmar (2001) that indicated that jurors overwhelmingly and spontaneously raised these forbidden topics during their deliberations, Greene et al. assumed that the more frequently their mock jurors discussed the *plaintiff's* insurance status or previous settlements reached with other defendants, the lower the damages awarded would be, and the more frequently they discussed the *defendant's* insurance status, the higher the damage award would be. They found that overall, 83 % of their mock juries mentioned the plaintiff's insurance status, and 75 % referred to the defendant's; 23 % talked about pretrial settlements. Only 4 % of the juries refrained from discussing any of the forbidden topics during their deliberation. While not statistically significant, the relationship between insurance status discussions and the damage award did trend in the predicted direction. There was also no relationship between settlement discussions and damages.

The finding that discussion of forbidden topics did not significantly affect the award size indicates that, although the topics did appear to be meaningful to jurors, their discussion did not actually relate to compensatory damages. Possibly, the discussion of these factors highlights to the mock jurors that they do not truly know the details and therefore should not consider them in their verdicts. Greene et al.

(2008) listed some examples of the type of comments that their mock jurors made referencing these topics: "We don't know. We're assuming he has insurance"; "Because it is so vague, we don't know what kind of insurance the plaintiff has"; "We were not really given information one way or the other, so we have to decide on the information we have" (Greene et al. 2008, pp. 213–216). Further research is needed on the comparison between jury discussions, verdict outcomes and damages when juries are blindfolded from these topics and when they are not.

610 Religious Beliefs or Opinions

Rule 610 states simply that "evidence of a witness's religious beliefs or opinions is not admissible to attack or support the witness's credibility." In *U.S. v. Sampol* (1980), the D.C. Court of Appeals described the purpose of this rule as a guard against prejudice that may occur as a result of disclosure of the witness's faith, in particular those who subscribe to unconventional or unusual religions.

There is some psychological research that supports this reasoning. Henri Tajfel is well known for his work on social categorization and intergroup discrimination. This research suggests that social identification and affiliation with others leads to more favorable attitudes toward the ingroup over any outgroup in order to enhance the collective self-esteem (e.g., Jackson and Hunsberger 1999; Tajfel and Turner 1979). In one study, Tajfel et al. (1971) induced intergroup categorization. All participants were asked to estimate the number of dots projected onto a screen. Participants in the "neutral" condition were told that some people consistently overestimate, whereas others underestimate, but that the kind of estimation did not relate in any way to accuracy of judgments. Participants in the "value" condition were told that some people are consistently more accurate than others. Those in the "neutral" condition were then divided into two groups and told that one consisted of those who tended to overestimate and the other of those who tended to underestimate. In the "value" condition, they were told that one group included those with more accurate judgments and the other those with less accurate judgments. All of these groups were actually randomly assigned. Later, participants were asked to assign rewards individually to either members of their own ingroup (other overestimators/underestimators) or the outgroup. In both conditions, participants showed discriminatory intergroup behavior, consistently choosing to assign rewards to other members of their ingroup. These results indicate that even arbitrarily assigned social categorization can lead to preferential treatment of those appearing to belong to the same social group, which supports the rationale behind excluding irrelevant evidence of a witness's religion. Examples of intergroup discrimination have been consistently shown in a variety of contexts and across numerous forms of real world categorization, such as race, socioeconomic status, nationality, and religion (e.g., Dickter and Bartholow 2007; Falk and Zehnder 2007; Hart et al. 2000; Schwartz and Struch 1989; Shayo and Zussman 2011).

Some studies have shown that in the context of intergroup relations, religious people are unlikely to act prosocially and come to the aid of outgroups and, instead, discriminate against them (Batson et al. 1993; Goldfried and Miner 2002). Jackson and Hunsberger (1999) examined intergroup bias in a religion context, comparing Christian participants to those who reported no current religious affiliation. Their participants completed questionnaires on, among other things, religious group identification, Christian orthodoxy, and attitudes toward four outgroups ("atheists," Christians," "people who do not believe in God," and "people who do believe in God"). Consistent with intergroup discrimination theories, they found that religious group identification predicted positive attitudes toward religious others and negative attitudes toward nonreligious others. They also found that those who were more religious had very positive attitudes toward other Christians, and those who believed in God had slightly negative attitudes toward atheists and nonbelievers. Conversely, those who were less religious showed generally positive attitudes toward all groups. It is possible that for the latter, group identification was less strong and affiliation was less likely, leading to a lack of ingroup/outgroup bias. Together, this research suggests that the enactment of Rule 610 may help to eliminate instances of ingroup/outgroup bias, particularly for defendants who may be treated less favorably based on their religion or lack thereof.

However, Rule 610 applies to credibility impeachment for all witnesses, not just defendants or plaintiffs. Ingroup preferential bias may occur, but how does this translate to credibility? Does information about the witness's religion, or lack thereof, actually alter credibility judgments? Saroglou et al. (2005) examined the relationship among religion, prosocial behavior, and honesty. In their first study, they found that when presented with decision hypotheticals whether to act prosocially, religiosity was positively associated with helping, but this was limited only to close relations and not to unknown targets. In subsequent studies, participants completed altruism and honesty scales and then provided these scales to two of their self-selected peers to complete about the participant. The results showed that religiosity was related to self-reported altruism and honesty and that religious targets were perceived as relatively altruistic and honest by their peers. Therefore, this study provides one example that religion could bolster the credibility of a witness. It is important to note that peer religion was not reported, but as they were selected by the participant, it is likely that the peers belonged to the same religious ingroup. Thus, further research is still needed to determine whether credibility due to religion is consistent across all religious groups, both of the target and the perceiver. Whether or not a witness was perceived as more or less credible based on the witness's religion would presumably depend on the agreement between the religion of the individual juror and witness.

One pervasive current view is that ingroup biases generate positive evaluations of the ingroup, but this does not necessarily lead to negative derogation of the outgroup (see Otten and Wentura 1999; Perdue et al. 1990). It would be important to determine if this holds true in cases of witness religiosity, particularly for those with unconventional and unusual religious beliefs. Regardless, while further research is required on witness credibility across religious groups, Rule 610 does

seem to correspond with research on intergroup discrimination. However, Rule 610 does not address the problem that intergroup discrimination can still occur across social categories other than religion, such as race and ethnicity, which are more clearly apparent than a witness's religion.

Rules 801-807 Hearsay

Hearsay is a particularly difficult topic to tackle within this chapter. It is a complex concept that lawyers and law students alike struggle to master. Perhaps due to its difficulty, psychologists have largely ignored it, but it is based, at least in part, on testable psychological assumptions (Thompson and Pathak 1999).

Hearsay has generally been defined as a statement made out of court that is offered for the truth of the matter asserted (Mauet and Wolfson 2011). For example, suppose that Joseph called Ryan and said that he saw the defendant leaving the victim's house holding a gun. Joseph would be expected to testify on that point himself, but if he was not available to do so, it would be defined as hearsay if Ryan testifies to Joseph's comment, because the statement was offered to prove that the defendant did actually leave the victim's house with a gun. It would not be hearsay, however, if offered to show why Ryan rushed to the victim's house. In the latter case, the statement is admitted only to show the effect it had on Ryan, the listener, it does not matter whether or not the statement that Joseph saw the defendant leaving with a gun was true

Hearsay is generally inadmissible because it lacks three crucial truth-testing tools. First, the original declarant (or speaker) is not under oath, because he is not the one in court to report what he saw. Secondly, the trier of fact is unable to perceive the declarant's demeanor, and thirdly, the declarant is not subject to cross-examination. If Joseph was in court to testify to what he saw, the trier of fact would be able to evaluate his demeanor and only be required to determine whether or not he was telling the truth. Hearsay complicates this process by adding a step. If Ryan testifies to Joseph's statement, the trier now not only has to discern whether Joseph was telling the truth without being able to see him, but also has to figure out if Joseph did actually tell Ryan that he saw the defendant or if Ryan is lying or mistaken. There are a number of exceptions to the hearsay prohibition, such as the admissibility of dying declarations (e.g., Ryan would probably be allowed to testify about Joseph's declaration if Joseph died shortly after making it), but another whole chapter (or book!) would be required to go into more detail about the various hearsay exceptions.

For psychological purposes, the probative value of hearsay depends on specific testable factors (Thompson and Pathak 1999). The declarant must correctly have perceived, remembered, and reported the event in question. Then the listener must also have perceived the statements correctly, remember them, and report them to the jury (of course assuming that neither one is lying). Most of us played broken telephone when we were children and know first-hand that information can easily

become distorted as it passes orally from one person to another, and psychological research confirmed the phenomenon over 80 years ago (Bartlett 1932). As the information passes through each individual, it is processed and filtered through his or her specific cognitive lens and influenced and altered through each person's own expectations, stereotypes, and schemas (Nisbett and Ross 1980). Based on these ideas, the rules against hearsay evidence seem to reflect research suggesting that hearsay may be inaccurate and greatly mangled by the time it reaches the courtroom.

Although some research has shown that jurors are influenced by hearsay testimony (e.g., Golding et al. 1997), interestingly, and perhaps because of the widespread "broken telephone" understanding of the dangers of hearsay evidence, some research has shown that mock jurors are aware that hearsay evidence should be viewed negatively, and they generally ignore it or give little weight to it (Schuller 1995). For example, Kovera et al. (1991) had participants view eyewitness testimony alone, or eyewitness testimony in conjunction with hearsay testimony. The hearsay testimony consisted of graduate students who had viewed the eyewitnesses' videotaped responses answering the same questions as the eyewitnesses. Participants rated eyewitnesses and their testimony more positively than the hearsay witnesses and their testimony. They also were more skeptical of the hearsay testimony, viewing it as less accurate and less useful.

Thompson and Pathak (1999) suggest that psychologists next need to test assumptions about the ways in which jurors respond to such hearsay evidence and whether some of the statements admissible under hearsay exceptions are unusually reliable. Because so much of the hearsay rules seems to be based on widely accepted beliefs about declarations, it is important to compare hearsay statements across different contexts. For example, it is necessary to examine whether a jury is likely to reject the validity of an inadmissible hearsay statement given in an ordinary context but accept its reliability if given under specific circumstances for which the FRE provides exceptions to the rule against hearsay, such as if the declarant's statements were made under the perception of imminent death. And while this last example is not an ethical one to manipulate experimentally, it is an open question whether or not individuals who believe they are on the brink of death would actually be honest under such exceptional circumstances.

Directions for Future Research

The previous sections have discussed the current state of psychological research in regard to the FRE and the ways in which the rules' assumptions do or do not correspond to the research. This section will examine several areas of evidence law that still require further research and will include some suggestions for how this research can and should be conducted.

Rule 606(b)

Federal Rule of Evidence 606(b)(1) states that "during an inquiry into the validity of a verdict or indictment, a juror may not testify about any statement made or incident that occurred during the jury's deliberations…The court may not receive a juror's affidavit or evidence of a juror's statement on these matters." Jury deliberation is one of the only aspects of the trial process that is kept hidden from the public. Court opinions for bench trials are written and published with an analysis detailing the judge's rationale for why the court reached its decision. The Seventh Amendment provides the right to a jury trial in most civil cases, and under the Sixth Amendment, all criminal defendants (except those charged with petty offenses) have a constitutional right to a public trial by an impartial jury. Secret trials are prohibited as a safeguard against any attempt to use the court as an instrument of persecution (*R.L.R. v. State* 1971). The jury room is an exception to this public scrutiny. It has been referred to as a black box for which the inputs (evidence and arguments) are highly regulated and the output (verdict) is made public, but the inner workings are carefully guarded and insulated from review (MacCoun 1987; *United States v. Benally* 2008). The purpose of this rule is to promote open discussion among jurors without fear of postverdict public scrutiny, retaliation, or harassment (Racist Juror Misconduct During Deliberations 1988; *Shillcutt v. Gagnon* 1987). However, this interest in deliberation privacy creates a barrier to a just and fair result for some parties when juror misconduct comes to light after the verdict has been determined. For example, in *Shillcutt v. Gagnon* (1987), after the defendant was found guilty, one juror reported that another juror had stated during the deliberation, "Let's be logical. He's black and he sees a seventeen year old white girl—I know the type." While this clearly appears to be a violation of the defendant's Sixth Amendment right to an impartial and unbiased jury, Rule 606(b) dictated that the juror who initially reported the misconduct was barred from testifying about the existence of the racist statement. Without any admissible evidence of misconduct, the defendant could not be granted a new trial.

However, comments made by jurors could impact the verdict decision making of other jurors. Under the theory of aversive racism, racist remarks by one juror may cause the other jurors to be less likely to find a Black defendant guilty. Aversive racism is a hypothesis that due to contemporary norms, most Whites believe that equality is important, but they may still harbor latent negative attitudes toward Blacks and other minorities (Dovidio and Gaertner 1991). The theory suggests that when race is salient, egalitarian norms and beliefs are triggered and Whites will attempt to suppress any expressions of bias, but when race is not salient, latent prejudice may arise. Kleynhans and Bornstein (2015) tested this theory by manipulating the presence of a racist juror remark. All participants read the same trial fact pattern with a Black defendant accused of aggravated assault. Participants then read a jury deliberation transcript with no reference to the defendant's race, a racially prejudiced comment from one juror, or a neutral comment regarding the defendant's race. The researchers found that those in the negative race condition were the least likely to find the defendant guilty, followed by the neutral race

condition; participants in the control condition, with no reference to race, were most likely to find the defendant guilty. This means that jurors are influenced by the comment—albeit in the defendant's favor—and not relying only on the evidence in the case. Rule 606(b) bars any testimony that could reveal this bias.

A recent Supreme Court ruling has held that 606(b) is not unconstitutional because the right to an impartial jury is still protected by other adequate safeguards (*Warger v. Shauers* 2014). The court stressed that "parties may bring to the court's attention evidence of bias *before* the verdict is rendered" (*Warger v. Shauers* 2014, p. 523; emphasis added). However, this is problematic, because clearly if defendants are unaware of the comment prior to the rendering of the verdict, they are unable to bring it to the attention of the court. The duty then falls solely to other jury members who are privy to the jury deliberation misconduct. Jurors might not feel comfortable reporting the misconduct of another juror until the trial is over and they are physically distant from the biased juror. Therefore, it is crucial to examine whether jurors are actually likely to bring the comment to the attention of the court in the permissible time, prior to the return of the verdict. If the research shows that they are unlikely to do so, then 606(b) may be barring the only realistic avenue of recourse for juror misconduct in the deliberation room.

Rule 603 Oath or Affirmation to Testify Truthfully

The entertainment industry has made the legal oath-taking process and raising one hand while placing the other on the Bible and swearing to tell "the truth, the whole truth and nothing but the truth, so help me God" very familiar to the average person. Rule 603 requires that prior to testifying, "a witness must give an oath or affirmation to testify truthfully. It must be in a form designed to impress that duty on the witness's conscience." The Advisory Committee Notes (1972) comment that the rule is specially designed to allow for flexibility for all witnesses, be they "religious adults, atheists, conscientious objectors, mental defectives [or] children." There is no special verbal formula required as long as the affirmation signals a solemn undertaking to tell the truth, and the oath does not require the Bible or any other exalted text.

The process and its familiarity lead to a number of interesting research questions. First, what is the effect of oath taking? How does it influence the witness, and how does it affect the perception of the trier of fact? In addition, due to the familiarity of the statement above, is there a difference in the perception and impact of the statement if the witness varies the statement and in particular, if the witness removes all reference to religion? This latter question is particularly relevant and interesting in relation to the research discussed above involving ingroup religious bias.

Some researchers have examined truth-telling in the context of embodied cognition, which suggests that people use concrete bodily sensations, the physical environment, and social context to make sense of abstract concepts, such as honesty (Barsalou 2009). Through metaphor, these concepts can become so linked with certain bodily movements that as soon as any component of the pattern is triggered,

the association is activated due to previous frequent connections (see Fetterman and Robinson 2013; Schnall et al. 2008; Williams and Bargh 2008). Parzuchowski et al. (2014) examined the metaphoric association of the heart with the concept of honesty: "from the heart," "cross my heart." They hypothesized that the motor movements of placing a hand over the heart would increase the accessibility of thought content associated with abstract honest behaviors, such that participants would be more likely to perceive honesty or be more honest even without explicitly thinking about the semantic meaning. They showed their participants a photo of a man with either his hands by his sides or one hand placed over his heart. Participants then read a number of statements that were ostensibly made by this man, some of which were highly improbable, such as "I have never argued with members of my family." Participants who saw the picture with his hand over his heart were more likely to perceive the man as honest. Parzuchowski et al. (2014) also found that participants who engaged in the behavior of placing their own hand over the heart were more likely to be honest than those who kept their hands on their hips. Thus, it is possible that the simple behavior of swearing an oath, especially when done with the prototypical movements, could lead to more honest testimony from the witness and a perception of greater witness credibility from the trier of fact.

The simple presence of religious words may also have an effect on prosocial behavior. Randolph-Seng and Nielsen (2007) used a conceptual priming technique in which participants were exposed to words that were either neutral or related to religion or sports. Participants were then asked to complete a circle test designed to measure cheating. The test required writing specific numbers inside a small circle while alone in a room with their eyes closed. Participants were induced to cheat by being provided with unrealistic performance expectations and a promise of additional extra credit for high performance. The researchers also measured religious orientation. They found that there were significant differences in cheating rates between conditions, with cheating rates of 44 % in the neutral, 50 % in the sports, and 0 % in the religious condition. Further, there were no significant differences based on the participant's religious orientation, suggesting that a stereotypical representation of religion triggers an automatic influence on prosocial behavior, regardless of individual religious belief. However, a recent meta-analysis found that while religious priming does have an affect across a variety of outcome measures, it does not reliably affect nonreligious participants (Shariff et al. 2016). Therefore, one question to be asked is whether or not a religious oath would affect a witness's honesty even if the witness is not religious. Moreover, does a religious oath have a greater impact than a simple promise to tell the truth? This is a particularly interesting question as witnesses are not required to refer to religion in their oaths.

The above research provides the hypotheses that witnesses would both be and be perceived as more honest when they give an oath prior to testifying. As all witnesses are required to swear some form of oath, it is less important to examine whether the oath itself leads to truth-telling, but more so to compare the impact of different forms of affirmations, religious or secular, on both the truth-telling of the witness and the perception of the juror. This could also presumably be influenced by the religion of the perceiver, as previously discussed.

Rule 505 and 506 Privileges [Not Enacted]

When the FRE were first proposed, they contained a number of rules that were ultimately rejected by Congress. These rejected rules or variations on them were nevertheless adopted in some jurisdictions. Many of these rejections involved rules pertaining to privileges. One example is Rule 506, which held communications to clergy as confidential and allowed both the speaker and the clergy to claim the privilege and refuse to disclose any part of the confidential communication. Instead, Congress enacted Rule 501, which stated generally that the common law governs claims of privilege. Therefore, while not officially binding, the rejected rules do provide a general guide to many judges, particularly as they are based on common law privileges. Thus, though not formally part of the FRE, the rejected rules are still applicable in many states, often through state statutes, and warrant discussion in this chapter.

Another rejected rule is Rule 505, which states that "an accused in a criminal proceeding has a privilege to prevent his spouse from testifying against him." Rule 505(b) also allows that the spouse may claim the privilege on his behalf. Conversely, in *Trammel v. United States* (1980), which was decided after the enactment of the FRE, the Supreme Court held that only the testifying spouse holds the privilege not to testify, and therefore the accused cannot prevent his or her spouse from testifying. Not all jurisdictions, however, have accepted *Trammel*, and in those that have not, the accused may still prevent the spouse from testifying. This includes some jurisdictions in which the FRE or the common law rules relating to privilege do not apply. While varying by jurisdiction, there are some exceptions to the testimonial privilege, such as crimes that were against the spouse, the spouse's property, or the children of either spouse.

A significant research question that has not been examined is the impact of invoking the privilege on the defendant's case. Although research into this specific question appears to be limited, the topic is similar to research into criminal defendants who invoke their Fifth Amendment right to remain silent, where "no person…shall be compelled in any criminal case to be a witness against himself." Legally, pleading the Fifth should not lead to an inference of either guilt or innocence, as both the guilty and the innocent alike have the right to plead (Hook 1957). However, research suggests that defendants who invoke their right to remain silent, either on the stand or by declining to take the stand, are judged as more guilty than those who do take the stand and testify (Shaffer and Case 1982).

Hendrick and Shaffer (1975) conducted the first study to determine whether pleading the Fifth leads to a stronger inference of guilt. They manipulated whether the defendant denied his guilt or pled the Fifth, and either denied to implicate another person or chose to remain silent about whether that person was guilty. The latter manipulation was to compare the effects of withholding information about another person versus withholding information about the self. The results indicated that those who pled the Fifth for the self and for the other were seen as the least moral as compared to those who did not plead the Fifth. Participants were also more

willing to indict those who pled the Fifth than those who denied their guilt. Participants were then asked about the likelihood that someone else was guilty, but no differences were found among conditions when the defendant either remained silent or denied the other's guilt. Therefore, overall, defendants were worse off when they chose to remain silent, but there was no effect on the perceived guilt of the other whether the defendant denied his guilt or remained silent.

If this study is extended to situations in which a witness invokes the spousal testimonial privilege in a case against his or her spouse, it can be presumed that the spouse would be perceived as less credible, but this provides little information about whether this increases perceptions of the defendant's guilt. There is some evidence that guilty verdicts and judgments are predicted by impressions of the defendant and of his or her character witnesses, and it is possible that this could also include the defendant's spouse, whether or not the spouse chooses to testify (Hunt and Budesheim 2004). These presumptions can be tested by manipulating whether the defendant's spouse testifies or invokes the spousal privilege and then evaluating verdict outcomes, witness and defendant credibility, and defendant guilt perceptions.

Conclusions

This chapter has examined the assumptions and implications of the FRE. While psychological research indicates that some of the rules do reflect reliable research, many of the rules do not seem to correspond with empirical evidence and can leave legal actors, especially criminal defendants, vulnerable. Researchers need to further evaluate the direct impact that use of the FRE may have on the actions of the parties and the perceptions of the triers of fact. To date, psycholegal researchers have concentrated on a small number of evidentiary rules (e.g., those related to relevance, prejudice, and expert testimony) and have paid little attention to many other evidence rules that make psychological assumptions. These literature gaps need empirical attention to ensure that all parties receive the most just procedures and substantive outcomes.

References

Aarts, H., & Dijksterhuis, A. (2000). Habits as knowledge structures: Automaticity in goal-directed behavior. *Journal of Personality and Social Psychology, 78*, 53–63.

Aarts, H., Verplanken, B., & Van Knippenberg, A. (1998). Predicting behavior from actions in the past: Repeated decision making or a matter of habit. *Journal of Applied Social Psychology, 28* (15), 1355–1374.

Alvarez, M. J., Miller, M. K., & Bornstein, B. H. (2016). "It will be your duty…:" The psychology of criminal jury instructions. In M. K. Miller & B. H. Bornstein (Eds.), *Advances in psychology and law* (Vol. 1, pp. 119–158). New York: Springer.

Arima, Y. (2013). Effect of word-list consistency on the correlation between group memory and group polarization. *Psychological Reports, 112*(2), 375–389.
Barsalou, L. W. (2009). Situating concepts. In P. Robbins & M. Aydede (Eds.), *Cambridge handbook of situated cognition* (pp. 236–263). New York: Cambridge University Press.
Bartlett, F. C. (1932). *Remembering: An experimental and social study*. Cambridge: Cambridge University.
Batson, C. D., Schoenrade, P., & Ventis, W. L. (1993). *Religion and the individual: A social-psychological perspective*. Oxford: Oxford University Press.
Bell, B. E., & Loftus, E. F. (1985). Vivid persuasion in the courtroom. *Journal of Personality Assessment, 49*, 659–664.
Best, A. (2009). *Examples & explanations: Evidence*. New York, NY: Aspen Publishers.
Black, R. (2011, June 21). Casey: The 403 strategy. [Web log post]. Retrieved from http://www.royblack.com/blog/casey-the-403-strategy/
Blau, T. H. (1998). *The psychologist as expert witness.* New York: John Wiley & Sons Inc.
Bornstein, B. H., & Greene, E. (2017). *The jury under fire: Myth, controversy, and reform*. New York: Oxford University Press.
Bornstein, B. H., & Miller, M. K. (2009). *God in the courtroom: Religion's role at trial.* Oxford: Oxford University Press.
Bornstein, B. H., & Nemeth, R. J. (1999). Jurors' perception of violence: A framework for inquiry. *Aggression and Violent Behavior, 4*, 77–92.
Bornstein, B. H., Rung, L. M., & Miller, M. K. (2002). The effects of defendant remorse on mock juror decisions in a malpractice case. *Behavioral Sciences & the Law, 20*, 393–409.
Bright, D. A., & Goodman-Delahunty, J. (2004). The influence of gruesome verbal evidence on mock juror verdicts. *Psychiatry, Psychology and Law, 11*, 154–166.
Bright, D. A., & Goodman-Delahunty, J. (2006). Gruesome evidence and emotion: Anger, blame, and jury decision-making. *Law and Human Behavior, 30*, 183–202.
Cleary, E. W. (Ed.) (1984). *McCormick's handbook on the law of evidence.* St. Paul: West.
Diamond, S. S., & Vidmar, N. (2001). Jury room ruminations on forbidden topics. *Virginia Law Review*, *87*, 1857–1915.
Dickter, C. L., & Bartholow, B. D. (2007). Racial ingroup and outgroup attention biases revealed by event-related brain potentials. *Social Cognitive and Affective Neuroscience, 2*, 189–198.
Douglas, K. S., Lyon, D. R., & Ogloff, J. R. (1997). The impact of graphic photographic evidence on mock jurors' decisions in a murder trial: probative or prejudicial? *Law and Human Behavior, 21*, 485–501.
Dovidio, J. F., & Gaertner, S. L. (1991). Changes in the expression and assessment of racial prejudice. In H. J. Knopke, R. J. Norrel, & R. W. Rogers (Eds.), *Opening doors: Perspectives on race relations in contemporary America* (pp. 119–148). Tuscaloosa, AL: The University of Alabama Press.
Dror, I. E., Kassin, S. M., & Kukucka, J. (2013). New application of psychology to law: Improving forensic evidence and expert witness contributions. *Journal of Applied Research in Memory and Cognition, 2*, 78–81.
Dumas, B. K. (2007). US pattern jury instructions: Problems and proposals. *International Journal of Speech Language and the Law, 7*, 49–71.
Falk, A., & Zehnder, C. (2007). *Discrimination and in-group favoritism in a citywide trust experiment*. IZA Discussion Paper 2765.
Faust, D., & Ziskin, J. (1988). The expert witness in psychology and psychiatry. *Science, 241*, 31–35.
Fetterman, A. K., & Robinson, M. D. (2013). Do you use your head or follow your heart? Self-location predicts personality, emotion, decision making, and performance. *Journal of Personality and Social Psychology, 105*, 316–334.
Gold, V. J. (1984). Limiting judicial discretion to exclude prejudicial evidence. *UC Davis Law Review, 18*, 59–96.
Goldfried, J., & Miner, M. (2002). Quest religion and the problem of limited compassion. *Journal for the Scientific Study of Religion*, *41*, 685–695.

Golding, J. M., Sanchez, R. P., & Sego, S. A. (1997). The believability of hearsay testimony in a child sexual assault trial. *Law and Human Behavior, 21*, 299–325.

Gonzales, M. H., Manning, D. J., & Haugen, J. A. (1992). Explaining our sins: Factors influencing offender accounts and anticipated victim responses. *Journal of Personality and Social Psychology, 62*, 958–971.

Gonzales, M. H., Peterson, J., Manning, D. J., & Wetter, D. (1990). Pardon my gaffe: Effects of sex, status, and consequence severity on accounts. *Journal of Personality and Social Psychology, 58*, 610–621.

Greene, E., & Dodge, M. (1995). The influence of prior record evidence on juror decision making. *Law and Human Behavior, 19*, 67–78.

Greene, E., Hayman, K., & Motyl, M. (2008). " Shouldn't we consider…?": Jury discussions of forbidden topics and effects on damage awards. *Psychology, Public Policy, and Law, 14*, 194–222.

Hans, V. P., & Doob, A. N. (1976). Section 12 of the Canada Evidence Act and the Deliberations of Simulated Juries. *Criminal Law Quarterly*, *18*, 235–253.

Hart, A. J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., & Rauch, S. L. (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *NeuroReport, 11*, 2351–2354.

Hendrick, C., & Shaffer, D. R. (1975). Effect of pleading the Fifth Amendment on perceptions of guilt and morality. *Bulletin of the Psychonomic Society, 6*, 449–452.

Hook, S. (1957). *Common sense and the fifth amendment*. New York: Criterion Books.

Hunt, J. S., & Budesheim, T. L. (2004). How jurors use and misuse character evidence. *Journal of Applied Psychology, 89*, 347–361.

Iyengar, S., & Westwood, S. J. (2014). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science, 59*, 690–707.

Jackson, L. M., & Hunsberger, B. (1999). An intergroup perspective on religion and prejudice. *Journal for the Scientific Study of Religion, 38*, 509–523.

James, W. (1890). *Habit*. New York: Henry Holt and Company.

Johnson, S. D. (1985). Religion as a defense in a mock-jury trial. *Journal of Social Psychology, 125*, 213–220.

Kassin, S. M., & Wrightsman, L. S. (Eds.). (1985). *The psychology of evidence and trial procedure* (pp. 67–94). Beverly Hills: Sage.

Kleynhans, A., & Bornstein, B. H. (2015). *The white man's court: The impact of rule 606(b) and racism in the jury room.* Paper presented at the annual meeting of the American Psychology Law Society in San Diego, CA.

Konecni, V. J. (1972). Some effects of guilt on compliance: A field replication. *Journal of Personality and Social Psychology, 23*, 30–32.

Kovera, M. B., Park, R. C., & Penrod, S. D. (1991). Jurors' perceptions of eyewitness and hearsay evidence. *Minnesota Law Review, 76*, 703–722.

Lieberman, J. D. (2009). The psychology of the jury instruction process. *Jury psychology: Social aspects of trial processes: Psychology in the courtroom, 1*, 129–155.

MacCoun, R. J. (1987). *Getting inside the black box: Toward a better understanding of civil jury behavior*. Santa Monica, CA: RAND.

Mauet, T. A., & Wolfson, W. D. (2011). *Trial evidence* (5th ed.). New York, NY: Aspen Publishers.

Miller, G. R., & Burgoon, J. K. (1982). Factors affecting assessments of witness credibility. *The Psychology of the Courtroom*, 169–196.

Mischel, W. (1969). Continuity and change in personality. *American Psychologist, 24*, 1012–1018.

Mueller, C. B., & Kirkpatrick, L. C. (2013). *Federal rules of evidence*. New York, NY: Wolters Kluwer.

Myers, D. G., & Bishop, G. D. (1971). Enhancement of dominant attitudes in group discussion. *Journal of Personality and Social Psychology, 20*, 386–391.

Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin, 83*, 602–627.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Eaglewood Cliffs, NJ: Prentice-Hall.

Ohbuchi, K. I., Kameda, M., & Agarie, N. (1989). Apology as aggression control: Its role in mediating appraisal of and response to harm. *Journal of Personality and Social Psychology, 56*, 219–227.

Otten, S., & Wentura, D. (1999). About the impact of automaticity in the Minimal Group Paradigm: Evidence from affective priming tasks. *European Journal of Social Psychology, 29*, 1049–1071.

Parzuchowski, M., Szymkow, A., Baryla, W., & Wojciszke, B. (2014). From the heart: Hand over heart as an embodiment of honest. *Cognitive Processing,* 1–8.

Perdue, C. W., Dovidio, J. F., Gurtman, M. B., & Tyler, R. B. (1990). Us and them: Social categorization and the process of intergroup bias. *Journal of Personality and Social Psychology, 59*, 475–486.

Peterson, D. R. (1968). *The clinical study of social behavior*. New York: Appleton-Century-Crofts.

Pickel, K. L. (1995). Inducing jurors to disregard inadmissible evidence: A legal explanation does not help. *Law and Human Behavior, 19*, 407–424.

Racist juror misconduct during deliberations. (1988). *Harvard Law Review*, *101*, 1595–1602.

Randolph-Seng, B., & Nielsen, M. E. (2007). Honesty: One effect of primed religious representations. *The International Journal for the Psychology of Religion, 17*, 303–315.

Regan, J. W. (1971). Guilt, perceived injustice, and altruistic behavior. *Journal of Personality and Social Psychology, 18*, 124–132.

R.L.R. v. State, 487 P.2d 27 (Alaska 1971).

Robbennolt, J. K. (2013). The effects of negotiated and delegated apologies in settlement negotiation. *Law and Human Behavior, 37*, 128–135.

Saroglou, V., Pichon, I., Trompette, L., Verschueren, M., & Dernelle, R. (2005). Prosocial behavior and religion: New evidence based on projective measures and peer ratings. *Journal for the Scientific Study of Religion, 44*, 323–348.

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin, 34*, 1096–1109.

Schuller, R. A. (1995). Expert evidence and hearsay: The influence of "secondhand" information on jurors' decisions. *Law and Human Behavior, 19*, 345–362.

Schwartz, S. H., & Struch, N. (1989). Values, stereotypes, and intergroup antagonism. In *Stereotyping and prejudice* (pp. 151–167). New York: Springer.

Shaffer, D. R., & Case, T. (1982). On the decision to testify in one"s own behalf: Effects of withheld evidence, defendant's sexual preferences, and juror dogmatism on juridic decisions. *Journal of Personality and Social Psychology, 42*, 335–346.

Shariff, A. F., Willard, A. K., Andersen, T., & Norenzayan, A. (2016). Religious priming: A meta-analysis with a focus on prosociality. *Personality and Social Psychology Review, 20*, 27–48.

Shayo, M., & Zussman, A. (2011). Judicial ingroup bias in the shadow of terrorism. *Quarterly Journal of Economics, 126*, 1447–1484.

Shillcutt v. Gagnon, 827 F.2d 1155, (7th Cir. 1987).

State of Florida v. Anthony, WL 4102581 (2011).

Steblay, N., Hosch, H. M., Culhane, S. E., & McWethy, A. (2006). The impact on juror verdicts of judicial instruction to disregard inadmissible evidence: A meta-analysis. *Law and Human Behavior, 30*, 469–492.

Tanford, S., & Cox, M. (1988). The effects of impeachment evidence and limiting instructions on individual and group decision making. *Law and Human Behavior, 12*, 477–497.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology, 1*, 149–178.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Montery, CA: Brooks/Cole.

Teitelbaum, L. E., Sutton-Barbere, G., & Johnson, P. (1983). Evaluating the prejudicial effect of evidence: Can judges identify the impact of improper evidence on juries. *Wisconsin Law Review*, 1147–1201.

Thompson, W. C., Fong, G. T., & Rosenhan, D. L. (1981). Inadmissible evidence and juror verdicts. *Journal of Personality and Social Psychology, 40*, 453–463.

Thompson, W. C., & Pathak, M. K. (1999). Empirical study of hearsay rules: Bridging the gap between psychology and law. *Psychology, Public Policy, and Law, 5*, 456–472.

Thornton, B., Kirchner, G., & Jacobs, J. (1991). Influence of a photograph on a charitable appeal: A picture may be worth a thousand words when it has to speak for itself. *Journal of Applied Social Psychology, 21*, 433–445.

Trammel v. United States, 445 U.S. 40, 100 S. Ct. 906 (1980).

United States v. Benally, 546 F.3d 1230, 1236–38 (10th Cir. 2008).

United States v. Sampol. 636 F.2d 621 (D.C. Cir. 1980).

Vernon, P. E. (1964). *Personality assessment: A critical survey*. New York: Wiley.

Warger v. Shauers, 134 S. Ct. 1491, 188 L. Ed. 2d 374 (2014).

Weissenberger, G. (1992). Supreme Court and the interpretation of the federal rules of evidence. *The Ohio State Law Journal, 53*, 1307–1339.

Williams, L. E., & Bargh, J. A. (2008). Experiencing physical warmth promotes interpersonal warmth. *Science, 322*, 606–607.

Wissler, R. L., & Saks, M. J. (1985). On the inefficacy of limiting instructions: When jurors use prior conviction evidence to decide on guilt. *Law and Human Behavior, 9*, 37–48.

Wolf, S., & Montgomery, D. A. (1977). Effects of inadmissible evidence and level of judicial admonishment to disregard on the judgments of mock jurors. *Journal of Applied Social Psychology, 7*, 205–219.

Wyer, R. S., Budesheim, T. L., Lambert, A. J., & Swan, S. (1994). Person memory and judgment: Pragmatic influences on impressions formed in a social context. *Journal of Personality and Social Psychology, 66*, 254–267.

An Examination of Whether Psychopathy Checklist-Revised (PCL-R) Evidence Satisfies the Relevance/Prejudice Admissibility Standard

David DeMatteo, Heath Hodges and Jaymes Fairfax-Columbo

Introduction

Psychopathy is a clinical construct defined by a cluster of interpersonal, affective, and lifestyle characteristics (Hare 1991, 1999; Hare and Neumann 2008), including arrogance, callousness, superficiality, manipulativeness, grandiosity, inability to form strong emotional bonds with others, a relative absence of guilt, irresponsibility, ignoring/violating social norms, and impulsivity (Cleckley 1976; Hare 1970; McCord and McCord 1964). These characteristics limit the ability to inhibit antisocial behaviors (Miller and Eisenberg 1988). As such, it is unsurprising that psychopathy is strongly associated with aggression and violence (Hart and Hare 1997), criminality and criminal justice system involvement (Hare 1999), and criminal recidivism (Hemphill et al. 1998; Salekin et al. 1996).

Though the base rate of psychopathy in the general population is low, perhaps about 1 % (Hare 1996), individuals with psychopathic characteristics are disproportionately represented in the criminal justice system, with approximately 16 % of male criminals meeting criteria for psychopathy (Kiehl and Hoffman 2011). Further, though it has long been recognized that not all individuals with psychopathic characteristics will come into contact with the criminal justice system (Babiak 1995; Cleckley 1976; Schneider 1958), and though "successful psychopaths" (Hall and Benning 2006) have been of increasing research interest in recent years, approximately 93 % of males meeting criteria for psychopathy in the United States are currently incarcerated or facing some type of criminal justice system supervision (Kiehl and Hoffman 2011).

D. DeMatteo (✉) · J. Fairfax-Columbo
Department of Psychology, Thomas R. Kline School of Law, Drexel University, 3141 Chestnut Street, Stratton Suite 119, Philadelphia, PA 19104, USA
e-mail: david.dematteo@drexel.edu

H. Hodges
Department of Psychology, University of Nebraska-Lincoln, Lincoln, NE, USA

B.H. Bornstein and M.K. Miller (eds.), *Advances in Psychology and Law*,
Advances in Psychology and Law 2, DOI 10.1007/978-3-319-43083-6_7

Perhaps the most widely utilized measure of psychopathy in criminal justice and forensic settings is the Psychopathy Checklist-Revised (PCL-R; Hare 1991, 2003). The PCL-R is a 20-item measure designed to assess psychopathic characteristics in correctional and forensic psychiatric populations. PCL-R administration consists of a semi-structured interview and review of collateral records (e.g., institutional files, criminal history, prior assessments). PCL-R Total scores range from 0 to 40, with scores of 30 and above suggesting that an individual is a psychopath (Hare 2003). Although the legal system is primarily interested in psychopathy as a categorical construct because of the need to make discrete placement decisions, there is no natural diagnostic cutoff for categorizing individuals as "psychopaths," and taxometric studies indicate that PCL-R scores represent an underlying dimensional construct rather than a latent taxon (e.g., Edens et al. 2006; Edens and Petrila 2006; Guay et al. 2007; Walters et al. 2007).

Early factor analyses of the PCL-R's precursor, the Psychopathy Checklist (Hare 1980), consistently resulted in a two-factor solution, with Factor 1 primarily reflecting the interpersonal and affective characteristics of psychopathy and Factor 2 primarily reflecting antisocial and socially deviant characteristics often seen among criminal offenders (Harpur et al. 1988, 1989). Confirmatory factor analyses with the PCL-R call into question the validity of the two-factor solution, with some researchers proposing three-factor (e.g., Cooke and Michie 2001; Hall et al. 2004) and four-factor models (e.g., Forth et al. 2003; Neumann et al. 2007; Vitacco et al. 2005).

The PCL-R has been found to be psychometrically sound (Cooke and Michie 1997; Hare 1991; Harpur et al. 1989), with its ability to identify individuals with psychopathic traits generalizable cross-culturally (Hare et al. 2000).[1,2] Research regarding the PCL-R's predictive ability has been extensive and a strong body of research suggests that higher PCL-R scores are associated with general and violent recidivism (DeMatteo et al. 2010; Singh et al. 2011; Yang et al. 2010), sexual recidivism (Hanson and Morton-Bourgon 2009; Hawes et al. 2013), institutional misconduct (Guy et al. 2005), community violence (Skeem and Mulvey 2001), treatment amenability in civil psychiatric (Skeem et al. 2002), and correctional settings (Ogloff et al. 1990; Rice 1997; Seto and Barbaree 1999). Of note, however, a 2004 systematic review of 24 studies found that the majority of studies of treatment amenability are of questionable design and that the existing research does

[1]The PCL-R has shown to be reliable and modestly valid in assessing psychopathy in females (Vitale and Newman 2001), but some have expressed concern about the PCL-R's ability to assess the construct of psychopathy in female offenders (Vitale et al. 2002).

[2]Several derivative measures of the PCL-R exist. The Psychopathy Checklist: Youth Version (Forth et al. 2003) is a 20-item measure designed to assess psychopathic traits in male and female adolescent offenders between the ages of 12 and 18. The Psychopathy Checklist: Screening Version (Hart et al. 1995) is a 12-item brief screening measure for the presence of psychopathy. Several self-report measures have also been created, including the P-SCAN (Hare and Hervé 1999), Self-Report Psychopathy Scale-Fourth Edition (Paulhus et al. in press), and Antisocial Process Screening Device (Frick and Hare 2001).

not provide adequate "evidence to conclude that high-scoring psychopaths have a negative response to treatment" (D'Silva et al. 2004, p. 175).

Due to these associations, PCL-R evidence is often presented via expert testimony to assist courts in making factual or legal determinations in several legal contexts, including sexually violent predator (SVP) commitment, juvenile transfer decisions, capital sentencing, general sentencing, mental state at the time of the offense, and determination of future dangerousness. Use of the PCL-R in United States courts has increased substantially in recent years (DeMatteo et al. 2014b). Two early reviews by DeMatteo and Edens (2006) and Walsh and Walsh (2006) found relatively little usage of the PCL-R from 1991 to 2004, noting usage in 87 and 76 cases, respectively. Use of the PCL-R increased dramatically during the 6-year period from 2005 to 2011, with DeMatteo et al. (2014a, b) finding that the PCL-R was used in 348 cases, which is a 300+ % increase in a time period half as long as the preceding one. Importantly, the results of these case law reviews likely provide underestimates of PCL-R use in court because legal databases typically only contain opinions that have reached the appellate stage of litigation.

Although it is encouraging that the legal system is increasingly willing to utilize science in helping juries and judges in their decision-making, questions of utility and potential misuse abound regarding the PCL-R's ability to assist in making factual and legal determinations in legal contexts (DeMatteo and Edens 2006; DeMatteo et al. 2014a, b; Edens 2001, 2006; Edens et al. 2001; Walsh and Walsh 2006). For example, DeMatteo and Edens (2006) reviewed the case of *United States v. Barnette* (2000), a capital case in which the defense attempted to present mitigating evidence, including the testimony of 3 experts and 12 character witnesses, that the defendant would adjust well to a structured prison environment and that he would not represent a future danger in prison. In rebuttal to this evidence, the prosecution presented evidence that Mr. Barnette was likely to represent a future danger if incarcerated for life instead of sentenced to death, based partly on his PCL-R score of 35, which placed him within the "psychopath" range. Mr. Barnette was sentenced to death, and although it is unclear if the testimony that Mr. Barnette was a psychopath was a primary factor in the jury's sentencing decision, the evidence was deemed admissible and presented to the jury.

The problem was that the prosecution's expert testimony regarding Mr. Barnette's future dangerousness was misleading. The prosecution expert testified that criminals with psychopathic traits are three times more likely to engage in future criminal behavior compared to criminals without those traits, and that although only approximately 20 % of incarcerated persons meet criteria for psychopathy, these individuals account for over half of violent crimes that are committed. Although these results might have applied to offenders with psychopathic traits at that time, scarce evidence existed at the time to suggest that these likelihoods applied to individuals with psychopathic characteristics in capital contexts. As such, the usefulness of this PCL-R evidence in rebutting the defense's claim that Mr. Barnette would not be a future danger was questionable, yet it was still

permitted to be presented to the jury. Given that the usefulness of Mr. Barnette's status as being highly psychopathic was questionable at best in determining his level of future dangerousness while incarcerated, why was this evidence deemed admissible and presented to the jury? How do United States courts determine which evidence factfinders are allowed to consider and which evidence they are not? The remainder of this chapter addresses the important but seldom-asked question of whether PCL-R-based future dangerousness testimony clears the relevance/prejudice hurdle and should be admissible for the legal contexts in which it is often presented. Although the PCL-R is used to aid decision-making in a number of legal contexts (DeMatteo and Edens 2006; DeMatteo et al. 2014a, b; Walsh and Walsh 2006), this chapter limits itself to consideration of this question only for the contexts in which the probative value and prejudicial impact of the PCL-R and psychopathy designation have been most rigorously researched in terms of predictive ability for the legal question at hand and the generation of labeling effects: capital sentencing, juvenile transfer, and SVP commitment. For each of these three contexts, the social science literature pertaining to the probative value (predictive ability) for the legal outcome of interest will be summarized, the social science literature relating to psychopathy labeling effects (prejudicial impact) will be summarized, and Federal Rules of Evidence will be applied to this research to examine if PCL-R evidence should be admissible for the purpose for which it is proffered.

Admissibility of Evidence

In federal courts, admissibility of evidence is governed by the Federal Rules of Evidence (FRE 1975). According to Federal Rules of Evidence for United States Courts and Magistrates Rule 1101, the FRE apply to all civil and criminal court cases and proceedings and to contempt proceedings, but they do not apply to judges' decisions regarding the preliminary admissibility of evidence, to grand jury hearings, and assorted other proceedings, such as sentencing, extradition, the issuing of warrants, preliminary examination in criminal cases, sentencing, granting/revoking probation or supervised release, and bail consideration. The FRE do not apply directly to the states, yet nearly every state has adopted the FRE either in whole, in part, or with minor modifications (Weinstein and Berger 2015). Due to lack of uniform adoption of FRE, evidentiary standards vary by jurisdiction. For example, though the FRE and FRE-derived evidentiary codes do not apply at capital sentencing in the federal system and in many states, some jurisdictions, such as Louisiana, apply evidentiary rules at sentencing (Louisiana Code of Evidence 2011; Louisiana Code of Criminal Procedure 1976).

As another example, the extent to which the FRE apply to juvenile transfer hearings also varies by jurisdiction. Although the majority of jurisdictions do not

strictly apply evidentiary and procedural rules at transfer hearings, evidentiary rules are strictly applied in a minority of jurisdictions (see Hill 1996, for a review). Reverse waivers begin in criminal court (because the juvenile is initially charged as an adult), which means the FRE are applicable *prima facie*. However, it is possible that a juvenile transfer hearing falls under the exception of "miscellaneous proceedings" under the applicability rules for the FRE, which is decided by legislatures or courts in each jurisdiction. For example, the Wisconsin Supreme Court held that the state rules of evidence apply in all criminal proceedings, including reverse waiver hearings (*State v. Kleser* 2010). In contrast, the Court of Special Appeals of Maryland held that courts have the discretion not to strictly apply evidentiary rules during reverse waiver hearings (*Gaines v. State* 2011).

Perhaps the biggest barrier to evidence making its way to factfinders and legal decision-makers for consideration is the relevance/prejudice hurdle. To clear this hurdle, the proffered evidence must satisfy the requirements for Federal Rules of Evidence for United States Courts and Magistrates 401, 402, and 403.

- Federal Rules of Evidence for United States Courts and Magistrates 401. Test for Relevant Evidence—Evidence is relevant if (a) it has any tendency to make a fact more or less probable that it would be without the evidence; and (b) the fact is of consequence in determining the action.
- Federal Rules of Evidence for United States Courts and Magistrates 402. General Admissibility of Relevant Evidence—Relevant evidence is admissible unless any of the following provides otherwise: the United States Constitution; a federal statute; these rules; or other rules prescribed by the Supreme Court. Irrelevant evidence is not admissible.
- Federal Rules of Evidence for United States Courts and Magistrates 403: Excluding Relevant Evidence for Prejudice, Confusion, Waste of Time, or Other Reasons—The court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence.

Considering these rules together, for evidence to be admissible, it must be relevant, meaning it is material to the issue at hand and the existence of the evidence provides some sort of predictive value in determining if a fact exists—in other words, the evidence must have *probative value* (Federal Rules of Evidence for United States Courts and Magistrates 401). This term is somewhat ambiguous, but is determined by a judge in light of the "facial definiteness of the evidence, the length of the chain of inferences, and the strength of those inferences" (Imwinkelried 1988, p. 905). If proffered evidence satisfies the probative value requirement, it is generally admissible to the factfinder; if it does not, it is disqualified from consideration. But even if evidence has probative value, it may be disqualified from consideration by the factfinder if it poses a risk of undermining the efficiency of the judicial process or undermining confidence in a factual or legal determination—in other words, if its presentation would have an overly *prejudicial*

impact on the factfinder (Federal Rules of Evidence for United States Courts and Magistrates 403). Such prejudicial impact "threatens to prompt inferential error by the trier of fact" (Imwinkelried 1988, p. 905). In short, for evidence to be admissible, it must have the capacity to assist the factfinder in decision-making and it cannot be overly prejudicial to the party against whom it is proffered.[3]

In applying Federal Rules of Evidence for United States Courts and Magistrates 401 and 403 to the use of the PCL-R, the analysis will vary depending on the legal context because the outcome of interest in each context varies. For juvenile transfer, the outcomes of interest are (1) adult criminality and (2) the youth's responsiveness to treatment because these are the factors most often considered in transfer hearings. Regarding capital sentencing, the outcomes of interest are (1) the prediction of future dangerousness via institutional misconduct and (2) malingering because prosecutors sometimes offer evidence of psychopathy to rebut an insanity defense (Edens et al. 2013). Concerning SVP proceedings, the outcome of interest is recidivism in the community over short- and long-term periods of time.

Concerning the Federal Rules of Evidence for United States Courts and Magistrates 401 analysis of the PCL-R's predictive ability for the outcome of interest in each context, PCL-R evidence will be considered probative if the extant literature clearly and unambiguously supports the PCL-R's ability to predict the outcome. Regarding the Federal Rules of Evidence for United States Courts and Magistrates 403 analysis as to the prejudicial impact of a psychopath label in each context, the label "psychopath" will be considered to have a prejudicial impact if the extant literature clearly and unambiguously supports labeling effects (i.e., biases created by a psychopathy diagnosis or the attribution of psychopathic traits to an individual). The PCL-R will be deemed inadmissible in contexts in which it is determined to have no probative value, or in contexts in which the prejudicial impact of a psychopathy label or the attribution of psychopathic traits substantially outweighs the probative value of the PCL-R for that context.

Juvenile Transfer

Juvenile transfer (or certification) is the process by which juvenile offenders are upgraded from being charged as juveniles in the juvenile justice system to being charged as adults in the criminal justice system. Transfers can occur in one of three ways. The first is statutory transfer, in which a juvenile is automatically charged as an adult if certain criteria are fulfilled (typically regarding offense type and the juvenile's age). The second is judicial transfer, in which juvenile court judges can exercise their discretion as to whether to retain jurisdiction over the juvenile or

[3]Although evidence may be deemed inadmissible because it fails to clear the 401–403 relevance prejudice hurdle, other evidentiary rules may also render evidence inadmissible. For example, because PCL-R evidence is presented via an expert witness, it also must satisfy the requirements of the rules that specifically pertain to expert evidence (FRE 701–706).

transfer the juvenile to adult court. The third is prosecutorial direct file, in which prosecutors have the discretion as to whether charges will be filed in the juvenile or adult system. Most states employ a mixture of these laws, and they allow for reverse waiver (or decertification), which is a process by which juveniles charged as adults can petition to have their case sent to juvenile court (Office of Juvenile Justice and Delinquency Prevention 2010).

For purposes of this chapter, judicial juvenile transfer is of primary concern because judges are interpreting and applying the rules of evidence. In *Kent v. United States* (1966), the United States Supreme Court outlined nine factors juvenile court judges might consider when making transfer decisions: (1) the seriousness of the alleged offense, (2) the level of danger the juvenile presents to the community, (3) the target of the offense (person or property), (4) the sophistication and maturity level of the juvenile, (5) the merit of the complaint, (6) whether the offense was violent and intentional, (7) whether the juvenile's associates in the offense were other juveniles or adults, (8) the juvenile's delinquent history, and (9) the juvenile's prospects for rehabilitation. A statutory review by Heilbrun, Leheny, Thomas, and Huneycutt (1997) revealed that most states incorporate several of these factors and related factors into their transfer statutes, with four factors in particular lending themselves to forensic mental health assessment: offense characteristics (psychological and behavioral aspects), psychopathology, treatment amenability, and risk assessment. Of these factors, risk assessment and treatment amenability can potentially be addressed by the PCL-R.

Probative Value

Though not as extensive as the adult literature, several meta-analyses suggest that adolescent PCL-R Total and Factor scores are positively correlated with future delinquency. In a meta-analysis of 13 studies, Edens and Campbell (2007) found PCL Total scores were predictive of adolescent institutional misconduct, incidents of institutional aggression, and incidents of institutional physical violence, with small to moderate effect sizes for all three outcomes, $r_w = 0.24$, 0.25, and 0.28, respectively. They also found that Factor 2 was more predictive of all three outcomes than Factor 1, although the differences were not statistically significant.

A 2007 meta-analysis by Edens, Campbell, and Weir examined the impact of scores on PCL measures tailored for adolescents [either modifications of the PCL or PCL-R based on recommendations from Forth et al. 1990 or the PCL: YV (Forth et al. 2003)] on general, violent, and sexual recidivism. They found significant small to moderate effects of PCL Total scores for general, $r_w = 0.26$, $p < 0.001$, and violent recidivism, $r_w = 0.23$, $p < 0.001$, but no significant effect for sexual recidivism. Both Factor 1 and Factor 2 showed significant ($p < 0.001$) small to moderate effects for general and violent recidivism, but no significant effect for sexual recidivism, with Factor 2, $r_w = 0.29$, 0.26, being more predictive than Factor 1, $r_w = 0.18$, 0.19, for general and violent recidivism. The moderator analysis for

gender was statistically significant for violent recidivism, $Q_B = 4.37$, $p < 0.05$, but the effect size for female offenders was not statistically significant and much lower than that of the males. Moreover, ethnicity significantly moderated effect sizes for Factor 2 in relation to violent recidivism, $Q_R = 5.57$, $p < 0.05$, with the relationship weakening when a higher proportion of non-White adolescents was included.

Finally, Olver, Stockdale, and Wormith (2009) conducted a meta-analysis of 28 studies utilizing the PCL: YV to predict general, nonviolent, violent, and sexual recidivism. Results indicated significant relationships between PCL: YV scores and general, nonviolent, and violent recidivism. Effect sizes were small to moderate, with general recidivism having an effect size of $r_w = 0.28$, $Q\ r_w = 45.17$, $p < 0.01$, nonviolent recidivism producing an effect size of $r_w = 0.16$, $Q\ r_w = 19.60$, $p < 0.05$, and violent recidivism showing an effect size of $r_w = 0.25$, $Q\ r_w = 34.39$, $p < 0.01$. However, PCL-YV scores produced a weak and nonsignificant effect, $r_w = 0.07$, $Q\ r_w = 1.11$, $p = \text{NS}$, on sexual recidivism. Interestingly, the meta-analysis revealed that PCL: YV effects on both violent and general recidivism were significantly larger in Canadian studies than in studies conducted outside of Canada. Canadian studies produced effect sizes of $r_w = 0.34$ for general recidivism and $r_w = 0.28$ for violent recidivism versus effect sizes of $r_w = 0.20$, $z = 3.60$, $p < 0.001$, and $r_w = 0.18$, $z = 2.53$, $p < 0.05$, respectively, for studies conducted outside of Canada. Taken together, these studies suggest that the PCL-R and PCL: YV are weak to moderate predictors of general, nonviolent, and violent recidivism, but are poor predictors of sexual recidivism.

As of the time of this writing, no meta-analyses exist regarding the relationship between PCL: YV scores and treatment amenability. However, a handful of studies regarding psychopathy and juvenile treatment amenability exist, and they have yielded mixed findings. Although several studies suggest that higher PCL: YV scores may negatively influence treatment amenability (Gretton et al. 2001; O'Neil et al. 2003; Spain et al. 2004), other studies suggest that PCL: YV scores are not associated with a juvenile's treatment experience (Caldwell et al. 2006, 2007).

Spain et al. (2004) examined the influence of PCL: YV scores on institutional misconduct for 85 adolescent male offenders in residential treatment. Results indicated that PCL: YV Total scores were a weak but significant predictor of total, $r = 0.27$, $p < 0.01$, physical, $r = 0.27$, $p < 0.05$, and verbal, $r = 0.24$, $p < 0.05$, misconduct. Although the PCL: YV showed predictive power for institutional misconduct, it was a weaker and more inconsistent predictor than the Antisocial Process Screening Device (Frick and Hare 2001) or Childhood Psychopathy Scale (Lynam 1997).

Employing a different take on treatment amenability, O'Neil et al. (2003) examined the influence of PCL: YV Total and Factor scores on attendance, quality of therapy participation, clean urine tests, consecutively clean urines tests, clinical improvement, and recidivism for 64 male adolescents in a partial hospital treatment program for court-adjudicated youth with evidence of a substance abuse problem. PCL: YV Total scores were negatively correlated with attendance, quality of participation, consecutively clean urine tests, and clinical improvement, $r = -0.42$, -0.50, -0.55, and -0.58, $p < 0.001$, respectively, and with submitting a clean urine

test, $r = -0.31$, $p < 0.01$. PCL: YV Total scores were also positively correlated with recidivism, $r = 0.33$, $p < 0.01$. Regarding factor scores, both Factor 1 and Factor 2 were negatively correlated with attendance, quality of participation, consecutively clean urine tests, and clinical improvement, $r = -0.35$, -0.45, -0.54, and -0.52 versus $r = -0.37$, -0.40, -0.41, and -0.49, $p < 0.001$, respectively. Factor 1 scores were negatively associated with submission of a clean urine test, $r = -0.32$, $p < 0.01$, but were not significantly related to recidivism, while Factor 2 showed the inverse in not being significantly associated with submitting a clean urine test but being positively related to recidivism, $r = 0.34$, $p < 0.01$.

Gretton et al. (2001) explored the impact of PCL: YV scores on escapes/attempted escapes from custody, violation of probation, and general, violent, and sexual reoffending for 220 adolescent males attending an outpatient sex offender treatment program in British Columbia. Instead of employing a regression analysis, Gretton et al. (2001) divided the males into three groups based on exhibition of high, medium, or low levels of psychopathic characteristics. Findings indicated that for escapes/attempted escapes, a greater percentage of high characteristic individuals (44.8 %) attempted escapes/escaped than medium characteristic individuals (10.8 %), who in turn represented a higher percentage than the low characteristic group, 0 %; $\chi^2 = 42.53$, $p < 0.001$. The trend was the same for probation violations, 69 % versus 46.8 % versus 16.3 %, $p < 0.001$, respectively. For reoffending, Gretton et al. (2001) employed point–biserial correlations to examine the impact of PCL: YV scores, finding that total PCL: YV scores were weakly and positively associated with both general and violent recidivisms, $r_{pbi} = 0.25$ and 0.19, $p < 0.01$, but they did not find PCL: YV scores to be a significant predictor of sexual reoffending.

In another study utilizing the MJTC intensive treatment model, Caldwell et al. (2007) explored the effect of PCL: YV scores on institutional misconduct and treatment performance (conforming behaviors to those expected by site staff) for 86 adolescent male offenders split into two groups based on median PCL: YV scores. Results revealed a significant main effect of time, with both the low- and high-psychopathy groups showing significant improvement from baseline to post-treatment, $F(1, 84) = 196.10$, $p < 0.01$. Taken together, the results of these studies present a mixed picture regarding the ability of the PCL:YV to predict treatment amenability in juveniles, though the Caldwell et al. (2006, 2007) studies suggest that, when matched with an appropriately tailored level of treatment, youth high in psychopathic traits can perform favorably in treatment, which suggests that higher levels of psychopathic traits do not necessarily predict poor treatment outcomes.

Prejudicial Impact

A handful of studies have examined the impact of the "psychopath" label and the attribution of psychopathic traits to juvenile offenders on perceptions of juvenile

danger risk and transfer decisions. Those studies examining samples of judges, probation officers, and clinicians are particularly informative because transfer decisions are made by judges whose decisions may be influenced by recommendations from juvenile probation officers and clinicians.

In an early study, Murrie et al. (2005) tested the influence of a juvenile psychopathy label on the decision-making of 260 juvenile probation officers (JPOs). JPOs were presented with a mock evaluation describing a juvenile offender that varied in terms of diagnostic label (psychopathy/conduct disorder/none), antisocial history (present/not present), and attribution of psychopathic traits (present/not present). The JPOs were asked their opinions about how likely they were to recommend various services (e.g., psychological services, intensive supervision/probation, secure residential placement, commitment to a juvenile facility, and transfer to adult court), and they were also asked about how likely the juvenile was to benefit from mental health services, to commit future crime, and to be a criminal as an adult. Results suggested that the psychopathy label had relatively little effect on JPOs' responses, but that the attribution of psychopathic traits influenced JPOs to perceive the juvenile to be significantly more likely to commit future non-violent crimes, $d = 0.41$, $p < 0.005$, and to be a criminal as an adult, $d = 0.43$, $p < 0.005$. In contrast, a history of antisocial behavior significantly influenced JPOs' responses in six areas (deferred prosecution, secure residential placement, commitment to juvenile correctional facility, transfer to adult court, likelihood of future general/violent crime, and likelihood of adult criminality), with large effect sizes for the latter three, $d = 0.79$, 0.58, and 0.61, respectively, and a medium effect size for transfer to adult court, $d = 0.38$.

Chauhan et al. (2007) examined the impact of the psychopathic label and the attribution of psychopathic traits to juvenile offenders among 83 Virginia district and juvenile court judges, 58 developmental experts, and 64 clinicians. Participants were randomly assigned to one of four conditions describing Michael, a 15-year-old juvenile offender accused of committing a serious violent crime: (a) Michael was given no label, but was given psychopathic traits; (b) Michael was given both the psychopathy label and psychopathic traits; (c) Michael was given a psychopathic label, but no psychopathic traits; and (d) Michael was given no psychopathy label and no psychopathic traits. Among developmental experts, results indicated a moderate effect of a psychopathy label on perceptions of amenability to treatment, $d = 0.64$, but a large effect of giving a juvenile offender both a psychopathy label and psychopathic traits, $d = 1.04$. Clinicians demonstrated a similar labeling effect, with a psychopathy label, regardless of the attribution of psychopathic traits, exerting a moderate effect on clinicians' perceptions of treatment amenability, $d = 0.56$. Generally, no labeling effect was found for judges, but post hoc analyses revealed a significantly larger effect among district court versus juvenile court judges in both conditions in which Michael was given a psychopathic label, with the effect being more pronounced in the label and traits condition, $d = 0.89$, $p < 0.05$ and $d = 1.07$, $p < 0.05$, respectively.

Murrie et al. (2007) examined the impact of antisocial behavior history (minimal vs. substantial), psychopathic personality traits (present vs. not present), and

diagnosis (psychopathy, conduct disorder, or no diagnosis) on 273 members of the National Council of Juvenile and Family Court Judges in a 2 × 2 × 3 design. Judges were provided with a mock evaluation of a juvenile in which the aforementioned variables were manipulated. They were asked about whether they would transfer the juvenile to adult court and the likelihood that the juvenile would benefit from mental health services, would be a criminal as an adult, and would pose a risk of violence to others. No effect of psychopathy diagnosis or attribution of psychopathic traits emerged regarding transfer decisions, but judges were significantly more likely to perceive the juvenile as being at increased risk for violence in the presence of psychopathic traits, $d = 0.22$, $p < 0.01$. Additionally, a significant interaction existed between the presentation of psychopathic traits and antisocial behavior ($p < 0.05$), indicating that there was a greater effect of psychopathic traits on perceptions of future violence when the juvenile had a minimal history of antisocial behavior, $d = 0.72$, $p < 0.001$, than when the juvenile had an extensive history of antisocial behavior, $d = 0.46$, $p < 0.01$.

Rockett et al. (2007) presented 109 clinicians with a mock psychological evaluation, with clinicians randomly assigned to a condition based on the same research design as Murrie et al. (2007). Results indicated a statistically significant two-way interaction between diagnostic label and history of antisocial behavior, with the psychopathy label exerting a significantly greater effect on clinicians' perceptions of violence risk than a label of conduct disorder, $d = 1.36$, $p < 0.01$. Additionally, the psychopathy label tended to produce a greater effect than no diagnosis, $d = 0.78$, though this result was not statistically significant. In the presence of an extensive history of antisocial behavior, however, the three label conditions did not differ. The attribution of psychopathic traits to the juvenile also influenced clinicians' perceptions of risk, with a significant main effect emerging in which clinicians were significantly more likely to perceive the juvenile to be at increased risk for violence when attributed psychopathic traits compared to when he was not attributed these traits, $d = 0.55$, $p = 0.04$.

Vidal and Skeem (2007) provided 204 juvenile probation officers with excerpts from a presentence investigation report that varied in terms of psychopathic label (present vs. not present), juvenile race (African-American vs. Caucasian), and the juvenile's child abuse history (severe abuse vs. no abuse). Juvenile probation officers were asked five questions: (1) the likelihood the juvenile would commit future criminal acts; (2) the likelihood the juvenile would commit future violent acts; (3) the likelihood the juvenile would pose a danger to society; (4) the likelihood the juvenile would become a criminal as an adult; and (5) the likelihood the juvenile would benefit from treatment services. Results indicated that the psychopathy label significantly influenced probation officers' perception of the juvenile as being at increased likelihood of future dangerousness ($p < 0.001$). Additionally, the psychopathy label produced a small effect on expectations of the juvenile's treatment amenability, partial $n^2 = 0.09$, with probation officers being significantly more likely to perceive the juvenile to be of lower treatment amenability in the presence of the psychopathy label versus the no label condition ($p < 0.001$).

Finally, Jones and Cauffman (2008) explored judges' (78 juvenile court, 22 adult court) perceptions of a juvenile's culpability, treatment amenability, dangerousness, and how restrictive of a placement the juvenile should be placed in. Judges were presented with a scenario depicting a juvenile charged with aggravated assault, and then randomly assigned to one of four juvenile description conditions: nonpsychopathic, psychopathic label, psychopathic traits, and psychopathic label and psychopathic traits. The latter three conditions were manipulated by providing judges with a mock expert report. Findings indicated that the judges perceived no difference in culpability among the four conditions, but they perceived the nonpsychopathic juveniles as significantly more amenable to treatment, partial $n^2 = 0.12$, $p < 0.05$, and in need of less-restrictive placements, partial $n^2 = 0.11$, $p < 0.05$, than juveniles described as psychopathic and presenting with psychopathic traits. Additionally, judges perceived the nonpsychopathic juvenile to be significantly less dangerous than juveniles in the psychopathic label, psychopathic traits, and psychopathic label and psychopathic traits conditions, partial $n^2 = 0.25$, $p < 0.0001$.

Taken together, these studies suggest that the label "psychopathy" can have a prejudicial impact in certain contexts, particularly when it comes to predictions and perceptions of future dangerousness. Given that future dangerousness—often phrased as risk to public safety in relevant transfer statutes—is a key criterion judges rely on when making transfer decisions, the potentially prejudicial impact of the psychopathy label can have deleterious effects on transfer decisions. With that said, it is important to note that judges appeared to be less influenced by the psychopathy label than other groups, such as JPOs and clinicians.

Admissibility: Conclusions and Recommendations

Regarding the probative value of the PCL, PCL-R, and PCL: YV to predict recidivism outcomes in juvenile offenders, research firmly suggests that PCL Total and Factor scores are poor predictors of sexual recidivism, but may be weak to moderate predictors of general, nonviolent, and violent recidivism (with Factor 2 tending to be more predictive than Factor 1) (Edens and Campbell 2007; Edens et al. 2007; Olver et al. 2009). However, PCL Total and Factor scores may be less predictive in female and minority samples (Edens et al. 2007). Interestingly, PCL scores may be more predictive of recidivism in Canadian samples than American samples, though that does not eliminate the weak to moderate predictive utility of the PCL for American samples (Olver et al. 2009). Findings are mixed regarding the probative value of the PCL: YV to predict treatment amenability, with some studies indicating that higher PCL: YV Total and Factor scores may be predictive of poor treatment outcomes (Gretton et al. 2001; O'Neil et al. 2003; Spain et al. 2004) and some studies finding no association (Caldwell et al. 2006, 2007).

Regarding prejudice, findings indicate that, in terms of perceived dangerousness/criminal tendency and treatment amenability, diagnostic labels and

the attribution of psychopathic traits to juvenile defendants may result in increased perceptions of the former and decreased perceptions of the latter across clinicians, developmental experts, and juvenile probation officers (Chauhan et al. 2007; Murrie et al. 2005; Rockett et al. 2007). Results also suggest that judges' perceptions of future dangerousness and treatment amenability may be influenced by labeling and trait attribution effects, particularly when a juvenile has little history of antisocial behavior (Murrie et al. 2007). However, this effect may be less influential for juvenile court judges (Chauhan et al. 2007). Notably, no studies revealed a labeling or trait attribution effect on judges' decisions to transfer, which is the ultimate outcome of interest for our purposes.

Given the weak to moderate predictive validity of the PCL, PCL-R, and PCL: YV in predicting juvenile recidivism, the mixed nature of the prejudicial findings, and the lack of a labeling of trait attribution effect on judges' transfer decisions, it is unlikely that admission of testimony based on the PCL, PCL-R, or PCL: YV for purposes of predicting future dangerousness would be substantially more prejudicial than probative; therefore, it appears to satisfy the relevance/prejudice hurdle and is likely admissible. However, given the mixed nature of the predictive ability of the PCL: YV in determining treatment amenability, research does not support the probative value of the PCL-R and PCL: YV to predict poor treatment amenability; as such, the PCL-R and PCL: YV would likely be deemed inadmissible due to irrelevancy under Federal Rules of Evidence for United States Courts and Magistrates 401 when offered for this purpose.

SVP Commitment

The PCL-R is used more frequently in SVP hearings than in any other type of legal proceeding (DeMatteo et al. 2014a, b; DeMatteo and Edens 2006; Walsh and Walsh 2006). In 1990, Washington became the first state to enact a civil commitment law for persons determined to be sexually violent predators (Community Protection Act 1990). The United States Supreme Court subsequently upheld the constitutionality of the Washington law (*Kansas v. Crane* 2002; *Kansas v. Hendricks* 1997). Accordingly, a habitual sex offender can be indefinitely detained following completion of the criminal sentence if he or she has a "mental abnormality" that predisposes him or her to sexually reoffend. SVP laws do not violate prohibitions against double jeopardy (i.e., being tried twice for the same charge) or *ex post facto* laws (i.e., retroactively illegalizing previously legal behavior) because the laws are intended to treat individuals rather than punish them. At the time this chapter was written, 20 states and the federal government have SVP laws (Joint Legislative Audit and Review Commission 2011; *United States v. Comstock* 2010). In 2006, the most recent year for which we have reliable data, 4534 persons in the United States were detained under SVP laws, and only 10.9 % had been conditionally released or discharged (Gookin 2007).

Probative Value

Although there is variability across SVP statutes, central to all of them is the issue of risk for future sexual violence (Janus and Prentky 2008; Petrila and Otto 2001). The PCL-R is routinely administered in SVP evaluations, and 75.6 % of experts who conduct these assessments consider psychopathy essential to the evaluation (Jackson and Hess 2007). Though sexual recidivism of any kind is a concern, of particular importance in SVP hearings is the expected likelihood the individual will sexually reoffend in the *long term* (Jackson 2008). Several state legislatures, including Florida, Iowa, Kansas, New Hampshire, South Carolina, and Texas, have specified that the focus of their SVP programs is the long-term management and treatment of sexually dangerous offenders. Therefore, we will review studies that have examined sexual recidivism over both short (under 10 years) and long (over 10 years) periods of time.

Evidence for the PCL-R to predict short-term sexual recidivism is not as ubiquitous as it is for general and violent recidivism, and most studies that yield significant results find modest effects. Rice et al. (1990) followed 54 adjudicated rapists who had been released from a psychiatric hospital for 46 months. They found that higher PCL-R Total scores were associated with more convictions for sexual offending post-release, $R = 0.36$, $p < 0.05$. Similar findings were found in a follow-up study that included offenders who had assaulted children (aged <14) and extended the follow-up period to 59 months, $R = 0.23$ (Quinsey et al. 1995). In a multi-national meta-analysis that included 13 studies with the PCL-R, Hanson and Morton-Bourgon (2005) found modest predictive validity for sexual recidivism, $d = 0.29$, over an average follow-up period of 5–6 years. A meta-analysis by Hawes et al. (2013) examined 20 PCL-R studies with sex offender samples, and they found that PCL-R Total scores predicted sexual recidivism much more strongly, $d = 0.40$, than reported by previous studies.

On the other end, Barbaree et al. (2001) did not find a significant effect for the PCL-R across an average interval of 4.5 years, but they found a significant effect for the Sex Offender Risk Appraisal Guide (Quinsey et al. 1998), which is a sexual risk assessment measure that incorporates the PCL-R. In an unpublished doctoral dissertation, Langton (2003) failed to find a significant main effect for PCL-R Total scores among a sample of Canadian inmates when sexual recidivism was operationalized as any new conviction for a contact offense involving a sexual element. This finding of null result was replicated in a published follow-up study with an average follow-up period of 5.5 years (Langton et al. 2006). Serin et al. (2001) followed sex offenders (child and adult victimizers) for 7 years and failed to observe a relationship between PCL-R Total scores and reconviction for a sexual offense. Among Swedish rapists, Sjöstedt and Långström (2002) found that PCL-R scores did not predict sexual reconviction over an average of 92 months.

Factor score comparisons are equally mixed. Walters (2003b) reviewed five PCL-R studies with sexual reoffense outcomes and found comparably small effect sizes for both factors. Hawes et al. (2013) concluded in their meta-analysis that

Factor 2 (d = 0.44), but not Factor 1, was significant for predicting sexual recidivism. Similar findings have been noted in other studies (Dietrich et al. 2007; Serin et al. 2001). By contrast, Olver and Wong (2006) noted a trend favoring Factor 1, although these findings were not significant.

A number of methodological limitations may help to explain the inconsistencies of the findings. First, studies vary in outcomes. Recidivism is alternately defined as charges for a sexual offense, convictions for a sexual offense, both/either, and convictions for crimes of a sexual nature. Second, sources for coding reoffense can influence base rates. For instance, Abel et al. (1987) observed one arrest for every 30 self-reported sexual assaults. Third, the PCL-R cutoffs used to create groups are inconsistent. For example, Quinsey et al. (1995) used a cutoff score of 19, whereas Dietrich et al. (2007) utilized a cutoff of 30+. Fourth, studies have alternative categories for sex offenders. Often, sex offenders are typified as "rapists" and "child molesters"; however, the legal standard for child victims varies by jurisdiction. Many researchers do not specify the statutory standard for their sample, making direct comparisons more difficult. Fifth, the sample for each study varies in terms of sex representation; although a majority of studies consider samples consisting only of men, several studies also examine outcomes for women. Finally, almost no research has been conducted on female sex offenders with the PCL-R.

Most studies looking at sexual recidivism and the PCL-R report an average follow-up period of between 4 and 7 years. Very few have tracked recidivism beyond this length of time and even fewer have follow-ups for sexual reoffense that exceed a decade after release. Although results are mixed in terms of long-term sexual recidivism (as described in the next paragraph), the weight of the evidence suggests PCL-R scores are meaningfully related to long-term sexual recidivism, particularly when considered in combination with "deviant sexual arousal" (Hanson and Morton-Bourgon 2009; Hawes et al. 2013). Therefore, it seems that PCL-R-based expert testimony is relevant to the courts' need to address future violence potential.

Rice and Harris (1997) followed sex offenders who had assaulted children, adults, or both over 10 years. The PCL-R Total score was predictive of sexual recidivism, but Factor scores were not reported. Hildebrand, de Ruiter, and de Vogel (2004) examined Dutch forensic psychiatric patients convicted of rape or sexual assault with an average post-release follow-up period of 11.8 years, and they found that PCL-R Total, Factor 1, and Factor 2 scores were significant predictors of sexual reconviction, with the Factor scores being comparably predictive. Parent et al. (2011) examined sexual recidivism among subtypes of sexual aggressors who had been evaluated for SVP civil commitment. Recidivism was measured annually for the entire length of the study (15 years), and offenders were either repeat or aggressive offenders (i.e., in the higher bracket of dangerousness). They found that the PCL-R successfully predicted sexual recidivism for either rapists or child molesters, but not for sex offenders who were both rapists and child molesters. Among nearly 600 sex offenders being evaluated for SVP commitment, Knight and Thornton (2007) looked at sexual recidivism at 3-, 10-, and 15-year follow-up intervals. Results revealed that PCL-R Total and Factor 2 scores significantly

predicted subsequent charges for serious sexual crimes at all follow-up periods. Of note is that Factor 1 was a nonsignificant predictor at the 3- and 10-year follow-up periods, but not the 15-year follow-up. Finally, Olver and Wong (2006) examined federally incarcerated sex offenders over a 10-year period and found that PCL-R Total and Factor scores were poor predictors of sexual recidivism. The results held when the offenders were divided into subtypes. Although neither was significant, Factor 1 appeared to contribute more to prediction of sexual reoffense than Factor 2.

Similar to studies on short-term sexual recidivism, some studies have found that PCL-R Total scores well below the diagnostic cutoff of 30 (i.e., 25 or 26) are still at least modestly predictive of sexual recidivism over the long-term (Hildebrand et al. 2004; Rice and Harris 1997). The role of PCL-R Factor scores is less clear because many studies did not include them in their analyses. All three studies examining PCL-R Factor scores reported an influence for Factor 1 at least comparable to that of Factor 2. In Knight and Thornton's (2007) study, Factor 1 was initially not predictive, but later surfaced as a significant predictor at the end of the 15-year follow-up period when Factor 2 scores began to drop. Both Hildebrand et al. (2004) and Olver and Wong (2006) found that prediction of sexual recidivism was more attributable to Factor 1 than Factor 2, albeit non-significantly. Research suggests the personality traits in Factor 1 are more stable than behaviors comprising Factor 2 (Harpur and Hare 1994; Porter et al. 2001), which may account for the greater predictive ability of Factor 1 scores over prolonged periods of time.

There are a number of methodological limitations to the studies discussed above. Although many studies considered subtypes of sex offenders, usually defined in terms of victims (e.g., adults only, children only, incestual, non-preferential), very few studies examined the predictive validity of the PCL-R for each group. Group analysis is important for understanding sex offenders because this population is not homogeneous; base rates for recidivism, sexual deviance, and psychopathy vary (Olver and Wong 2006; Porter et al. 2000). For several studies (Rice and Harris 1997; Olver and Wong 2006; Hildebrand et al. 2004), follow-up periods were assessed at fixed periods even though offenders were released at variable times (i.e., variable follow-ups), yielding less accurate predictions (Harris and Rice 2006). Recent studies have attempted to account for this statistically (e.g., Cox regression analysis; Hildebrand et al. 2004). Early studies did not track recidivism at intervals over time, but recent studies are accounting for longer term prediction. Because recidivism rates drop and accurate prediction becomes more difficult as time passes, it is important to assess the PCL-R's predictive accuracy at different intervals across time. Understanding the trajectory of risk and the temporal importance of the PCL-R (Total and Factor scores) is important in terms of long-term risk assessment.

Prejudicial Impact

Several factors may exert biasing effects in SVP hearings, including the perceived attractiveness of the sex offender (Esses and Webster 1988), emotionally evocative

victim statements (Jackson et al. 2004), clinical expert testimony (Krauss et al. 2011), the victim's age (Guy and Edens 2006), and the decision-maker's gender (McCabe et al. 2010). Because nearly one-quarter (23.6 %) of SVPs are classified as "psychopaths" (i.e., PCL-R score of 30+; Vess et al. 2004), it is germane to the issue of SVP commitment to examine the biases associated with this label.

Guy and Edens (2003) supplied undergraduate students with the same information provided to jurors in Texas SVP trials and asked them to decide if the defendant should be committed under Texas's SVP act. The perpetrator was male and his victims were female. Among males, commitment decisions were not influenced by the psychopathy label, but females were more likely to commit those described by the prosecution as psychopathic. In a follow-up study, Guy and Edens (2006) manipulated the victim's age. Although a gender effect was again observed, it did not contribute above the bias created when the victim was a child (aged 8 or 10).

Lieberman and Krauss (2009) compared biases of the psychopathy label to another diagnosis in SVP hearings. Undergraduate students were asked to make a commitment decision for a male offender diagnosed with psychopathy or a paraphilia who had an extensive sexual offense history and victimized children of both genders. The psychopathy condition used the term "psychopath" but described criteria for Antisocial Personality Disorder from the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (American Psychiatric Association 2013). Contrary to Guy and Edens (2006), a psychopathy diagnosis created bias, and both men and women in the psychopathy condition were more likely to support commitment.

Taken together, it appears the psychopathy label creates a bias under certain conditions in SVP hearings. Women may be more likely than men to be influenced by psychopathy, but this gender effect may be subsumed by the emotion evoked when the victim is a child. Guy and Edens (2006) found that both men and women had a strong assumption that the sex offender was psychopathic even before reading the expert witness's testimony, and these perceptions significantly increased after being presented with expert testimony.

These studies share a number of limitations. For all studies, transcripts were used rather than realistic simulations such as videotaped SVP hearings and mock deliberations, and the participants were undergraduates. Although the ecological validity of such samples is debatable (McCabe et al. 2010), critics have claimed that ecological validity is crucial for juror data to be meaningful (Diamond 1997). Indeed, some court cases have refused experimental evidence on the grounds of misrepresented sample populations (*Lockhart v. McCree* 1986).

Admissibility: Conclusions and Recommendations

Extant research evidence suggests that the ability of the PCL-R to predict short-term sexual recidivism is ambiguous. Therefore, the PCL-R does not appear to have

probative value when offered to predict short-term sexual recidivism, and it should be deemed inadmissible due to irrelevance when offered for this purpose. Although there appears to be tentative evidence for the PCL-R to modestly predict long-term sexual recidivism, only three studies have examined follow-up periods in excess of 10 years. Evidence for labeling effects suggests that there is not sufficient prejudicial impact to substantially outweigh the probative value of the PCL-R for predicting long-term sexual recidivism. In light of few probative value studies and inconclusive evidence of prejudicial impact, the PCL-R is likely to be admitted for purposes of long-term SVP commitment considerations.

Capital Sentencing

Capital sentencing is the process by which criminal offenders are sentenced to death or life in prison after being convicted of a capital offense. Due to the severe liberty deprivation inherent in capital sentencing, the United States Supreme Court has ruled that many constitutional restrictions exist in the imposition of the death penalty: (1) the death penalty cannot be mandatorily imposed (*Roberts v. Louisiana* 1976); (2) the death penalty can only be imposed for crimes involving death (*Kennedy v. Louisiana* 2008); (3) the death penalty cannot be imposed on individuals who were juveniles when they committed their offenses (*Roper v. Simmons* 2005), individuals who are intellectually disabled (*Atkins v. Virginia* 2002), or individuals who are not competent to be executed (*Ford v. Wainwright* 1986; *Panetti v. Quarterman* 2007); (4) the jury, not the judge, must determine the presence of aggravating factors (*Ring v. Arizona* 2002), and the death penalty must be based on individualized consideration (*Lockett v. Ohio* 1978), and must have allowed for consideration of all mitigating circumstances (*Eddings v. Oklahoma* 1982; *Tennard v. Dretke* 2004); and (5) the death penalty cannot be issued arbitrarily (*Furman v. Georgia* 1972), meaning that legislatures must provide juries with objective criteria to direct and limit the imposition of the death penalty, typically accomplished via consideration of aggravating and mitigating factors (*Gregg v. Georgia* 1976).[4]

The PCL-R is used in capital cases in two primary ways. First, the PCL-R is sometimes used by the prosecution during the guilt/innocence phase of capital trials

[4]Note, however, that while schemes involving the consideration of aggravating and mitigating circumstances predominate, not all jurisdictions follow such a scheme. For example, in choosing between a punishment of life or death at sentencing, Texas law asks jurors to consider the following question after finding a defendant guilty of a capital offense: "Whether, taking into consideration all of the evidence, including the circumstances of the offense, the defendant's character and background, and the personal moral culpability of the defendant, there is a sufficient mitigating circumstance or circumstances to warrant that a sentence of life imprisonment without parole rather than a death sentence be imposed" (TEX. CODE CRIM. PROC. ANN. art. 37.071 (West 2015).

to rebut a defense of insanity by demonstrating that the defendant is malingering (Edens et al. 2013). Second, the PCL-R is often used in the sentencing phase of capital trials to assess the defendant's likelihood of being dangerous in the future. In response to *Furman*, most states scrambled to adopt a set of guidelines with which to provide sentencing juries in capital cases; these schemes often involve jury consideration of a broad range of mitigating factors and a more narrow set of statutorily defined aggravating factors (Sevilla 1999). Several states allow future dangerousness to serve as an aggravating factor for the death penalty, and the PCL-R is frequently used by attorneys and experts as a means for assessing this (DeMatteo et al. 2014b; DeMatteo and Edens 2006; Edens and Cox 2012; Walsh and Walsh 2006). Despite long-standing questions over the ability of mental health professionals to accurately and reliably identify those offenders who will be dangerous in the future, the United States Supreme Court has upheld the submission of expert testimony in this area (*Barefoot v. Estelle* 1983).

Probative Value

In capital trials, the two outcomes of interest that the PCL-R is used to inform are malingering, or the exaggeration or feigning of psychiatric symptoms for secondary gain, and future dangerousness. Limited research provides some support for an association between PCL-R-measured psychopathy and malingering. In a study comparing 18 successful insanity malingerers to 18 non-malingering insanity acquittees, Gacono, Meloy, Sheppard, Speth, and Roske (1995) found that the malingering group had significantly higher PCL-R Total, Factor 1, and Factor 2 scores than the non-malingerers (34.9 vs. 19.4, 14.8 vs. 6.3, and 15.8 vs. 10.6, $p < 0.001$, respectively). Similarly, in a review of hospital records for 374 men in a maximum-security psychiatric hospital categorized as being either low, medium, or high in psychopathy based on PCL-R score, Heinze and Vess (2005) identified 18 individuals as being diagnosed malingerers, and they found that a significantly greater percentage of malingerers were in the high-psychopathy group, $\chi^2(1) = 6.344$, $p = 0.012$. This result increased in significance when comparing just the high and low groups, $\chi^2(1) = 7.016$, $p = 0.008$. Although supportive of the association between PCL-R scores and malingering, both studies used small sample sizes and malingering was determined by previous diagnosis.

Kucharski et al. (2006) examined the relationship between PCL-R scores and malingering among 207 male criminal defendants referred for competency to stand trial. The defendants were split into low-, medium-, and high-psychopathy groups, with malingering assessed by scores on the Minnesota Multiphasic Personality Inventory (MMPI-2; Butcher et al. 1989) Infrequency (*F*), *F–K* Index, and Infrequency Psychopathology (*F*(*p*)) subscales; scores on the Psychological Assessment Inventory (PAI; Morey 1991) Negative Impression (NIM), Rogers Discriminant Function (RDF), and Malingering Index subscales; and score on the Structured Interview of Reported Symptoms (SIRS; Rogers et al. 1991). Findings

indicated that the high-psychopathy group exhibited significantly greater malingering than the low-psychopathy group on the MMPI-2 F, $F(2, 189) = 8.43$, $p < 0.001$, *F–K*, $F(2, 189) = 10.20$, $p < 0.001$, and $F(p)$, $F(2, 189) = 7.19$, $p = 0.001$, subscales; additionally, the high-psychopathy group exhibited significantly greater malingering as determined by the PAI NIM subscale, $F(2, 164) = 6.63$, $p = 0.002$, and SIRS, $F(2, 107) = 6.18$, $p = 0.003$.

Concerning future dangerousness, several meta-analyses have examined the PCL-R's ability to predict institutional infractions. Walters (2003a) compared the PCL/PCL-R and Lifestyle Criminality Screening Form (Walters et al. 1991) in terms of predicting institutional adjustment. Results indicated that both measures were weak to moderate predictors of institutional adjustment, with the PCL/PCL-R being slightly more predictive, $k = 15$, $r_w = 0.27$ vs. $k = 7$, $r_w = 0.23$, respectively. However, confidence intervals for the two instruments overlapped substantially, 95 % $CI_w = 0.23–0.32$ and 0.16–0.29, respectively. Further, Walters (2003a) examined only total instances of institutional adjustment as opposed to any subcategories of institutional adjustment. Similarly, Walters (2003b) examined the ability of PCL/PCL-R factor scores to predict institutional adjustment, finding that Factor 1 was a weak predictor of violent infractions, $k = 14$, $r_w = 0.12$, $Q = 24.90$, $p < 0.05$, while Factor 2 was a stronger predictor, $k = 22$, $r_w = 0.22$, $Q = 27.24$, $p < 0.05$. Further, Factor 2 predicted institutional violence better than Factor 1 (which approached significance), but both Factor 1 and Factor 2 more strongly predicted institutional adjustment as a whole, $r_w = 0.18$ and 0.27, as opposed to violent institutional adjustment.

Guy et al. (2005) examined the ability of PCL-R Total and Factor scores to predict instances of institutional misconduct, finding that Total scores were a weak predictor of physical violence, $k = 22$, $r_w = 0.17$, $Q = 24.93$, ns, but a stronger predictor of verbal/destructive aggression, $k = 15$, $r_w = 0.26$, $Q = 16.25$, ns, and aggression in general, $k = 31$, $r_w = 0.23$, $Q = 50.95$, $p < 0.01$. In terms of Factor 1, findings indicated that it was a weak predictor of physical violence, $k = 16$, $r_w = 0.14$, $Q = 13.41$, ns, and general aggression, $k = 22$, $r_w = 0.15$, $Q = 15.78$, ns, but a slightly stronger predictor of verbal/destructive aggression, $k = 9$, $r_w = 0.20$, $Q = 6.93$, ns. Results indicated Factor 2 was a weak predictor of physical violence, $k = 16$, $r_w = 0.15$ $Q = 16.52$, ns, but a slightly stronger predictor of general aggression, $k = 22$, $r_w = 0.20$, $Q = 29.25$, ns, and a stronger predictor of verbal/destructive aggression, $k = 9$, $r_w = 0.24$, $Q = 9.26$, ns. Although PCL-R Total, Factor 1, and Factor 2 scores were somewhat predictive of institutional aggression, they tended to be better predictors of institutional misconduct, both aggressive and nonaggressive, $r_w = 0.29$, 0.21, and 0.27, respectively, than institutional aggression. Leistico et al. (2008) conducted a similar analysis, and findings indicated that PCL-R Total, Factor 1, and Factor 2 scores produced moderate to large effect sizes regarding institutional infractions both generally and specifically (nonviolent vs. violent), with Total scores producing mean weighted effect sizes of $d_w = 0.53$, 0.59, and 0.46; Factor 1, $d_w = 0.41$, 0.37, and 0.40; and Factor 2, $d_w = 0.51$, 0.60, and 0.57, respectively. Taken together, these results suggest that the PCL-R has some ability to predict malingering, though the evidence base is

small; additionally, a mixed evidence base provides ambiguous results for the usefulness of the PCL-R in predicting institutional violence.

Prejudicial Impact

To date, only a few studies have examined whether labeling a defendant a "psychopath" renders jurors more likely to sentence a capital defendant to death, or to perceive a capital defendant as insane or as a continuing threat to society. Regarding insanity, one study to date has examined the role of labeling effects in mock juror perceptions of insanity. Rendell, Huss, and Jensen (2010) provided 383 undergraduate students with a trial transcript based on the facts of *United States v. Weed* (2004), in which the defendant was charged with second-degree murder for killing a federal employee. Transcripts varied according to a 3 (prosecution rebuttal argument: no disorder/personality disorder/psychopath) × 2 (basis for the defense expert's diagnosis of defendant as schizophrenic: biological evidence/performance on traditional psychological tests) × 2 (defense case strength: moderately strong/moderately weak) design, and participants were asked to render a verdict of guilty or not guilty by reason of insanity (NGRI). Results indicated significant main effects for both defense strength and for defense expert's basis for a schizophrenia diagnosis, but found no significant differences between the three prosecution rebuttal evidence conditions (which contained the psychopathy manipulation) in terms of mock jurors' decisions to render an NGRI verdict, $\chi^2(1, 382) = 2.486$, $p = 0.288$, $n^2 = 0.078$.

Several studies have investigated the role of labeling and the attribution of psychopathic traits to defendants in capital offenses. Edens et al. (2004) presented 301 undergraduate students with a case summary describing a homicide offense and a summary of expert testimony that varied in terms of the defendant's diagnosis (psychopathy/psychosis/no diagnosis) and defendant risk level for future violence (low/high); in the psychopathy condition, the defendant's violence risk was described as deriving from PCL-R Factor 1 traits. Participants were also provided with a simplified version of the legal criteria that must be met to sentence a defendant to death in Texas. Results revealed a main effect in terms of diagnosis such that defendants with psychopathy and psychosis were judged to be more dangerous than defendants without a diagnosis, $t(167) = 2.78$, $p < 0.05$, $d = 0.43$, and $t(161) = 2.52$, $p < 0.05$, $d = 0.40$, respectively; however, defendants with psychopathy and psychosis did not differ from each other, $t(142) = 2.78$, $p =$ ns, $d = 0.04$. Further, though risk level significantly influenced perceptions of dangerousness in the no diagnosis condition, $t(92) = 3.68$, $p < 0.001$, $d = 0.07$, this effect did not hold for defendants with psychopathy, $t(73) = 0.02$, $p =$ NS, $d = 0.22$, suggesting that mock jurors' perceptions of dangerousness were based mainly on diagnostic label for defendants with psychopathy. However, the generalizability of these results may be limited because a post-test measure suggested that

many of the mock jurors had a poor understanding of the legal criteria necessary to impose the death penalty.

Edens et al. (2005) sought to rectify that limitation by replicating Edens et al. (2004) with 233 undergraduate students who were provided with clearer jury instructions indicating that the defendant should be put to death only if no mitigating factors were present and the prosecution proved beyond a reasonable doubt that the defendant would represent a continuing threat to society due to a high propensity to commit future violence. In terms of perceived dangerousness, results generally mirrored Edens et al. (2004), with the exception that no significant difference existed between low- and high-risk defendants in any of the diagnostic conditions. Diagnostic label did, however, influence mock jurors' decisions to impose the death penalty on the defendant, with defendants with psychopathy being significantly more likely to be sentenced to death than either defendants without a diagnosis, $\beta = 0.91$, $p = 0.01$, $R = 0.13$, or defendants with psychosis, $\beta = 1.25$, $p < 0.01$, $R = 0.19$.

Cox et al. (2010) provided 144 undergraduate students with vignettes describing the trial phase of a capital offense and then measured their perceptions of the defendant's dangerousness and their decision to sentence the defendant to life without parole (LWOP) or death. All vignettes contained identical descriptions of the trial phase, but differed in regards to expert testimony presented according to a 2 (diagnostic label: no label vs. psychopath) $\times$ 2 (danger risk: low vs. high) design. It should be noted that in both diagnostic conditions, psychopathic traits were attributed to the defendant; however, in the psychopath condition, the defendant was described as being labeled a psychopath due to the results of the PCL-R as interpreted by a court-appointed clinician. Results revealed no significant impact of diagnostic label on either decisions to impose the death penalty or perceptions of future violence, but there was a significant difference between the low- and high-violence risk groups in terms of mock jurors' perceptions of whether the defendant would commit another violent crime, $F(3, 140) = 14.62$, $p = 0.001$, $n^2 = 0.24$, or murder, $F(3, 140) = 6.14$, $p = 0.001$, $n^2 = 0.13$, if given LWOP. Further, a chi-square analysis revealed a significant difference between the number of participants expected to impose the death penalty (using a chi-square test of independence) and the number of participants that actually imposed the death penalty, $\chi^2(3, n = 144) = 19.21$, $p = 0.001$, Cramer's $V = 0.378$), with the difference occurring between defendants described as high- or low-violence risk, regardless of diagnostic label.

Finally, Edens et al. (2013) aggregated data from the control conditions of the Edens et al. (2003)[5] and Edens et al. (2005) studies in addition to data from an unpublished master's thesis (Davis 2003) to explore the incremental predictive

[5]This study examined the influence of the psychopathy label on mock jurors' perceptions of a juvenile capital offender. This study was conducted before *Roper v. Simmons*, which is the 2005 United States Supreme Court case that abolished the juvenile death penalty. As this study was targeted specifically at a juvenile capital offender, which can now no longer exist, its results are not explored in detail here.

validity of subcomponents of psychopathy as a construct and individual PCL-R items on mock jurors' attitudes toward capital punishment. This resulted in a total sample of 362 undergraduate students. Each study asked participants to rate the extent to which someone like the defendant presented in the stimulus materials for the study would be likely to display similar behavior, symptoms, and traits to those encapsulated by the PCL-R. The stimulus materials in the Edens et al. (2005) and Davis (2003) studies were based on the facts from the *Barnette* case, while the stimulus materials in the Edens et al. (2003) study were based on a Missouri newspaper article about a juvenile capital offender. Receiver Operating Characteristic (ROC) curve analyses indicated strong effects for both global psychopathy scores and for Factor 1 scores in all three studies, with the lowest area under the curve (AUC) and significance level being AUC = 0.66 ($p < 0.05$) for general offending and AUC = 0.72 ($p < 0.01$) for Factor 1. Additionally, when the aggregate sample was split by scale and subscale midpoint, only 15 % of participants supported a death sentence when the psychopathy score was less than 20 while 43 % supported a death sentence when the psychopathy score was greater than 20, $\chi^2 = 29.18$, $p < 0.001$. Further, when the researchers controlled for the predictive validity of Factor 1 scores, psychopathy ratings failed to significantly predict support for the death penalty, which suggests that Factor 1 scores primarily drove support for the death penalty in the study.

Taken together, these studies suggest that labeling a capital defendant a "psychopath" makes jurors more likely to vote for a death sentence. Further, the label "psychopath" makes jurors more likely to perceive the capital defendant as insane or dangerous. These unintended consequences of the label "psychopath" could therefore have significant implications in capital cases.

Admissibility: Conclusions and Recommendations

Regarding prosecutorial assertions of malingering as a rebuttal to the insanity defense, research suggests that PCL-R Total and Factor scores may have some probative value in identifying malingerers, though the research is thin and often limited by small sample sizes. Regarding prejudice, the lone study exploring the effect of a psychopathy label on juror decisions regarding insanity indicates that a label of psychopathy predicts juror behavior no better than either a label of disordered personality or no label at all. Thus, given that PCL-R scores are at least somewhat probative of malingering and a label of psychopathy does not appear to produce differential rates of guilty verdicts among mock jurors, the use of the PCL-R should/would likely be admissible for prosecutorial rebuttal of the insanity defense.

Regarding institutional misconduct, some meta-analyses indicate that PCL-R Total and Factor scores are moderate predictors of institutional violence, but the weight of the evidence suggests that PCL-R scores are weak predictors of institutional violence and are better predictors of general institutional misconduct.

Therefore, the PCL-R has limited probative value when offered to predict institutional violence, and it should be deemed inadmissible due to irrelevance when offered for this purpose. Regarding prejudice, though one study found no impact for diagnostic labels, the accumulated literature suggests a significantly greater likelihood to view psychopathic or mentally ill offenders as dangerous, with concomitant increased rates of imposition of a death sentence. Given the mixed findings regarding *both* probativeness and prejudice for the PCL-R in terms of violent institutional misconduct and labeling effects, it is difficult to conclude that PCL-R-based expert testimony is *substantially* more prejudicial than probative. Nevertheless, if PCL-R evidence is determined by a judge to be relevant, it should be disqualified as overly prejudicial given the magnitude of the liberty interest at stake.

Conclusions

Due to the increased use of the PCL-R within legal contexts, this chapter reviewed the scientific literature regarding the probative value (predictive ability) and prejudicial impact (psychopathy labeling effects) of the PCL-R in three legal contexts: juvenile transfer, SVP commitment, and capital sentencing. An application of Federal Rules of Evidence for United States Courts and Magistrates 401–403 to these literature bases suggests PCL-R evidence is likely to satisfy the relevance/prejudice test regarding recidivism in juvenile transfer proceedings, long-term sexual recidivism for SVP proceedings, and malingering in capital sentencing. However, PCL-R evidence is unlikely to satisfy the relevance/prejudice test for treatment amenability in juvenile transfer proceedings, for short-term sexual recidivism in SVP commitment proceedings, and for predicting future institutional dangerousness in capital contexts.

Regarding juvenile transfer decisions, the scientific literature suggests that the PCL-R yields modest prediction for most outcomes, including general and violence recidivism and institutional misconduct, but the evidence for sexual recidivism is rather weak. The evidence concerning treatment amenability is ambiguous. The prejudicial impact for perceptions of dangerousness and treatment amenability exists, but the research does not support any prejudicial impact for the main outcome of judges' transfer decisions. Therefore, the PCL-R is likely to be admissible when offered to support concerns regarding recidivism, but should not be admissible when offered to show poor treatment amenability.

The PCL-R is likely to be admitted for SVP proceedings. It has demonstrated predictive ability for sexual recidivism over long-term periods of time, but has ambiguous predictive ability regarding short-term sexual recidivism. Preliminary evidence suggests a psychopathy diagnosis increases perceptions of dangerousness and commitment decisions, but the extent to which this generalizes to other mental illnesses or applies to psychopathic traits has been unexplored. Regardless, limitations in the rigor of labeling studies suggest that while prejudicial impact exists

for a psychopathy label, this prejudicial impact does not substantially outweigh the probative value of the PCL-R for long-term sexual recidivism. As such, the PCL-R is likely to be admitted in this context when offered as evidence of long-term sexual recidivism.

Finally, in capital sentencing contexts, the PCL-R is a modest predictor of institutional infractions, but an ambiguous (at worst) and very weak (at best) predictor of institutional violence. The evidence suggests that perceptions of dangerousness in the institution may be influenced by any mental illness diagnosis, but that the psychopathy label and possibly psychopathic traits render a death sentence more likely. Given the PCL-R's limited predictive ability regarding institutional violence by capital offenders, an outcome with a very low base rate,[6] and research supporting labeling effects, the prejudicial impact of a psychopathy label substantially outweighs the probative value of the PCL-R with respect to future dangerousness in the form of institutional violence. Regarding malingering, although the research is limited, the PCL-R appears to distinguish malingerers from non-malingerers, and there does not appear to be a psychopathy labeling effect regarding malingering. Therefore, the PCL-R is likely to be admissible in capital sentencing contexts if offered by the prosecution to rebut an insanity defense, but it is unlikely to be admitted if offered as proof of future dangerousness while incarcerated.

In terms of the practical impact of this analysis, the focus of this chapter was to help facilitate a more uniform and research-informed approach among legal decision-makers regarding the admissibility of PCL-R evidence in the contexts in which it is commonly offered. An inference that arises from our findings is that the PCL-R's admissibility can be challenged in some situations on grounds of irrelevancy or overly prejudicial impact in each of the three contexts that were reviewed. Attorneys should be challenging admission of the PCL-R on grounds of irrelevance during transfer proceedings when offered as evidence that a juvenile is not amenable to treatment. Attorneys should challenge admission of the PCL-R on grounds of irrelevance when offered as evidence of potential short-term sexual recidivism in SVP proceedings. Attorneys should challenge admission of the PCL-R on grounds of irrelevance when offered as evidence of future institutional dangerousness in capital sentencing contexts. Finally, if irrelevancy challenges to the PCL-R regarding short-term sexual recidivism in SVP commitment proceedings or regarding future dangerousness in capital sentencing are not successful, the PCL-R's admission should be challenged on grounds of undue prejudice.

Unfortunately, research indicates that challenges are rare. In the DeMatteo et al. (2014b) case law review, of the 215 SVP cases in which PCL-R evidence was offered, it was challenged in just 13 (6 %) cases, and on 401–403 grounds in only 6 (3 %) cases. For capital sentencing, PCL-R evidence was challenged in only 2 of

[6]Estimates place the base rate of institutional violence in general among capital offenders at under 20 %, and of institutional violence involving serious injury or death at only 2 and 1 %, respectively (Cunningham et al. 2011; Sorensen and Pilgrim 2000; Sorensen and Cunningham 2007; Sorensen and Cunningham 2010).

the 12 cases in which it was offered (16 %), and only one of those two challenges was on 401-403 grounds. Although none of the SVP 401–403 challenges were successful, the 401–403 challenge for the capital case was successful. Additionally, in the DeMatteo and Edens (2006) case law review, two capital cases involved 401–403 challenges and both were successful. Although it is not prudent to extrapolate from such a small sample, these findings provide some evidence that judges may be willing to disqualify PCL-R-based expert testimony on 401–403 grounds when challenges are made.

Our review of the literature and application of the relevance/prejudice hurdle to the three legal contexts covered in this chapter has several limitations. We evaluated the issues of probative value and prejudicial impact through the scientific literature on risk assessment, treatment amenability, malingering, and labeling effects, but correspondence between the psychological research and these legal constructs is not perfect. For instance, the concept of dangerousness as used by the legal system does not have a direct analog in psychology; psychologists use risk for recidivism to inform this construct. Similarly, the legal system conceptualizes "treatment" differently than the mental health system (Tate and Redding 2005). In the legal context, treatment is often coterminous with "rehabilitation" and directly corresponds to a reduction in recidivism. By contrast, mental health treatment can pertain to any number of outcomes linked to functional impairment, such as attendance, quality of participation, program compliance, and clinical improvement. Recidivism is one of many proxies for determining whether treatment is effective.

Concerning probative value, many of our findings were pulled from meta-analyses due to the expansive literature on the PCL-R and recidivism. Although meta-analyses are a convenient method for making sense of extant findings, criticisms have been levied against meta-analyses. Studies within a meta-analysis can vary considerably in terms of study design, outcome, setting, sample, statistical technique, and time at risk (Douglas et al. 2006), so meta-analyses may produce heterogeneous coefficients (Gendreau et al. 2002; Walters 2003a) or exclude relevant studies (Hemphill and Hare 2004). Additional considerations are warranted for the outcome of violence. Broader definitions of violence, longer follow-up periods, and self-report measures of violence may artificially inflate base rates of recidivism (Hemphill et al. 1998; Monahan et al. 2001). In light of the heterogeneity of findings and blending of violent and non-violent outcomes, researchers have cautioned against relying solely on meta-analyses for determining the strength of the PCL-R to predict violence (DeMatteo et al. 2010). Additionally, we were unable to address the relevance/prejudice hurdle for certain understudied populations, such as females in the risk assessment literature. Continued research with diverse samples may help explain mixed findings in areas like sexual recidivism.

Research on labeling effects remains scarce. A notable limitation with many of these studies is that psychological outcomes, such as perception of future dangerousness, were either measured with one item, which threatens construct validity, or grouped with other items to form clusters, which impedes discriminant validity. At best, these clusters demonstrated moderate internal consistency. Such variability among items may indicate

construct differences between perceptions of future dangerousness and other related concepts, such as punishment decisions and predictions of lifetime criminality. Future studies should strive to design a priori items that cluster well together rather than grouping items post hoc. Future research should also explore interactions among psychopathy and other biasing factors (e.g., victim statements) or evidence (e.g., eyewitness testimony), employ more sophisticated outcome measures (e.g., multi-item constructs, reverse worded items), and distinguish between perceptions about the offender and the legal decisions. Studies should better replicate the legal process using jury deliberation or mock trials (Diamond 1997). Non-representative samples, such as undergraduates, may not be representative of jury venires in certain legal contexts (McCabe et al. 2010), and the United States Supreme Court has expressed "serious doubts" about findings based on unrepresentative jury samples that omit group deliberation (*Lockhart v. McCree* 1986, p. 171).

This chapter provides an approach for examining the admissibility of the PCL-R in different legal contexts, but it may not be appropriate to apply this approach in individual cases. The PCL-R's probative value has yet to be established for mentally ill individuals across all types of risk assessment (e.g., sexual dangerousness) and predictive accuracy for female offenders and female psychiatric patients is not well established. Also, the research on labeling effects suggests prejudice may influence legal decision-makers differently depending on the legal question being addressed, the role of the decision-maker, and the type of label being used.

A point of concern that has been highlighted by the findings in this chapter is the problematic use of the moniker "psychopath" by mental health professionals. There are two grounds for why this should be actively discouraged. First, professional standards discourage the use of diagnostic labels to avoid impressions that the person as a whole is defective, equating the individual with his or her disorder, and engendering negative stereotypes (American Psychological Association 2010). There has been a gradual shift in the literature to use mental disorders as qualifiers rather than as labels. However, there has not been this shift within psychopathy research despite the stakes for having such a diagnosis in the legal system. The field continues to use the term "psychopath" in clinical practice, expert testimony, research, and even in the PCL-R manual. Second, an abundance of research indicates that psychopathic traits occur on a continuum and do not represent a discrete taxon (e.g., Edens et al. 2011; Guay et al. 2007; Lynam and Derefinko 2006). According to FRE Rule 702, an expert's testimony must be based on sufficient evidence and the product of reliable scientific principles or methods. Using a categorical label that is not empirically supported may therefore encroach upon the federal parameters for expert opinions.

The findings of this chapter underscore the importance of competent testimony by clinicians. The PCL-R may satisfy the relevance/prejudice test, but it is the responsibility of the expert witness to communicate the limitations of this evidence and use language that is consistent with evidence, ethical principles, and legal standards. Such practice would comport more closely with applicable ethical standards and ultimately offer more meaningful information to legal decision-makers.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, D.C.

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC.

Abel, G. G., Becker, J. V., Mittelman, M., Cunningham-Rathner, J., Rouleau, J. L., & Murphy, W. D. (1987). Self-reported sex crimes of nonincarcerated paraphiliacs. *Journal of Interpersonal Violence, 2*, 3–25. doi:10.1177/088626087002001001

Atkins v. Virginia, 536 U.S. 304 (2002).

Barbaree, H. E., Seto, M. C., Langton, C. M., & Peacock, E. J. (2001). Evaluating the predictive accuracy of six risk assessment instruments for adult sex offenders. *Criminal Justice and Behavior, 28*, 490–521. doi:10.1177/009385480102800406

Babiak, P. (1995). When psychopaths go to work: A case study of an industrial psychopath. *Applied Psychology: An International Review, 44*, 171–188. doi:10.1111/j.1464-0597.1995.tb01073.x

Barefoot v. Estelle, 463 U.S. 880 (1983).

Butcher, J., Dalstrom, W., Graham, J., Tellegen, A., & Kraemmer, B. (1989). *Manual for administering and scoring the MMPI-2*. Minneapolis, MN: University of Minnesota Press.

Caldwell, M. F., McCormick, D., Umstead, D., & Van Rybroek, G. (2007). Evidence of treatment progress and therapeutic outcomes among adolescents with psychopathic features. *Criminal Justice and Behavior, 34*, 573–587. doi:10.1177/0093854806297511

Caldwell, M., Skeem, J., Salekin, R., & Van Rybroek, G. (2006). Treatment response of adolescent offenders with psychopathy features: A 2-year follow-up. *Criminal Justice and Behavior, 33*, 571–596. doi:10.1177/0093854806288176

Chauhan, P., Reppucci, N. D., & Burnette, M. L. (2007). Application and impact of the psychopathy label to juveniles. *International Journal of Forensic Mental Health, 6*, 3–14. doi:10.1080/14999013.2007.10471245

Cleckley, H. (1976). *The mask of sanity* (5th ed.). St. Louis, MO: Mosby.

Community Protection Act, R.C.W. §71.09 (1990).

Cooke, D. J., & Michie, C. (1997). An item response theory evaluation of Hare's Psychopathy Checklist. *Psychological Assessment, 9*, 2–13. doi:10.1037/1040-3590.9.1.3

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment, 13*, 171–188. doi:10.1037/1040-3590.13.2.171

Cox, J., DeMatteo, D. S., & Foster, E. E. (2010). The effect of the psychopathy checklist–revised in capital cases: Mock jurors' responses to the label of psychopathy. *Behavioral Sciences and the Law, 28*, 878–891. doi:10.1002/bsl.958

Cunningham, M. D., Sorensen, J. R., Vigen, M. P., & Woods, S. O. (2011). Correlates and actuarial models of assaultive prison misconduct among violence-predicted capital offenders. *Criminal Justice and Behavior, 38*, 5–25. doi:10.1177/0093854810384830

D'Silva, K., Duggan, C., & McCarthy, L. (2004). Does treatment really make psychopaths worse? A review of the evidence. *Journal of Personality Disorders, 18*, 163–177. doi:10.1521/pedi.18.2.163.32775

Davis, K. M. (2003). *The influence of the label 'psychopath' on mock juror decisions in a capital trial*. Unpublished master's thesis, Sam Houston State University, Huntsville, TX.

DeMatteo, D., & Edens, J. F. (2006). The role and relevance of the Psychopathy Checklist-Revised in court: A case law survey of U.S. courts (1991–2004). *Psychology, Public Policy, and Law, 12*, 214–241. doi:10.1037/1076-8971.12.2.214

DeMatteo, D., Edens, J. F., Galloway, M., Cox, J., Smith, S. T., & Formon, D. (2014a). The role and reliability of the Psychopathy Checklist-Revised in U.S. sexually violent predator evaluations: A case law survey. *Law and Human Behavior, 38*, 248–255. doi:10.1037/lhb0000059

DeMatteo, D., Edens, J. F., Galloway, M., Cox, J., Smith, S. T., Koller, J. P., et al. (2014b). Investigating the role of the Psychopathy Checklist-Revised in United States case law. *Psychology, Public Policy, and Law, 20*, 96–107. doi:10.1037/a0035452

DeMatteo, D., Edens, J. F., & Hart, A. (2010). The use of measures of psychopathy in violence risk assessment. In R. K. Otto & K. S. Douglas (Eds.), *Handbook of violence risk assessment* (pp. 19–40). New York, NY: Taylor & Francis.

Diamond, S. S. (1997). Illuminations and shadows from jury simulations. *Law and Human Behavior, 21*, 561–671. doi:10.1023/A:1024831908377

Dietrich, A. M., Smiley, W. C., & Frederick, C. (2007). The roles of childhood maltreatment and psychopathy in sexual recidivism of treated sex offenders. *Journal of Aggression, Maltreatment and Trauma, 14*, 19–31. doi:10.1300/J146v14n03_02

Douglas, K. S., Vincent, G. M., & Edens, J. F. (2006). Risk for criminal recidivism: The role of psychopathy. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 533–544). New York, NY: Guilford Press.

Eddings v. Oklahoma, 455 U.S. 104 (1982).

Edens, J. F. (2001). Misuses of the Hare Psychopathy Checklist-Revised in court. *Journal of Interpersonal Violence, 16*, 1082–1093. doi:10.1177/088626001016010007

Edens, J. F. (2006). Unresolved controversies concerning psychopathy: Implications for clinical and forensic decision making. *Professional Psychology: Research and Practice, 37*, 65–69. doi:10.1037/0735-7028.37.1.59

Edens, J. F., & Campbell, J. S. (2007). Identifying youths at risk for institutional misconduct: A meta-analytic investigation of the Psychopathy Checklist measures. *Psychological Services, 4*, 13–27. doi:10.1037/1541-1559.4.1.13

Edens, J. F., Campbell, J. S., & Weir, J. M. (2007). Youth psychopathy and criminal recidivism: A meta-analysis of the psychopathy checklist measures. *Law and Human Behavior, 31*, 53–75. doi:10.1007/s10979-006-9019-y

Edens, J. F., Colwell, L. H., Desforges, D. M., & Fernandez, K. (2005). The impact of mental health evidence on support for capital punishment: Are defendants labeled psychopathic considered more deserving of death? *Behavioral Sciences and the Law, 23*, 603–625. doi:10.1002/bsl.660

Edens, J. F., & Cox, J. (2012). Examining the prevalence, role, and impact of evidence regarding antisocial personality, sociopathy, and psychopathy in capital cases: A survey of defense team members. *Behavioral Sciences and the Law, 30*, 239–255. doi:10.1002/bsl.2009

Edens, J. F., Davis, K. M., Fernandez Smith, K., & Guy, L. S. (2013). No sympathy for the devil: Attributing psychopathic traits to capital murderers also predicts support for executing them. *Personality Disorders: Theory, Research, and Treatment, 4*, 175–181. doi:10.1037/a0026442

Edens, J. F., Desforges, D. M., Fernandez, K., & Palac, C. A. (2004). Effects of psychopathy and violence risk testimony on mock juror perceptions of dangerousness in a capital murder trial. *Psychology, Crime and Law, 10*, 393–412. doi; 10.1080/10683160310001629274

Edens, J. F., Guy, L. S., & Fernandez, K. (2003). Psychopathic traits predict attitudes toward a juvenile capital murderer. *Behavioral Sciences and the Law, 21*, 807–828. doi:10.1002/bsl.567

Edens, J. F., Marcus, D. K., Lilienfeld, S. O., & Poythress, N. G. (2006). Psychopathic, not psychopath: Taxometric evidence for the dimensional structure of psychopathy. *Journal of Abnormal Psychology, 115*, 131–144. doi:10.1037/0021-843X.115.1.131

Edens, J. F., Marcus, D. K., & Vaughn, M. G. (2011). Exploring the taxometric status of psychopathy among youthful offenders: Is there a juvenile psychopath taxon? *Law and Human Behavior, 35*, 13. doi:10.1007/s10979-010-9230-8

Edens, J. F., & Petrila, J. (2006). Legal and ethical issues in the assessment and treatment of psychopathy. In C. Patrick (Ed.), *Handbook of psychopathy* (pp. 573–588). New York, NY: Guilford Press.

Edens, J. F., Petrila, J., & Buffington-Vollum, J. K. (2001). Psychopathy and the death penalty: Can the Psychopathy Checklist-Revised identify offenders who represent "a continuing threat to society?". *The Journal of Psychiatry and Law, 29*, 433–481.

Esses, V. M., & Webster, C. D. (1988). Physical attractiveness, dangerousness, and the Canadian Criminal Code. *Journal of Applied Social Psychology, 18*, 1017–1031. doi:10.1111/j.1559-1816.1988.tb01190.x

Federal Rules of Evidence for United States Courts and Magistrates. (1975). St. Paul, MN: West Publishing Company.

Ford v. Wainwright, 477 U.S. 399 (1986).

Forth, A. E., Hart, S. D., & Hare, R. D. (1990). Assessment of psychopathy in male young offenders. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2*, 342–344. doi:10.1037/1040-3590.2.3.342

Forth, A. E., Kosson, D., & Hare, R. D. (2003). *The Hare PCL: Youth version*. Toronto, ON: Multi-Health Systems.

Frick, P. J., & Hare, R. D. (2001). *The antisocial process screening device*. Toronto, ON: Multi-Health Systems.

Furman v. Georgia, 408 U.S. 238 (1972).

Gacono, C. B., Meloy, J. R., Sheppard, K., Speth, S., & Roske, A. (1995). A clinical investigation of malingering and psychopathy in hospitalized insanity acquittees. *Bullet of the American Academy of Psychiatry and the Law, 23*, 387–397.

Gaines v. State, 28 A.3d 706 (Md. Ct. Spec. App. 2011).

Gendreau, P., Goggin, C., & Smith, P. (2002). Is the PCL-R really the "unparalleled" measure of offender risk? *Criminal Justice and Behavior, 29*, 397–426. doi:10.1177/0093854802029004004

Gookin, K. (2007). *Comparison of state laws authorizing involuntary commitment of sexually violent predators: 2006 updated, revised*. (Document No. 07–08-1101). Retrieved from the Washington State Institute for Public Policy website: http://www.wsipp.wa.gov/rptfiles/07-08-1101.pdf

Gregg v. Georgia, 428 U.S. 153 (1976).

Gretton, H. M., McBride, M., Hare, R. D., O'Shaughnessy, R., & Kumka, G. (2001). Psychopathy and recidivism in adolescent sex offenders. *Criminal Justice and Behavior, 28*, 427–449. doi: 10.1177/009385480102800403

Guay, J., Ruscio, J., Knight, R. A., & Hare, R. (2007). A taxometric analysis of the latent structure of psychopathy: Evidence for dimensionality. *Journal of Abnormal Psychology, 116*, 701–716. doi:10.1037/0021-843X.116.4.701

Guy, L. S., & Edens, J. F. (2003). Juror decision-making in a mock sexually violent predator trial: Gender differences in the impact of divergent types of expert testimony. *Behavioral Sciences and the Law, 21*, 215–237. doi:10.1002/bsl.529

Guy, L. S., & Edens, J. F. (2006). Gender differences in attitudes toward psychopathic sexual offenders. *Behavioral Sciences and the Law, 24*, 65–85. doi:10.1002/bsl.665

Guy, L. S., Edens, J. F., Anthony, C., & Douglas, K. S. (2005). Does psychopathy predict institutional misconduct among adults? A meta-analytic investigation. *Journal of Consulting and Clinical Psychology, 73*, 1056–1064. doi:10.1037/0022-006X.73.6.1056

Hall, J. R., & Benning, S. D. (2006). The "successful" psychopath. In C. Patrick (Ed.), *Handbook of psychopathy* (pp. 459–478). New York, NY: Guilford.

Hall, J. R., Benning, S. D., & Patrick, C. J. (2004). Criterion-related validity of the three-factor model of psychopathy: Personality, behavior, and adaptive functioning. *Assessment, 11*, 4–16. doi:10.1177/1073191103261466

Hanson, K. R., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology, 73*, 1154–1163. doi:10.1037/0022-006X.73.6.1154

Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment, 21*, 1–21. doi:10.1037/a0014421

Hare, R. D. (1970). *Psychopathy: Theory and research*. New York, NY: Wiley.

Hare, R. D. (1980). A research scale for the assessment of psychopathy in criminal populations. *Personality and Individual Differences, 1*, 111–120. doi:10.1016/0191-8869(80)90028-8

Hare, R. D. (1991). *The Hare Psychopathy Checklist-Revised*. Toronto, ON: Multi-Health Systems.

Hare, R. D. (1996). Psychopathy: A clinical construct whose time has come. *Criminal Justice and Behavior, 23*, 25–54. doi:10.1177/0093854896023001004

Hare, R. D. (1999). Psychopathy as a risk factor for violence. *Psychiatric Quarterly, 70*, 181–197. doi:10.1023/A:1022094925150

Hare, R. D. (2003). *The Hare Psychopathy Checklist-Revised technical manual* (2nd ed.). Toronto, ON: Multi-Health Systems.

Hare, R. D., Clark, D. C., Grann, M., & Thornton, D. (2000). Psychopathy and the predictive validity of the PCL-R: An international perspective. *Behavioral Sciences and the Law, 18*, 623–645. doi:10.1002/1099-0798(200010)18:5<623:AID-BSL409>3.0.CO;2-W

Hare, R. D., & Hervé, H. F. (1999). *Hare P-Scan. Research version*. New York, NY: Multi-Health Systems Inc.

Hare, R. D., & Neumann, C. S. (2008). Psychopathy as a clinical and empirical construct. *Annual Review of Clinical Psychology, 4*, 217–246. doi:10.1146/annurev.clinpsy.3.022806.091452

Harpur, T. J., & Hare, R. D. (1994). Assessment of psychopathy as a function of age. *Journal of Abnormal Psychology, 103*, 604–609. doi:10.1037/0021-843X.103.4.604

Harpur, T. J., Hakstian, A. R., & Hare, R. D. (1988). Factor structure of the psychopathy checklist. *Journal of Consulting and Clinical Psychology, 56*, 741–747. doi:10.1037/a0028986

Harpur, T. J., Hare, R. D., & Hakstian, R. (1989). A two-factor conceptualization of psychopathy: Construct validity and implications for assessment. *Psychological Assessment, 1*, 6–17. doi:10.1037/1040-3590.1.1.6

Harris, G., & Rice, M. (2006). Treatment of psychopathy: A review of empirical findings. In C. Patrick (Ed.), *Handbook of psychopathy* (pp. 555–572). New York, NY: Guilford.

Hart, S. D., Cox, D. N., & Hare, R. D. (1995). *The Hare psychopathy checklist: Screening version (PCL: SV)*. Toronto, ON: Multi-Health Systems.

Hart, S. D., & Hare, R. D. (1997). Psychopathy: Assessment and association with criminal conduct. In D. M. Stoff, J. Breiling, & J. D. Maser (Eds.), *Handbook of Antisocial Behavior* (pp. 22–35). New York, NY: Wiley.

Hawes, S. M., Boccaccini, M. T., & Murrie, D. C. (2013). Psychopathy and the combination of psychopathy and sexual deviance as predictors of sexual recidivism: Meta-analytic findings using the Psychopathy Checklist-Revised. *Psychological Assessment, 25*, 233–243. doi:10.1037/a0030391

Heilbrun, K., Leheny, C., Thomas, L., & Huneycutt, D. (1997). A national survey of U.S. statutes on juvenile transfer: Implications for policy and practice. *Behavioral Science and the Law, 15*, 125–149. doi:10.1002/(SICI)1099-0798(199721)15:2<125:AID-BSL265>3.0.CO;2-R

Heinze, M. C., & Vess, J. (2005). The relationship among malingering, psychopathy, and the MMPI-2 validity scales in maximum security forensic psychiatric inpatients. *Journal of Forensic Psychology Practice, 5*, 35–53. doi:10.1300/J158v05n03_02

Hemphill, J. F., & Hare, R. D. (2004). Some Misconceptions about the Hare PCL-R and risk assessment: A reply to Gendreau, Goggin, and Smith. *Criminal Justice and Behavior, 31*, 203–243. doi:10.1177/0093854803261326

Hemphill, J. F., Hare, R. D., & Wong, S. (1998). Psychopathy and recidivism: A review. *Legal and Criminological Psychology, 3*, 139–170. doi:10.1111/j.2044-8333.1998.tb00355.x

Hildebrand, M., de Ruiter, C., & de Vogel, V. (2004). Psychopathy and sexual deviance in treated rapists: Association with sexual and nonsexual recidivism. *Sexual Abuse: A Journal of Research and Treatment, 16*, 1–24. doi:10.1023/B:SEBU.0000006281.93245.de

Hill, M. F. (1996). Applicability of rules of evidence to juvenile transfer, waiver, or certification hearings. *American Law Reports (5th), 37*, 703.

Imwinkelried, E. J. (1988). Meaning of probative value and prejudice in Federal Rule of Evidence 403: Can Rule 403 be used to resurrect the common law of evidence? *Vanderbilt Law Review, 41*, 879–907.

Jackson, R. L. (2008). Sex offender civil commitment: Recommendations for empirically guided evaluations. *The Journal of Psychiatry and Law, 36*, 389–429.

Jackson, R. L., & Hess, D. (2007). Evaluation for civil commitment of sex offenders: A survey of experts. *Sexual Abuse: A Journal of Research and Treatment, 19*, 425–448. doi:10.1177/107906320701900407

Jackson, R. L., Rogers, R., & Shuman, D. W. (2004). The adequacy and accuracy of sexually violent predator evaluations: Contextualized risk assessment in clinical practice. *International Journal of Forensic Mental Health, 3*, 115–129. doi:10.1080/14999013.2004.10471201

Janus, E. S., & Prentky, R. A. (2008). Sexual predator laws: A two-decade retrospective. *Federal Sentencing Reporter, 21*, 90–97. doi:10.1525/fsr.2008.21.2.90

Joint Legislative Audit and Review Commission (2011). *Review of the civil commitment of sexually violent predators: Report to the Assembly of Virginia*. Retrieved from http://jlarc.virginia.gov/meetings/November11/SVP.pdf

Jones, S., & Cauffman, E. (2008). Juvenile psychopathy and judicial decision making: An empirical analysis of an ethical dilemma. *Behavioral Sciences and the Law, 26*, 151–165. doi:10.1002/bsl.792

Kansas v. Hendricks, 521 U.S. 346 (1997).

Kansas v. Crane, 534 U.S. 407 (2002).

Kiehl, K. A., & Hoffman, M. B. (2011). The criminal psychopath: History, neuroscience, treatment, and economics. *Jurimetrics, 51*, 355–397.

Kennedy v. Louisiana, 554 U.S. 407 (2008).

Kent v. United States, 383 U.S. 541 (1966).

Knight, R. A., & Thornton, D. (2007). *Evaluating and improving risk assessment schemes for sexual recidivism: A long-term follow-up of convicted sexual offenders* (Document No. NCJ 217618). Retrieved from https://www.ncjrs.gov/pdffiles1/nij/grants/217618.pdf

Krauss, D. A., McCabe, J. G., & Lieberman, J. D. (2011). Dangerously misunderstood: Representative jurors' reactions to expert testimony on future dangerousness in a sexually violent predator trial. *Psychology, Public Policy, and Law, 18*, 18–49. doi:10.1037/a0024550

Kucharski, L. T., Duncan, S., Egan, M. A., & Falkenbach, D. M. (2006). Psychopathy and malingering of psychiatric disorder in criminal defendants. *Behavioral Sciences and the Law, 24*, 633–644. doi:10.1002/bsl.661

La. C.Cr.P. art. 905.2. (1976).

La. Code Ev. 1101. (2011).

Langton, C. M. (2003). *Contrasting approaches to risk assessment with adult male sexual offenders: An evaluation of recidivism prediction schemes and the utility of supplementary clinical information for enhancing predictive accuracy*. Unpublished doctoral dissertation. University of Toronto.

Langton, C. M., Barbaree, H. E., Harkins, L., & Peacock, E. J. (2006). Sex offenders' response to treatment and its association with recidivism as a function of psychopathy. *Sexual Abuse: A Journal of Research and Treatment, 18*, 99–120. doi:10.1177/107906320601800107

Leistico, A. M. R., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the hare measures of psychopathy to antisocial conduct. *Law and Human Behavior, 32*, 28–45. doi:10.1007/s10979-007-9096-6

Lieberman, J. D., & Krauss, D. A. (2009). The effects of labeling, expert testimony, and information processing mode on juror decisions in SVP civil commitment trials. *Journal of Investigative Psychology and Offender Profiling, 6*, 25–41. doi:10.1002/jip.91

Lockett v. Ohio, 438 U.S. 586 (1978).

Lockhart v. McCree, 476 U.S. 162 (1986).

Lynam, D. R. (1997). Pursuing the psychopath: Capturing the fledgling psychopath in a nomological net. *Journal of Abnormal Psychology, 106*, 425–438. doi:10.1037//0021-843X.106.3.425

Lynam, D. R., & Derefinko, K. J. (2006). Psychopathy and personality. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 133–155). New York, NY: Guilford Press.

McCabe, J., Krauss, D., & Lieberman, J. (2010). Reality check: A comparison of college students and community samples of mock jurors in a simulated sexual violent predator civil commitment. *Behavioral Sciences and the Law, 28*, 730–750. doi:10.1002/bsl.902

McCord, W. & McCord, J. (1964). *The psychopath: An essay on the criminal mind.* Princeton, NJ: Van Nostrand.

Miller, P. A., & Eisenberg, N. (1988). The relation of empathy to aggressive and externalizing/antisocial behavior. *Psychological Bulletin, 103*, 324–344. doi:10.1037/0033-2909.103.3.324

Monahan, J., Steadman, H., Silver, E., Appelbaum, P. S., Robbins, P. C., Mulvey, E. P., et al. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York, NY: Oxford University Press.

Morey, L. C. (1991). *Personality assessment inventory: Professional manual*. Tampa, FL: Psychological Assessment Resources.

Murrie, D. C., Boccaccini, M. T., McCoy, W., & Cornell, D. G. (2007). Diagnostic labeling in juvenile court: How do descriptions of psychopathy and conduct disorder influence judges? *Journal of Clinical Child and Adolescent Psychology, 36*, 228–241. doi:10.1080/15374410701279602

Murrie, D. C., Cornell, D. G., & McCoy, W. K. (2005). Psychopathy, conduct disorder, and stigma: Does diagnostic labeling influence juvenile probation officer recommendations? *Law and Human Behavior, 29*, 323–342. doi:10.1007/s10979-005-2415-x

Neumann, C. S., Kosson, D. S., & Salekin, R. T. (2007). Exploratory and confirmatory factor analysis of the psychopathy construct: Methodological and conceptual issues. In H. Hervé & J. C. Yuille (Eds.), *The psychopath: Theory, research, and practice* (pp. 79–104). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Office of Juvenile Justice and Delinquency Prevention, U.S. Department of Justice. (2010). *Juvenile transfer laws: An effective deterrent to delinquency?* Retrieved from https://www.ncjrs.gov/pdffiles1/ojjdp/220595.pdf.

Ogloff, J. R., Wong, S., & Greenwood, A. (1990). Treating criminal psychopaths in a therapeutic community program. *Behavioral Sciences and the Law, 8*, 181–190. doi:10.1002/bsl.2370080210

O'Neil, M. L., Lidz, V., & Heilbrun, K. (2003). Adolescents with psychopathic characteristics in a substance abusing cohort: Treatment process and outcomes. *Law and Human Behavior, 27*, 299–313.

Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2009). Risk assessment with young offenders: A meta-analysis of three assessment measures. *Criminal Justice and Behavior, 36*, 329–353. doi:10.1177/0093854809331457

Olver, M. E., & Wong, S. C. P. (2006). Psychopathy, sexual deviance, and recidivism among sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 18*, 65–82. doi:10.1177/107906320601800105

Panetti v. Quarterman, 551 U.S. 930 (2007).

Parent, G., Guay, J., & Knight, R. A. (2011). An assessment of long-term risk of recidivism by adult sex offenders: One size does not fit all. *Criminal Justice and Behavior, 38*, 188–209. doi:10.1177/0093854810388238

Paulhus, D. L., Neumann, C. S., & Hare, R. D. (in press). *Self-report psychopathy scale: Version IV*. Toronto, ON: Multi-Health Systems.

Petrila, J., & Otto, R. K. (2001). Admissibility of expert testimony in sexually violent predator proceedings. In A. Schlank (Ed.), *The sexual predator: Legal issues, clinical issues, and special populations* (Vol. 2, pp. 3:1–3:25). Kingston, NJ: Civic Research Press.

Porter, S., Birt, A. R., & Boer, D. P. (2001). Investigation of the criminal and conditional release profiles of Canadian federal offenders as a function of psychopathy and age. *Law and Human Behavior, 25*, 647–661.

Porter, S., Fairweather, D., Drugge, J., Herv´e, H., Birt, A., & Boer, D. P. (2000). Profiles of psychopathy in incarcerated sexual offenders. *Criminal Justice and Behavior, 27*, 216–233. doi:10.1177/0093854810388238

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, D.C.: American Psychological Association.

Quinsey, V. L., Rice, M. E., & Harris, G. T. (1995). Actuarial prediction of sexual recidivism. *Journal of Interpersonal Violence, 10*, 85–105. doi:10.1177/0093854810388238

Rendell, J. A., Huss, M. T., & Jensen, M. L. (2010). Expert testimony and the effects of a biological approach, psychopathy, and juror attitudes in cases of insanity. *Behavioral Sciences and the Law, 28*, 411–425. doi:10.1002/bsl.913

Rice, M. E. (1997). Violent offender research and implications for the criminal justice system. *American Psychologist, 52*, 414–423. doi:10.1037/0003-066X.52.4.414

Rice, M. E., & Harris, G. T. (1997). Cross-validation and extension of the violence risk appraisal guide for child molesters and rapists. *Law and Human Behavior, 21*, 231–241.

Rice, M. E., Harris, G. T., & Quinsey, V. L. (1990). A follow-up of rapists assessed in a maximum-security psychiatric facility. *Journal of Interpersonal Violence, 5*, 435–448. doi:10.1177/088626090005004001

Ring v. Arizona, 536 U.S. 584 (2002).

Rockett, J. L., Murrie, D. C., & Boccaccini, M. T. (2007). Diagnostic labeling in juvenile justice settings: Do psychopathy and conduct disorder findings influence clinicians? *Psychological Services, 4*, 107–122. doi:10.1037/1541-1559.4.2.107

Roberts v. Louisiana, 428 U.S. 153 (1976).

Rogers, R., Gillis, J. R., Dickens, S. E., & Bagby, R. M. (1991). Standardized assessment of malingering: Validation of the Structured Interview of Reported Symptoms. *Psychological Assessment, 3*, 89–96. doi:10.1037/1040-3590.3.1.89

Roper v. Simmons, 543 U.S. 551 (2005).

Salekin, R. T., Rogers, R., & Sewell, K. W. (1996). A review and meta-analysis of the psychopathy checklist and Psychopathy Checklist-Revised: Predictive validity of dangerousness. *Clinical Psychology: Science and Practice, 3*, 203–215. doi:10.1111/j.1468-2850.1996.tb00071.x

Schneider, K. (1958). *Psychopathic personalities*. London, UK: Cassell.

Serin, R. C., Mailloux, P. B., & Malcolm, P. B. (2001). Psychopathy, deviant sexual arousal, and recidivism among sexual offenders. *Journal of Interpersonal Violence, 16*, 234–246. doi:10.1177/088626001016003004

Seto, M. C., & Barbaree, H. E. (1999). Psychopathy, treatment behavior, and sex offender recidivism. *Journal of Interpersonal Violence, 14*, 1235–1248. doi:10.1177/088626099014012001

Sevilla, C. M. (1999). Antisocial personality disorder: Justification for the death penalty? *Journal of Contemporary Legal Issues, 10*, 247–262.

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review, 31*, 499–513. doi:10.1016/j.cpr.2010.11.009

Sjöstedt, G., & Långström, N. (2002). Assessment of risk for criminal recidivism among rapists: A comparison of four different measures. *Psychology, Crime and Law, 8*, 25–40. doi:10.1177/0886260509336961

Skeem, J. L., Monahan, J., & Mulvey, E. P. (2002). Psychopathy, treatment involvement, and subsequent violence among civil psychiatric patients. *Law and Human Behavior, 26*, 577–603.

Skeem, J. L., & Mulvey, E. P. (2001). Psychopathy and community violence among civil psychiatric patients: Results from the MacArthur violence risk assessment study. *Journal of Consulting and Clinical Psychology, 69*, 358–374. doi:10.1037//0022-006X.69.3.358

Sorensen, J. R., & Cunningham, M. D. (2007). Operationalizing risk: The influence of measurement choice on the prevalence and correlates of prison violence among incarcerated murderers. *Journal of Criminal Justice, 35*, 546–555. doi:10.1016/j.jcrimjus.2007.07.007

Sorensen, J., & Cunningham, M. D. (2010). Conviction offense and prison violence a comparative study of murderers and other offenders. *Crime and Delinquency, 56*, 103–125. doi:10.1177/0011128707307175

Sorensen, J. R., & Pilgrim, R. L. (2000). An actuarial risk assessment of violence posed by capital murder defendants. *Journal of Criminal Law and Criminology, 9*, 1251–1270.

Spain, S. E., Douglas, K. S., Poythress, N. G., & Epstein, M. (2004). The relationship between psychopathic features, violence and treatment outcome: The comparison of three youth measures of psychopathic features. *Behavioral Sciences and the Law, 22*, 85–102. doi:10.1002/bsl.576

State v. Kleser, 786 N.W.2d 144 (Wis. 2010).

Tate, D., & Redding, R. (2005). Mental health and rehabilitative services in juvenile justice: System reforms and innovative approaches. In K. Heilbrun, N. Goldstein, & R. Redding (Eds.), *Juvenile delinquency: Prevention, assessment, and intervention* (pp. 134–160). New York, NY: Oxford University Press.

Tennard v. Dretke, 542 U.S. 274 (2004).

United States v. Barnette, 211 F .3d 803 (4th Cir. 2000).

United States v. Comstock, 560 U.S. 126 (2010).

United States v. Weed, 389 F.3d 1060 (10th Cir. 2004).

Vess, J., Murphy, C., & Arkowitz, S. (2004). Clinical and demographic differences between sexually violent predators and other commitment types in a state forensic hospital. *Journal of Forensic Psychiatry and Psychology, 15*, 669–681. doi:10.1177/0306624X06287644

Vidal, S., & Skeem, J. L. (2007). Effect of psychopathy, abuse, and ethnicity on juvenile probation officers' decision-making and supervision strategies. *Law and Human Behavior, 31*, 479–498. doi:10.1007/s10979-006-9077-1

Vitacco, M. J., Rogers, R., Neumann, C. S., Harrison, K. S., & Vincent, G. (2005). A comparison of factor models on the PCL-R with mentally disordered offenders: The development of a four-factor model. *Criminal Justice and Behavior, 32*, 526–545. doi:10.1177/0093854805278414

Vitale, J. E., & Newman, J. P. (2001). Using the Psychopathy Checklist-Revised with female samples: Reliability, validity, and implications for clinical utility. *Clinical Psychology: Science and Practice, 8*, 117–132. doi:10.1093/clipsy.8.1.117

Vitale, J. E., Smith, S. S., Brinkley, C. A., & Newman, J. P. (2002). The reliability and validity of the Psychopathy Checklist-Revised in a sample of female offenders. *Criminal Justice and Behavior, 29*, 202–231. doi:10.1177/0093854802029002005

Walsh, T., & Walsh, Z. (2006). The evidentiary introduction of Psychopathy Checklist-Revised assessed psychopathy in U.S. courts: Extent and appropriateness. *Law and Human Behavior, 30*, 493–507. doi:10.1007/s10979-006-9042-z

Walters, G. D. (2003a). Predicting criminal justice outcomes with the psychopathy checklist and lifestyle criminality screening form: A meta-analytic comparison. *Behavioral Sciences and the Law, 21*, 89–102. doi:10.1002/bsl.519

Walters, G. D. (2003b). Predicting institutional adjustment and recidivism with the psychopathy checklist factor scores: A meta-analysis. *Law and Human Behavior, 27*, 541–558.

Walters, G. D., Duncan, S. A., & Mitchell-Perez, K. (2007). The latent structure of psychopathy: A taxometric analysis of the Psychopathy Checklist-Revised in a heterogeneous sample of male prison inmates. *Assessment, 14*, 270–278. doi:10.1177/1073191107299594

Walters, G. D., White, T. W., & Denney, D. (1991). The lifestyle criminality screening form: Preliminary data. *Criminal Justice and Behavior*, *18*, 406–418. doi:10.1177/0093854891018004003

Weinstein, J. B., & Berger, M. A. (2015). *Weinstein's evidence manual: A guide to the Federal Rules of Evidence based on Weinstein's Federal Evidence* (10th ed.). Danvers, MA: Matthew Bender & Company Inc.

Yang, M., Wong, S. C., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin, 136*, 740–767. doi:10.1037/a0020473

A Synthetic Perspective on the Own-Race Bias in Eyewitness Identification

John Paul Wilson, Michael J. Bernstein and Kurt Hugenberg

In August 1997, a White woman in Louisiana suffered a physical attack and attempted rape at the hands of a Black stranger. Before the end of the night, police had apprehended a suspect, a man named Nathan Brown. The victim identified him as her attacker. Although Brown had four witnesses to corroborate his alibi, he was convicted and sentenced to 25 years in prison almost exclusively on the basis of the victim's eyewitness testimony. Brown served 17 of these 25 years before attorneys with the Innocence Project demonstrated that the DNA profile on the victim's garments was not a match to Brown (Cates and Trigg 2014). Brown was exonerated in 2014, and the actual perpetrator was soon found. Brown's experience of unjust incarceration unfortunately is only one of the most recent in a long line of people who have been wrongfully convicted due to faulty eyewitness testimony. Since DNA testing was first used in post-conviction appeals, over 320 DNA-related exonerations have occurred, and over 72 % of these involved eyewitness misidentification. Further, a disproportionate number of misidentifications involved a victim of a different race than the perpetrator—an *other-race* face misidentification (Scheck et al. 2000). Such cases of eyewitness misidentification have created a great deal of interest in how to improve eyewitness memory, and how to reduce errors in eyewitness identification. Although the need for new research remains

This work was supported by NSF grants BCS 0951463 and 1423765, awarded to the third author.

J.P. Wilson (✉)
Montclair State University, Montclair, NJ, USA
e-mail: johnpaulw@gmail.com

M.J. Bernstein
Penn State Abington, Abington, PA, USA

K. Hugenberg (✉)
Miami University, Oxford, OH, USA
e-mail: hugenbk@muohio.edu

B.H. Bornstein and M.K. Miller (eds.), *Advances in Psychology and Law*,
Advances in Psychology and Law 2, DOI 10.1007/978-3-319-43083-6_8

great, we now know a great deal about the psychology of eyewitness memory. In this chapter, we will focus on one specific aspect of eyewitness memory, the own-race bias (ORB) in face memory, or the tendency for perceivers to have greater difficulty in making other-race face identifications than own-race face identifications.

Humans are quite adept at processing and remembering individual faces (Kanwisher et al. 1997), especially when those faces have become quite familiar. For example, Bahrick et al. (1975) demonstrated that participants even 35 years after high school could recognize as many as 90 % of their high school classmates, and could recognize 73 % at 40 years postgraduation. More than any other stimulus, we are skilled at encoding the perceptual configurations of faces and binding them to individual identities (Haxby et al. 2000), in part because humans appear to have specialized neural structures dedicated to processing faces (e.g., Rangarajan et al. 2014). Despite this great facility with faces, we often fail to recognize the faces of strangers or confuse novel faces with those of people we have encountered only briefly before (Duchaine and Nakayama 2006; Hancock et al. 2000). In the domain of eyewitness identification, this means that few witnesses would mistake the identity of a friend, co-worker, or lover; however, witnesses of crimes that occur quickly or that are perpetrated by strangers are subject to commonplace errors and strong biases in recognition.

The fallibility of face memory quite obviously has severe repercussions for the criminal justice system. Eyewitnesses who fail to accurately remember the facial appearance of suspects are at risk of contributing to the wrongful imprisonment of innocent people, as well as failing to identify the actual perpetrators of crimes. Even more troublesome than the general inaccuracy of face memory is the particularly high level of inaccuracy that is reliably observed for suspects *who are not the same-race as witnesses* (Sporer 2001). This ORB in face recognition represents a substantial problem for the legal system.

This substantial problem stems, to a great extent, from the fact that eyewitnesses wield a great degree of influence over criminal justice decisions (e.g., Loftus 1974). Jurors tend to believe personal testimony from people who can point to suspects on the basis of their own experience (Semmler et al. 2012). Unfortunately, recent evidence has demonstrated that eyewitness errors, and especially other-race eyewitness errors, occur with troubling frequency. In the present writing, we hope to represent a new perspective on the ORB in face memory and the ways in which its influence can be reduced in eyewitness identification.

The Own-Race Bias and Eyewitness Identification

Past research has demonstrated two effects of race on face memory, both of which have been referred to as an ORB (see Meissner and Brigham 2001). First, the classic definition of the ORB is a tendency for perceivers to have greater *sensitivity* in memory for own-race than other-race faces. In essence, perceivers are better at

discriminating between faces they have seen before and faces they have not seen before, if the target is of their own-race. A second effect of race on face memory is that perceivers tend to have a more stringent *criterion* for reporting they have seen an own-race, relative to an other-race face. In short, perceivers favor "errors of commission" for other-race faces (saying they have seen the target when really they have not), but "errors of omission" for own-race faces (saying they have not seen an own-race face when they really have). Both findings can have important implications for eyewitness identification, but in different ways. The effects of sensitivity are straightforward: improving sensitivity will improve the veracity of eyewitness testimony. However, changing the criterion effects involve a trade-off. Making a criterion more liberal (i.e., saying "yes, that's the guy" more often) will lead to more convictions of *both* innocent and guilty parties; conversely making a criterion more conservative will have the opposite effect of fewer convictions of both innocent and guilty parties.

In this work, we focus primarily on sensitivity effects (except as noted), for two reasons. First, there has been more research examining how race influences sensitivity (which we refer to as the ORB), and is the primary focus of many major models of the ORB. Second, and as noted above, whereas improving sensitivity is an unmitigated good (improved sensitivity means perceivers can better distinguish between the innocent and actual perpetrators), changing criterion is a difficult trade-off that is almost certainly fraught with judgments about the value of protecting the wrongly accused versus incarcerating the guilty.

For several decades, researchers have been interested in the ORB—the tendency to have better memory for own-race relative to other-race faces. For example, Malpass and Kravitz (1969) published a now-classic paper showing that White participants showed superior recognition memory for White target photos relative to Black target photos. Since that paper was published, hundreds of samples representing thousands of participants have shown this ORB to be robust across populations, stimuli, and research context (Meissner and Brigham 2001). Although much research in this area has been conducted in laboratory settings, the ORB is of obvious relevance to eyewitness memory, and it has been established in more ecologically valid field studies as well (see Sporer 2001). Despite this high level of relevance to criminal justice outcomes, recent advances in the science of the ORB have yet to influence policies for eyewitness lineups and testimony. In this review, we turn first to describing the ORB itself, and synthesize the current state of the scientific research on this recognition bias. We then turn to describing how one might consider scaling up practical interventions, drawn from the research literature, that have been demonstrated to reduce the ORB.

Notably, we will only provide a limited treatment of the state of research regarding eyewitness lineup construction and procedures for interacting with witnesses in lineup situations. Creating unbiased lineups and avoiding biasing witnesses, of course, is of critical importance to accurate eyewitness identification. Relevant to this, recent research has demonstrated that other-race sensitivity cannot be improved by simple changes to lineup procedures. Instead, the ORB in sensitivity is in large part a symptom of how faces are processed *at encoding*

(Meissner et al. 2005a)—that is, how faces are processed during and immediately after the crime was witnessed. Interventions at retrieval can indeed shift the response criterion (e.g., Meissner et al. 2005a, b), but improving the discriminability of faces seems to be an encoding phenomenon (e.g., Young et al. 2010). Thus, improving accuracy for other-race face recognition is most likely to occur with interventions undertaken *before* encoding. After the eyewitness event, one's chances for actually improving the accuracy of memory for the target are quite small. At this point, the best that can be expected is to employ best practices in de-biasing lineup construction to reduce false identifications in a general sense. Although the fact that de-biasing attempts post-encoding are ineffective at reducing the ORB may seem to paint a bleak picture, there are indeed promising interventions that could promote the reduction of other-race misidentifications. These interventions, implemented before face encoding, may result in a meaningful reduction in the tendency to conflate innocent foils with the perpetrators of crimes.

The ORB across Paradigms

We first wish to briefly address the methods commonly used for studying the ORB. Experimental procedures in this area typically fall into one of two general categories: serial face presentation in a recognition task and the eyewitness lineup paradigm. Some researchers favor the latter, as it arguably provides more ecological validity—a better match between the experiment and the real world—with regard to eyewitness situations (Lindsay and Wells 1983). In the former, more commonly used serial presentation task, participants are merely asked to view a series of faces in an encoding phase, and then subsequently to undergo a test phase in which they view faces that they saw previously, intermixed with novel faces, and to indicate whether each face is old or new. In the latter, participants witness an event and are later asked to pick the perpetrator from a lineup. Importantly, both types of methods have reliably been associated with the ORB (Meissner and Brigham 2001). Put simply, both seem to tap into the same basic cognitive processes that make own-race faces easier to recognize than other-race faces. Importantly, because both procedures appear to generate an identical ORB, many ORB researchers favor the serial presentation because of its additional advantages. To wit, the serial presentation allows great precision in calculating indices of both recognition sensitivity (sensitivity to whether a target was seen or not at encoding) and response bias (overall tendency to indicate whether targets were seen or not at encoding) in terms of signal detection (Green and Swets 1966; MacMillan and Creelman 2004). In other words, each target in the testing phase was either seen or not at encoding, and all possible responses are easily categorized as either correct identifications of a person who was previously seen (hits), incorrect identifications of a person who was not previously seen (false alarms), failures to identify a person who was previously seen (misses), and correctly indicating that a person was not seen (correct rejections). Furthermore, this paradigm tends to allow for the greater

experimental control that is necessary to develop a well-specified model of the ORB. Notably, some recent research has employed a "middle road" approach, combining characteristics of both serial and lineup procedures (Lane and Meissner 2008). For example, Meissner et al. (2005b) developed a lineup paradigm that gives participants multiple opportunities to identify suspects in target-present and target-absent arrays while providing signal detection data that can inform face memory theory (see also Evans et al. 2009; Jackiw et al. 2008). Moving forward, such an approach will be fruitful for considering how to integrate basic and applied research in order to increase eyewitness accuracy.

Two Primary Causes of the ORB

Since the publication of the seminal work by Malpass and Kravitz (1969), researchers have debated the causal origins of the ORB. Although many people appear to spontaneously believe the ORB is a manifestation of racism—antipathy toward a racial outgroup – in actuality this does not appear to be the case (Meissner and Brigham 2001). Instead, researchers have proposed two primary causes of the ORB: first, many people lack *experience* with other-race faces, and second people tend to think of others in terms of social *categories* rather than as individuals, attending to what makes people similar rather than what makes them unique individuals. Each of these causes, by itself, can reduce outgroup face recognition. Considered together, they create a circumstance in which other-race face recognition is severely disadvantaged (Hugenberg et al. 2013; Young and Hugenberg 2012). These two causes combine to produce effect sizes for the other-race effect that are reliably larger than for other classes of "Own-Group Biases" (such as those based on arbitrarily assigned groups, e.g., Bernstein et al. 2007), and as such, result in the troubling race-based biases we see play out in other-race eyewitness identification (Hugenberg et al. 2013).

An experience deficit. It has long been argued that the ORB is rooted in differential experience with own-race and other-race faces (Diamond and Carey 1986; Malpass et al. 1973). This *experience hypothesis* argues that the ORB begins with de facto racial segregation—even multicultural societies such as the U.S. and the UK are quite racially segregated (Goldsmith and Blakely 2010), and many countries are highly racially homogenous (Wan et al. 2015). According to the experience-based perspective, this pattern of racial segregation leads most people to have more contact, often *much* more contact, with own-race individuals than with other-race individuals (MacLin and Malpass 2001; Meissner and Brigham 2001). This differential contact is then argued to lead to differential experience in processing own-race and other-race faces—people have more practice recognizing same-race relative to other-race faces—and this differential experience gives rise to greater recognition accuracy for own-race faces.

That experience would lead to better recognition of own-race faces seems straightforward, but exactly how practice with own-race faces results in the ORB is

not a settled matter. Some argue that experience elicits a superior process—configural processing—that allows people to efficiently extract information from own-race but not from other-race faces (e.g., Caharel et al. 2011; Rhodes et al. 1989; Tong et al. 2014). According to this perspective, by extracting information about the relationships among facial characteristics (e.g., the location of the eyes relative to the nose and mouth), rather than attending to a single feature at a time, perceivers are better able to efficiently extract identity information from own-race faces. However, insofar as perceivers have less experience with other-race faces, they are less able to engage in this highly efficient configural processing, and therefore have weaker memories for other-race faces (Hancock and Rhodes 2008).

One common paradigm for investigating these differences in configural processing is using a face inversion task. To the extent that own-race faces are processed more configurally than are other-race faces, own-race face processing should be more impacted by presenting the face upside-down (because inversion disrupts the typical eyes-over-nose-over-mouth configuration of faces). A number of studies using this inversion paradigm have shown that own-race faces are processed more configurally than are other-race faces (Michel et al. 2006; Rhodes et al. 1989; Sangrigoli and de Schonen 2004). Moreover, Hancock and Rhodes (2008) found that greater self-reported contact with racial outgroup members was associated both with a reduction in the ORB and an increase in configural coding of other-race faces. Importantly, the other-race difference in configural coding was a significant predictor of the ORB in recognition memory. However, other recent research has demonstrated that people actually process both own- *and* other-race faces configurally (Bukach et al. 2012). Related to this finding, own- versus other-race differences in face memory may be reflected in quantitative, rather than qualitative, differences in processing. Harrison et al. (2014) found that participants processed both ingroup and outgroup faces configurally but still showed the ORB in recognition (Harrison et al. 2014). Taken together, there is some recent evidence that calls into question whether configural processing is a primary cause of the ORB; however, many leading models do conclude that configural processing plays a key role in generating this bias (Hancock and Rhodes 2008; Rossion and Michel 2011).

Other theorists (e.g., Valentine 1991; Valentine and Endo 1992), however, argue that the differential own-race versus other-race recognition is a matter of representation in memory rather than of processing style. These representational models argue that by having an extensive number of own-race faces represented in memory, one's memory becomes biased such that own-race faces are more easily discriminated from one another. Faces are thought to be represented in a multidimensional face space. Although there are different face space accounts of the ORB, these models each rely on the density of face exemplars that are psychologically represented in one's face space, with denser exemplar representations for own—than other-race faces (Valentine 1991; Nosofsky 1986). These representational accounts are distinct from processing accounts, but both rely on the idea that perceivers have more experience with own-race faces than other-race faces.

These different perspectives underlying the experience hypothesis are quite distinct from one another in important ways, but they share a central prediction

about the ORB: more experience with other-race faces should improve other-race memory. Initial evidence in favor of this hypothesis appeared quite mixed. Indeed, a number of classic studies have found no relationship between past experience with other-race faces and the magnitude of the ORB (e.g., Malpass and Kravitz 1969; Ng and Lindsay 1994). More concerning still was Meissner and Brigham's (2001) meta-analysis that analyzed the relationship between other-race exposure and the ORB across nearly 5000 participants, finding that past experience with other-race faces only accounts for a small portion of the variance in face memory. Finally, in some recent work, the null relationship between past experience and the ORB has been accompanied by a criterion shift, such that more experience predicts more "yes" responses in lineup presentations, without a corresponding increase in discrimination accuracy (Wylie et al. 2015). Taken together, a number of studies have called into question the putative relationship between differential expertise and the ORB.

A number of recent studies, however, do appear to support the case that differential experience with own-race and other-race faces can play a role in accounting for the ORB, and in some cases may be a primary cause for the ORB (e.g., Wan et al. 2015). For example, manipulating experience with other-race faces can affect the ORB; specifically, giving perceivers practice at perceptual discrimination can reduce the magnitude of the ORB, at least temporarily (Malpass et al. 1973). Furthermore, lifelong training with other-race faces can actually reverse the direction of the ORB. Sangrigoli et al. (2005) studied ethnically Korean perceivers who had been adopted by and grown up with White European families. These ethnically Korean (but culturally European) participants showed a reversal of the ORB as adults, such that they were better at recognizing White than Asian faces. In fact, a number of theorists argue that there is a critical period in infancy in which babies' attention attunes to facial features commonly experienced in their environment. After this period of development, other-race faces (faces that presumably do not benefit from common exposure and experience) will become less discriminable (e.g., Kelly et al. 2007). Despite the recent evidence supporting this experience hypothesis, it is worth noting that even in samples in which contact is predictive of the ORB, it is difficult to conceptually separate intergroup contact from other confounding factors such as motivation to individuate, or desire for intergroup contact. Because of this, merely exposing perceivers to other-race faces does not appear to be a sufficient solution for fully eliminating the ORB in eyewitness memory.

We argue that the existing evidence suggests that contact and experience alone cannot be the only components of a strategy to reduce the ORB in eyewitness identification. However, we do argue that effective strategies will involve perceptual experience with racial outgroups. Critically, though, this experience must be paired with motivation to process those racial outgroup members as *individuals*. Recent research has shown that experience and motivation act conjointly to reduce the ORB (Young and Hugenberg 2012). Often, perceivers do have enough experience to show relatively high accuracy for faces of various races (Hugenberg et al. 2007). Indeed, it may often be somewhat more difficult to accurately encode

other-race faces, but with sufficient motivation and effort, one can do so. This means that an additional key question is whether perceivers have the *awareness, motivation, and opportunity* to sufficiently encode those other-race faces. In fact, social categorization is often at the root of failures to use the other-race experience that perceivers actually do possess.

Categorical thinking. More recent research has begun to expand our knowledge beyond what we know from the experience deficit outlined above. This approach begins with what we know from decades of social cognitive research: people rely on categories when thinking about other people. This tendency to think categorically about others is automatic, pervasive, and occurs spontaneously in most contexts (Macrae and Bodenhausen 2000; see also Freeman and Ambady 2011). Upon seeing someone, perceivers quickly and efficiently categorize them according to salient characteristics such as age, sex, and race (e.g., Brewer 1988; Ito and Urland 2003; Macrae et al. 2005; Mason et al. 2006). Further, category information is very often activated when perceiving faces in naturalistic contexts. Although there are exceptions (Macrae et al. 1997, 1999; Quinn et al. 2009, 2010), category activation occurs in nearly all contexts in which face memory is relevant—and almost certainly in the case of eyewitness memory. In fact, people are very *un*likely to make memory errors across the lines of categories (i.e., people tend not to misremember a White person as Black, or a woman as a man), demonstrating how powerful and prevalent these categories are in person memory (e.g., Stangor et al. 1992; Taylor et al. 1978).

Although this tendency to use social categories can be beneficial to perceivers in that it simplifies person construal, it can be damaging when individual identities must be remembered. It is only when perceivers are motivated to think carefully about others that they tend to move past categories to look for *individuating* information—that is, it takes some effort to attend to what discriminates an unfamiliar individual from others in their group (Brewer 1988; Fiske and Neuberg 1990; Hugenberg et al. 2010). One effect of such categorization is that it can essentially make faces that belong to the same category seem more similar to one another, or to otherwise "blend together." Rather than being encoded as individuals, people are encoded as category members. Importantly, stronger category activation can lead to weaker recognition (Susa et al. 2010; Wolff et al. 2014; Young et al. 2009).

A number of studies indicate that social categories can influence how faces are processed and recognized. In one series of studies, Corneille et al. (2004) investigated memory distortions toward racial prototypes. Using faces morphed along continua from White to North African and White to Asian, they found that memory for previously seen faces was distorted toward the ethnic prototype. Participants who saw a face at encoding that was 70 % North African/30 % White were more likely at recognition to select an 80 or 90 % North African face as the target face, rather than the correct 70/30 face. This suggests that perceivers misremembered faces as being more racially prototypical than they actually were. These distortions did not occur for highly ambiguous (50/50) faces, perhaps because such distortions only happen when there is less ambiguity about the target category. In short, memory for faces was biased toward the category that was activated during encoding.

Other research has investigated these memory distortions in a recognition paradigm like those used in most ORB studies. MacLin and Malpass (2001) presented Hispanic participants with images of ethnically ambiguous Hispanic/Black faces. These authors introduced a critical manipulation, however, to induce participants to categorize the faces, they added hairstyles to each face that were either typical of Hispanic or Black people. In a subsequent recognition task, participants showed more accurate recognition of previously seen "Hispanic" faces than "Black" faces. Despite the fact that the ethnicity of the actual faces had not been altered, top-down category effects induced by hairstyle drove biases in face memory. In fact, such a manipulation has effects beyond face memory; in a follow-up investigation, MacLin and Malpass (2003) found that faces categorized as Black due to hairstyle were even judged to have darker skin tone than faces with Hispanic hairstyles.

In a related set of experiments, Pauker et al. (2009) found that memory for White–Black biracial faces was moderated by the racial label attached to the faces. Participants recognized own-race faces more accurately than other-race and biracial faces, but when the biracial faces were labeled as either Black or White, they elicited recognition consistent with their labels. Put simply, labeling biracial faces as own-race or other-race distorted memory toward the category label. Furthermore, simply encouraging participants to include biracial individuals in their ingroup led to better memory for biracial faces. In follow-up work, Pauker et al. (2013) demonstrated that not only can changing perceptions of biracial targets (via labeling) change memory, but that changing perceptions of the *self* can have analogous effects. Here, White/Black biracial participants wrote essays either about how they were connected to their mother's or their father's ethnic identity, thereby priming either their White or their Black identity. Adopting different racial identifications dictates face memory—biracial participants who were induced to think of their White identity showed better memory for White faces, whereas biracial participants induced to think of their Black identity showed better memory for Black faces. Perhaps even more strikingly, the impact of social categorization on memory for racially ambiguous faces can even be seen among young children (Gaither et al. 2014). These researchers measured the extent to which White children between 4 and 9 years of age think about race in essentialist (i.e., highly biological rather than cultural) terms. They found that children who used essentialist thinking remembered racially ambiguous faces as poorly as Black faces. However, children who did not rely on essentialist thinking remembered White and ambiguous faces more accurately than Black faces. In critical ways, perceived membership in a social category drives perceptions of and processing of faces. In many cases, this differential processing results in worse recognition for racial and ethnic outgroup members.

This work illuminates another influence of categorization: categories themselves can signal that the unique, individuating characteristics of outgroup members are unnecessary to process. The individual identities of members of some groups just seem less personally relevant than others (Rodin 1987; Wilson et al. 2014). With regard to the ORB, the identities of racial outgroup members (i.e., other-race faces) often seem less relevant than the identities of racial ingroup members (i.e., own-race

faces). Unless someone has a reason to move beyond a category label for other-race faces, attending to the individual characteristics of other-race faces becomes much less likely. Of course, some eyewitness contexts themselves may provide a reason to attempt to remember a person's face—many crime suspects may attempt to encode the details of perpetrators. However, mere processing motivation alone is insufficient to eliminate the ORB (Hugenberg et al. 2007). Instead, perceivers must attempt to *individuate* faces—that is, to pay attention to what *distinguishes* this face from other similar faces, thereby moving beyond the category information in processing. Thus, the often reported attempt to "remember the perpetrator" by an eyewitness in a real crime situation will not necessarily lead to an individuated face representation because perceivers simply apply effort, without applying that effort to encoding what makes a face unique (see Hugenberg et al. 2010).

Further, not all crimes are obvious, even to those in the context. A bank robber with a threatening note may seem quite important to the teller, which may motivate encoding, but may seem perfectly normal to others in line. Finally, in other crime contexts, there may be components of the situation that may make focusing on the individuating characteristics of suspects faces relatively challenging (e.g., threat, darkness, or the presence of a weapon), that can similarly distract attention from face individuation, an issue to which we return later. Thus when perceivers are distracted or not otherwise motivated, this differential tendency to encode the individuating characteristics of own-race faces can have a demonstrable effect on memory.

Hugenberg et al. (2007) provided some of the first evidence that was explicitly in support of this motivational perspective. In a straightforward procedure, they found that the commonly observed ORB could be eliminated by warning participants about the ORB, and instructing them to attend to what differentiates among the faces of racial outgroup members (see Appendix for exact instructions). Participants who received these individuation instructions before encoding showed no ORB, while participants who were simply asked to attend to the faces showed the well-replicated deficit for other-race recognition. Similar effects can be obtained by offering financial incentives for strong other-race recognition (e.g., Kawakami et al. 2014), although this may be most effective when perceivers can control the length of encoding (DeLozier and Rhodes 2014).

Although not replicated in all research (e.g., Bornstein et al. 2013), these and related findings are among a growing wave of social cognitive-oriented investigations of the ORB. Broadly speaking, the growing body of evidence suggests that perceiver and target factors that are likely to elicit individuation will attenuate or eliminate the ORB. Although some OGBs may tend to be of a larger magnitude than others, membership in a variety of outgroups, not just groups defined by race, seems sufficient to reduce recognition. For example, Bernstein et al. (2007) conducted research in which group membership was defined either by university affiliation (same or different university) or arbitrarily assigned membership to a group supposedly defined by personality characteristics. In both of these examples, White perceivers showed differential recognition for White faces according to these group distinctions. Ingroups defined by university and even personality can apparently yield biases in memory that appear analogous to race, perhaps by

changing the nature of how people encode faces (e.g., more attention to ingroup eyes; Kawakami et al. 2014; but see McDonnell et al. 2014 for more mixed results). Similar effects have been observed across religious boundaries (e.g., Mormons better recognize fellow Mormons, holding race constant; Rule et al. 2010) and across sexual orientation (e.g., straight perceivers better recognize straight targets; Rule et al. 2007).

Other researchers have used inventive manipulations in order to further explore the extent to which social group-based categorical thinking beyond race per se can create biases in face processing and memory that are quite similar to previously observed biases based on race. In some such work, Van Bavel et al. (2011) assigned White participants to teams—they would be members of either the Leopards or the Tigers. These participants viewed both same- and other-group faces that were either Black or White. These researchers found that group membership, but not race, moderated face memory. Furthermore, this same work found that group membership, but not race, was associated with differential activation of the fusiform face area (FFA), which is a region of the temporal lobe that has been found to play a role in individuated processing of faces (e.g., Furl et al. 2011; Tarr and Gauthier 2000). Other work has shown that even arbitrarily assigned group categorizations (e.g., assigning people into groups on the basis of whether the over- or underestimate the number of dots in an array) can influence processing of faces as early as 170 ms after exposure (Ratner and Amodio 2013). Results such as these suggest that other-race targets may be individuated if they are part of a superordinate ingroup that crosses racial boundaries (see also Hehman et al. 2010, but see Kloth et al. 2014 for a different conclusion). In short, categorizing someone as an ingroup member (even when that individual is of a different race) can improve recognition, whereas categorizing that same person as an outgroup member can reduce recognition. These findings suggest that social categorization of targets as outgroup members may lead to worse recognition through decreased motivation, which leads to less efficient encoding of targets.

Taken together, this body of research indicates that social categories themselves, even beyond prior experience with faces, can both create and eliminate the ORB. Moreover, categories other than race, including even arbitrary group memberships, can create other-race like effects. Importantly, however, this research indicates that motives outside of category information, elicited by the behavior of the targets or even instructions to perceivers, can reduce or even eliminate the ORB. This may be leveraged to reduce eyewitness errors in the future by allowing us to think more comprehensively about the specific causes of and solutions to group-based memory bias.

A Synthetic Model of the ORB

These two causes of the ORB outlined above have been supported across dozens of experiments. As previously noted, each individual cause can account for some

variance in the ORB, but in concert they may be particularly powerful in explaining how the ORB is created, and how it might be eliminated (Young and Hugenberg 2012). These perspectives also share one basic commonality that should be critical in forming interventions for reducing the ORB: they both show that the ORB is related to a failure to process outgroup targets as individuals. Whether this failure is rooted in lack of experience, or in lack of attention to the individual identity because of a focus on the social category, perceivers are not identifying what uniquely distinguishes one face from another. In suggesting interventions for reducing the ORB in eyewitness memory, we will therefore focus on *individuation*, or the process of attending to what differentiates between faces. Furthermore, based on an integration of findings from the literature, we focus on interventions that synthesize both experience and motivation in reducing the ORB (Hugenberg et al. 2013).

We must emphasize that each of these two causes of the ORB focuses on the *encoding* phase of face memory. That is, a purely experience-based, a purely motivational, and a synthetic perspective all focus on the processes involved with initially learning a face (e.g., seeing it at the scene of a crime), rather than what happens when attempting to later recognize a target (e.g., during a police lineup). As previously noted, at least in terms of recognition sensitivity, it seems that the ORB is an encoding phenomenon—it arises during the actual encounter with the face (see Young et al. 2010). This means that there is little one can do after witnessing a crime to *improve* recognition sensitivity for perpetrators' faces, other than to attempt subsequent recognition as early as possible after witnessing a crime (as with all memories, face memory degrades over time, Charman and Wells 2013). It is possible to change race-based *criterion* effects after encoding—thus, we can make perceivers less likely to "false alarm" (i.e., say "yes" to an unseen face) but at the cost of fewer correct identifications (i.e., "hits")—in a number of ways. However, given the practical trade-off between these effects as discussed above, we focus primarily on encoding.

Improving Face Recognition: Focus on Sensitivity

The previous section may seem to lead to the pessimistic inference that once a crime is witnessed, there is nothing one can do to improve other-race recognition sensitivity. However, this conclusion would mask a key point of intervention. There are multiple effective interventions that can be employed before witnessing a crime (or when trying to encode any face) that can substantially reduce the ORB by influencing recognition sensitivity. As such, in the following section we will focus both on general strategies for improving eyewitness identification and on ORB-specific strategies designed to reduce race-based deficits in face memory. These general influences are not specific to the ORB, and therefore we will not discuss them at length. Rather, we focus primarily on the specific strategies that have proven effective at reducing the ORB at encoding. For both the general

influences on recognition and the race-specific strategies, we focus primarily on the first time point at which interventions can be undertaken—at encoding (i.e., during or before a crime, when influences on sensitivity are possible). However, we include a secondary focus on interventions at the point of recognition (i.e., during lineup or courtroom identification, when criterion only criterion shift is possible).

Non-Race-Specific Influences on Sensitivity

Although we will not include an exhaustive discussion, it is important to note that face recognition and the inferences that we can make about eyewitness memory based on self-report are subject to many influences beyond race (see Wells and Olson 2003 for a review). We will briefly discuss a few of these broad issues.

Perceptual quality and psychological state. Various perceptual factors can hinder perceivers' ability to accurately encode faces and therefore can make accurate eyewitness identifications problematic. One key factor is the quality of the viewing conditions in which the face is initially encoded (Shapiro and Penrod 1986). Viewing a face for just a brief time (Bornstein et al. 2012; Light et al. 1979; Maclin et al. 2001) or at a long distance (Malpass and Devine 1981) can often elicit poor subsequent recognition. These influences reduce the ease with which a face is processed, often involving suboptimal viewing conditions, which then translate into poor memory. More recent work has shown that some qualities of the external environment can have even more specific effects on recognition. For example, ambient darkness may reduce accuracy overall in eyewitness situations, but it also has been shown to result in poor accuracy specifically for faces exhibiting angry expressions (Nakashima et al. 2014). This latter work has not been extended to other-race targets but it does suggest that law enforcement should be cautious about assessing face memory in situations characterized by darkness or threat, common components of eyewitness scenarios (Doleac and Sanders 2015; Felson and Poulsen 2003).

To this same end, researchers also know of psychological influences on face encoding, leading to weak memory in spite of high-quality viewing conditions. For example, both distraction and stress have been shown to undermine face recognition. One typical example of a situation that can elicit both distraction and stress is the presence of a firearm or a weapon during a crime—known as the *weapon focus effect*. Loftus (1979), Loftus et al. (1987) and others (Cutler et al. 1987) have found that the presence of a firearm during a crime reduces eyewitnesses' subsequent recognition accuracy. Early research suggested that perceivers focus on weapons for longer than other items in a scene (Loftus et al. 1987), potentially at the cost of recognition of the perpetrators. Other research more directly investigated the effects of weapon focus on identity recognition. For example, Cutler et al. (1987) found that perceivers who viewed a video of a robbery involving a gun were less accurate in a subsequent lineup than those who viewed a robbery in which the gun was not visible. While the salience of a weapon was detrimental to accurate recognition, it should be noted that even in the no-weapon condition, only 46 % of perceivers made correct identification judgments.

Weapon focus effects have been observed both with guns (e.g., Hulse and Memon 2006) and other weapons such as knives (Pickel 2009). A growing body of work has corroborated these effects. Meta-analytically, weapon focus effects are moderate (g = 0.53; Fawcett et al. 2013) and reliable.

Eyewitness confidence and accuracy. Earlier work on the topic found that the confidence of an eyewitness is only weakly related to their recognition *accuracy* (see Sporer et al. 1995, for a review). Early field research supported this finding. For example, in a study carried out by Brigham et al. (1982) store clerks were asked to identify "customers" who had been in their store 2 hours earlier. Surprisingly, clerks were correct less than 50 % of the time. When including clerks who did not attempt to guess, accuracy was only 34 % (chance responding was 16.7 % in this case). Though not truly analogous to a crime attempt, the data from this study demonstrate the extent to which people may be inaccurate at recalling facial identities in real-life settings. Especially alarming is the fact that 85 % of clerks reported being confident enough that they would testify in court.

Fortunately, more recent work has found a much stronger relationship between confidence and accuracy (Martire and Kemp 2009; Neuschatz et al., in press). A similar correspondence between confidence and accuracy has been found using calibration analysis, in which researchers look at both resolution (similar to the point biserial correlation between confidence and accuracy) and calibration curves in which confidence is plotted against accuracy (Palmer et al. 2013). The latter allows researchers to determine how well responses are calibrated (e.g., with perfect calibration, identifications made with 80 % confidence would be correct 80 % of the time). These recent findings are encouraging, as eyewitness confidence is quite impactful in jury decisions (e.g., Cutler et al. 1990b; Lindsay 1994).

The confidence-accuracy relationship is relevant here because it is important to know whether the link between confidence and accuracy differs for same- and other-race targets. Unfortunately, little research has examined this question. However, in one recent study, Wylie et al. (2015) found some evidence that the confidence–accuracy relationship may be somewhat stronger for same-race than other-race targets, but only among younger adults. Other work has found no difference in the confidence-accuracy relationship between same- and other-race identifications (Smith et al. 2001). More work with larger sample sizes should be conducted with an eye toward determining whether confidence is indeed a better predictor of accuracy for same-race eyewitness identifications.

Race-Specific Factors Influencing Race-Based Biases in Face Recognition

Although it is informative to look to the general literature on factors influencing face recognition and eyewitness identification, it is most important in the present work to consider race- and group-specific factors that have been found relevant to eyewitness identification sensitivity. It is these factors that are most centrally

informed by a synthetic perspective on face memory and the group-based biases that frequently arise.

Individuation motivation influences sensitivity. As we have previously noted, one of the central features of recent theorizing on the ORB is that categorical thinking disrupts the extraction of unique individuating information in others. The use of categories (e.g., race and profession) draws attention to similarities and deemphasizes information that differentiates individual faces from one another (see Levin 1996, 2000), or even leads to disregarding faces altogether (Rodin 1987). Category-based deficits in face recognition unfold across a wide variety of categories (Hugenberg et al. 2010, 2013), suggesting that group-based bias in face memory poses a significant challenge.

There is some cause for optimism, however. A number of studies have shown that experimental interventions can induce perceivers to deemphasize categories and encode unique facial identities in the laboratory (Hugenberg et al. 2007; Rhodes et al. 2009; Van Bavel et al. 2011; Young et al. 2010; Young and Hugenberg 2012). For example, the previously reported experiment conducted by Hugenberg et al. (2007) showed that individuation instructions can reduce the ORB. White perceivers were notified before encoding that the ORB occurs often, and that they should attempt to closely attend to what differentiates individual other-race faces from one another. This instruction was simple and brief, yet effective. The effectiveness of these same instructions has been replicated across multiple experiments and laboratories (e.g., Rhodes et al. 2009; Young et al. 2010; Young and Hugenberg 2012), albeit not all (e.g., Bornstein et al. 2013). Indeed, a full understanding of the boundary conditions of such manipulations is yet to be achieved. However, it is the case that individuation motivation can occur without explicit instructions in some situations. As previously described, Bernstein et al. (2007) found that mere ingroup/outgroup distinctions (such as university affiliation or even arbitrarily assigned personality types) were sufficient to create superior recognition for ingroup relative to outgroup faces, and Van Bavel et al. (2011) extended this work to show that lab-based ingroups yield superior recognition even for other-race targets (see also Hehman et al. 2010; cf. Kloth et al. 2014). In other work, anticipating future interpersonal interactions with outgroup members (including racial outgroups) is associated with more accurate recognition (Wilson et al. 2014). Taken together, this research indicates that a shared group membership at encoding can substantially improve recognition, and attenuate or even eliminate the ORB.

Other-race individuation training over time influences sensitivity. A simple reading of experience-based perspectives on the ORB would suggest that a training program involving repeated exposure to other-race races would be beneficial in reducing the ORB in eyewitness identification. Because the core cause of the ORB, according to this perspective, is a lack of experience differentiating among other-race faces, one solution to this deficit might be a program of exposure to other-race faces. Unfortunately, it is unlikely that such a strategy, absent the motivation to actually individuate those faces, would work effectively over time. First, training has been found to be of limited effectiveness in reducing the ORB (see Brigham et al. 2007). Further, as we noted above, Meissner and Brigham's (2001) meta-analysis finds that self-reported exposure to other-race faces accounts

for only about 2 % of the variance in the ORB. Indeed, there are examples using real groups that indicate that real-world other-race contact can fail to create improvements in other-race recognition. For instance, Ng and Lindsay (1994) find that Whites living in Singapore showed just as large of an ORB for East Asian faces as Whites living in North America, despite the clear difference in other-race contact.

More recent research (e.g., Tanaka and Pierce 2009) appears to have offered more clarity on this contact problem: accurate other-race recognition requires *high-quality* contact. Just living among those of another race does not necessarily guarantee a large amount of individuated other-race interaction. Rather, if a White person, for example, lives in Singapore but still works and lives primarily among other Whites, he or she may have little motivation to discriminate among Asian faces. In this vein, Tanaka and Pierce (2009) found that training participants to classify other-race targets according to *individual identity* led to better subsequent recognition accuracy than training participants to classify faces by ethnicity. Their approach is based on the idea that perceivers' default strategy with other-race faces is to categorize them according to race. However, if a learning task requires that the faces be encoded according to individual identity, perceivers will show enhanced individuation, even for other-race faces. One cannot simply expose people to other-race faces and expect to see improved recognition. Rather, there should be an explicit focus on individuated encoding. Tanaka and Pierce's training (2009) also led to measurable differences in neural activity in addition to the behavioral improvement; when viewing the faces, participants who were induced to classify other-race faces by individual identity showed neural activity associated with experience. As such, researchers can lead people to remember other-race faces as though they are experts with a focus on individuation.

Race-based biases in lineup presentation. Although this review is not centrally focused on post-encoding influences, which tend to influence criterion but not sensitivity, it is important to briefly summarize the strategies that have been found to manage eyewitness errors. One strategy that has been widely suggested, and subsequently adopted in many police departments (Jonsson 2007), is the use of sequential lineups; rather than presenting witnesses with suspects simultaneously as one group, investigators present witnesses with suspects one at a time. Despite some evidence for the effectiveness of sequential lineups (e.g., Lindsay and Wells 1985), more recent investigations have questioned their advantages. For example, Meissner et al. (2005b) argued that sequential lineups result in higher selectivity in choosing targets. Perceivers become more likely to avoid picking an incorrect target, but they also become less likely to correctly identify a perpetrator when the target is present. In technical terms, participants' overall recognition sensitivity may not be improved, but they do show a shift in response bias toward becoming more selective. Other work has found that the sequential lineup advantage only occurs under limited circumstances, such as when the lineup construction is biased or the suspect's location in the lineup comes near the end (Carlson et al. 2008). In a recent assessment of the literature on sequential lineups, Malpass et al. (2009) argued that the overall body of research does not justify a mandated move to sequential lineups.

Most central to the current work, some research has specifically examined the ORB in sequential versus simultaneous lineup procedures. For example, Wylie et al. (2015) found that, as in other research, younger participants used a more liberal criterion, being more likely to identify suspects in simultaneous lineups, but only for Black targets. Older participants showed the opposite pattern, showing a more liberal criterion for Black targets in sequential lineups. Furthermore, contributing to the lack of clarity regarding sequential versus simultaneous lineups, Wylie et al. (2015) found that accuracy was greater for simultaneous lineups.

Further, it is also important to point out that race-based biases on the part of those who *construct* the police lineups can lead to post-encoding influences, even for perceivers who have encoded other-race faces quite well. Police lineups are most effective when the nonsuspect "foils" in the lineup have been chosen well. Foils should resemble the suspect as much as possible. When this occurs, selection of the suspect is more diagnostic of a correct identification. When the suspect is highly distinct from the foils, on the other hand, less can be concluded from the identification. Consider a hypothetical of a White eyewitness who observes a Black suspect committing a crime. In an extreme case in which there is only one Black target in the subsequent lineup, an identification will be made solely on the basis of his racial category information, without any need for individuating information. Here, there can be no reasonable confidence that a correct identification was made. Although this hypothetical is unrealistic, it is illustrative of the need for foils that are similar in appearance to the suspect, as it is easy to imagine other ways in which the suspect may be distinct from the foils (e.g., taller, more heavyset, only person with facial hair). Lineups in which the foils do not share salient characteristics with the target, then, pose a problem.

The benefit of unbiased lineups to eyewitness identification more broadly is actually questionable. For example, Gronlund et al. (2012) found no differences in discriminability between fair lineups and biased lineups. However, there may still be benefits that are specific to the OGB. For example, Brigham and Ready (1985) asked White and Black participants to construct lineups using both Black and White suspects. They found that White lineup constructors were less selective about who was included in a lineup of Black faces, whereas Black lineup constructors were less selective of who was included in a lineup of White faces. Participants set different thresholds for inclusion based on target race, and they tended to select other-race foils who were easy to reject because they did not resemble the target. This created an own-race effect in lineup construction—lineups made across racial lines (e.g., a White person constructing a lineup of Black faces, or vice versa) were less fair than were same-race lineups. Furthermore, Brigham and Brandt (1992) found that Black officers found lineups less useful than White officers, particularly when the lineups consisted of Black targets. In light of this work, multiple researchers (e.g., Sporer 2001; Wells and Olson 2001) have recommended avoiding other-race lineup construction, and instead recommend that police who make lineups be of the same-race as the suspect in the investigation. We echo this recommendation.

Policy Suggestions for Reducing the ORB

The ORB poses a formidable challenge because its locus is at encoding. As such, it is difficult to reduce in many real eyewitness identification situations. Although perceivers can manage errors, there is little that one can do to improve face memory after encoding. However, we propose that future research investigate two primary possibilities, drawn from research that has demonstrated reductions in the ORB, as contenders for implementation for public policy: (1) individuation training for police and (2) unbiased construction of police lineups. We close this section by briefly discussing legal interventions prescribing caution in the use of eyewitness testimony.

Other-Race Individuation Training for Police

Police officers commonly serve as eyewitnesses (Lipton 1996), and there is no reason to believe that their memory performance will tend to be qualitatively different from other eyewitnesses. Although some legal opinions have taken the position that police are particularly good eyewitnesses (Lipton 1996), others have concluded that police may be no more effective at eyewitness identification (Billig and Milner 1976). However, some research in more naturalistic settings has indicated that police do show better recall than civilians, though only under conditions of relatively long exposure, such as when a target asks for directions (Clifford and Richards 1977). Regardless, the empirical work on this topic is limited to just a small handful of studies. More central to our purposes, none of the work comparing police to civilian eyewitnesses has involved race as a variable of interest.

Because police do serve as eyewitnesses to crimes and their aftermath, it is sensible to target them for individuation training. Furthermore, because police training is so highly proceduralized, it is much more plausible to engage police in a systematic individuation training intervention than it would be to target civilians more broadly. As such, our primary recommendation is to consider engaging police officers in other-race individuation training. Such training has been effective in laboratory settings at reducing the ORB, and we propose that preliminary programs could investigate how such training might be implemented amongst police departments, and to study the training's efficacy in improving face memory among officers.

Specifically, based partially on the research of Tanaka and Pierce (2009), and informed by the synthetic perspective on the OGB, we recommend testing a system in which police get sustained training discriminating among other-race faces. As previously discussed, Tanaka and Pierce asked participants to learn to give separate responses to a series of other-race faces. In this task, each individual target person was assigned an individual response. This type of training, in which there is explicit emphasis on individual identity, improves other-race recognition. Unfortunately, past research has indicated that the effectiveness of other-race training of this sort does decay over time (Goldstein and Chance 1985). This suggests that the training

should be repeated over time; even less than 15 min of such discrimination training can yield benefits in subsequent other-race recognition (Elliott et al. 1973; see also McGugin et al. 2011); however as with most skill training, it may need to occur repeatedly in order to be effective over a longer time course.

We also recommend that this individuation training program include an emphasis on using an *individuation mindset* when engaging with suspects. As previously discussed, Hugenberg et al. (2007) found that merely informing participants that the ORB exists, and then instructing them to attend to what differentiates individuals among other-race faces, led to a substantial increase in the recognition of other-race faces (see also Rhodes et al. 2009; Young et al. 2010; Young and Hugenberg 2012, for replications of this effect). Thus, when interacting with criminal suspects, it is important to attend to the suspect with an individuation mindset. People seem to do this spontaneously for same-race faces but apparently need to be reminded to do so for other-race faces (Hugenberg et al. 2010). Importantly, the work finding an association between anticipated interaction and the ORB (Wilson et al. 2014) suggests that such training may be especially beneficial, as it will provide not only motivation to individuate others, but also increased attention to the fact that interactions with outgroup targets are likely to occur. Insofar as police departments can make this a habit during police training, they may well be able to reliably reduce eyewitness errors at least amongst police officers.

These recommendations for investigating the implementation of Other-Race Individuation Training are based on a robust and growing body of evidence. Although researchers are constantly refining our understanding of what exactly underlies the ORB, the existing evidence for how to reduce it is strong enough that preliminary programs can now be designed and tested in an attempt to do so. As we have pointed out previously, we are not recommending a more general prejudice-reduction training strategy. Although it would be plausible to expect that racial prejudice would be a predictor of the ORB, a large body of research has thus far has demonstrated little if any relationship between prejudice and the ORB (see Meissner and Brigham 2001). Although there are certainly other important and timely reasons for increasing racial harmony between police and citizens, especially in the current climate emphasizing the need for trust amongst citizens and their police forces, we do not believe that prejudice reduction would meaningfully reduce the ORB independent of indirect effects of such training on individuation motivation. As such, we would suggest a more specific strategy that plays upon cognitions and motivations regarding the individuation of outgroup faces during encoding. Because this recommendation is based on laboratory research, it would be important to first conduct smaller scale pilot work. As with all such public policies, we propose that first preliminary research be conducted with local police departments, giving the ability to refine the training and understand the efficacy of the techniques when employed in an applied setting, before such techniques are scaled to wider implementation. Such collaborations between researchers and police departments have been effective in other recent attempts to improve officer performance (e.g., Sim et al. 2013), and could provide a model here.

We also wish to emphasize that we make these recommendations for police officers primarily because they commonly serve as eyewitnesses to crimes, and because police training is systematic and rigorous. As a result of these factors, training with police groups likely represents the most efficient use of resources for an individuation training program. Although the evidence suggests that individuation training could work for community members more broadly, limited resources would best be spent first targeting a specific group of people who are known to be frequent witnesses in criminal cases. Indeed, if evidence shows that police can benefit from a program of individuation training, larger scale efforts can be designed and implemented. More broadly, it would also be sensible to attempt to engage other high-frequency witnesses in individuation training. For example, clerks at convenience stores and bank tellers may be likely to witness crimes at a higher frequency than most citizens. Police departments could benefit from encouraging local merchants to participate in eyewitness training programs. Even for crimes that do not happen inside stores, having local citizens who are more capable of identifying suspects in crimes that occur in their vicinity could be beneficial.

Unbiased Construction of Police Lineups

In addition to our motivational recommendations, we echo the recommendations of Sporer (2001) and Wells and Olson (2001) to have police and investigators only create lineups for suspects of the same-race as the officer who constructs them. Although the issue of unintentional race-based bias in suspect lineups has long been recognized (Brigham and Ready 1985), recent controversial events in the United States have highlighted the unfortunate reality that many communities largely made up of minority citizens have police forces that are overwhelmingly White (Ashkenas and Park 2014). In the absence of an explicit policy mandating same-race lineup construction, we can expect this to continue to be a sustained problem in our criminal justice system, especially in instances in which the witness and the suspect are not the same-race.

As we discussed previously, Brigham and Ready (1985) found that people tend to select foils that resemble the suspect less closely when constructing other-race (relative to same-race) lineups. As such, we recommend that all police departments implement a policy of same-race lineup construction whenever possible. This policy is simple and straightforward, and there is clear reason to believe that it can reduce costly errors in eyewitness identification.

Legal Interventions: Caution in Using Eyewitness Testimony

For the reasons that we have summarized thus far, one of the most basic recommendations that we can make is that eyewitness testimony about strangers should be

treated with caution, and given the substantial evidence for poor other-race face recognition, other-race eyewitness testimony should be treated with even more caution than same-race testimony. To wit: a meta-analysis showed that an innocent Black suspect is 55 % more likely to be falsely identified by a White eyewitness than a Black eyewitness (Wells and Olson 2001). The research strongly indicates that this is not only a bias in the minds of Whites—the ORB can occur across any racial or ethnic distinction (e.g., McKone et al. 2012). We strongly encourage courts to consider the possibility of specifically and explicitly addressing other-race identification problems. Some courts have done so, recognizing how critical it is to allow the testimony regarding the unreliability of eyewitness testimony in general (Fradella 2006). We believe that this is not enough, and some states have recognized this. For instance, New Jersey's Supreme Court has mandated that juries be instructed specifically of the elevated likelihood of inaccuracies in other-race identifications when those identifications constitute a critical piece of evidence in the case and are not corroborated by additional evidence (*New Jersey v. Cromedy* 1999). More recently, the same court mandated a list of other reforms in how jurors are instructed, including whether the lineup was double blind, whether the witness was told that the suspect was present in the lineup, and whether there were situational characteristics that could hamper accurate memory (*State v. Henderson* 2011). The data overwhelmingly support that such actions should be implemented broadly.

Although this may seem like an uncontroversial recommendation, some have argued against such instructions, claiming that they are likely to be harmful. For example, Bartolomey (2001) argued that jury instructions regarding other-race identifications are overly general and are useless with regard to the specific identification in the case. In addition to claiming that the evidence for the ORB is "hardly overwhelming," Bartolomey argues that jurors are actually aware of the fallibility of perceivers. Identification testimony is presented in the context of a body of evidence that all must be assessed in order to prove that the suspect is guilty beyond a reasonable doubt. Thus, eyewitness testimony must be questioned just as any other evidence presented, but we should not attempt to call into question other-race identifications because of the existence of a general phenomenon. It is also argued that such a jury instruction would be unfairly biased, as it will always favor the defendant.

We disagree with these assertions. The evidence for the ORB is reliable, robust, and overwhelming (Meissner and Brigham 2001). In actual eyewitness situations, this translates into a real and systematic bias that unduly harms suspects who are not of the witness' ethnicity. Furthermore, the fallibility of eyewitnesses and the ORB are far from a matter of common sense. In fact, research has shown that jurors do not tend to understand the ways in which memory is fallible as a matter of common sense (Schmechel et al. 2006). Despite our support for instructions regarding the ORB, most research has found that instructions tend to have little effect on jurors' sensitivity to the quality of the witnessing conditions; rather, such instructions may be more likely to make jurors more broadly skeptical of eyewitness testimony (Papailiou et al. 2014). Other research has found no effect of instructions (Cutler et al. 1990). Some work using a more in-depth presentation did find

increased sensitivity to the quality of identification conditions (Pawlenko et al. 2013). Unfortunately, there has been very little work on the influence of instructions specific to the ORB. Regardless, based on what we know about the influence of race on memory, increased skepticism in cases of other-race identifications is likely a desirable outcome in the absence of more effective retrieval influences on other-race face discrimination.

Conclusion

Other-race eyewitness identification poses a real problem for the legal system. As researchers have known now for many decades, the ORB is a robust and replicable phenomenon. Despite its status as a popular topic of study among scholars, policymakers are certainly far from eliminating its influence in behavior and in eyewitness misidentifications. More is currently known about its causes than has previously been the case, and researchers and policymakers should harness this knowledge in order to improve the reliability of eyewitness identification. Perhaps even more importantly, this knowledge should be used to reduce the misuse of other-race identifications. Courts and court systems have slowly begun to embrace the idea that they must explicitly acknowledge the problems inherent in eyewitness testimony, and some have even allowed for special treatment of other-race testimony. Despite this, errors remain common. We have proposed some promising methods for reducing the ORB in eyewitness identification in the present chapter, including both strategies for improving eyewitness memory (e.g., with individuation training), and for managing errors and their likely impact. We strongly believe that both approaches are necessary if we are to reduce other-race eyewitness misidentifications to an acceptable level. Centrally, our perspective argues for a clear role of social motivational factors in eyewitness memory. We hope that such research will be informative to legal scholars and policymakers in the coming decades.

Appendix

Hugenberg, Miller, and Claypool's Individuation Instructions:

> Previous research has shown that people reliably show what is known as the Own-Race Bias (ORB) when learning faces. Basically, people tend to confuse faces that belong to other races. For example, a White learner will tend to mistake one Black face for another. Now that you know this, we would like you to try especially hard when learning faces in this task that happen to be of a different race. Do your best to try to pay close attention to what differentiates one particular face from another face of the same-race, especially when that face is not of the same-race as you…
>
> Remember, pay very close attention to the faces, especially when they are of a different race than you in order to try to avoid this Own-Race Bias.

References

Ashkenas, J., & Park, H. (2014, September 4). The race gap in America's Police Departments. *The New York Times*. Retrieved from http://www.nytimes.com/interactive/2014/09/03/us/the-race-gap-in-americas-police-departments.html

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104*, 54–75. doi:10.1037/0096-3445.104.1.54

Bartolomey, D. (2001). Cross-racial identification testimony and what not to do about it: A comment on the cross-racial jury charge and cross-racial expert identification testimony. *Psychology, Public Policy, and Law, 7*, 247–252. doi:10.1037/1076-8971.7.1.247

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science, 18*, 706–712. doi:10.1111/j.1467-9280.2007.01964.x

Billig, M., & Milner, D. (1976). A spade is a spade in the eyes of the law. *Psychology Today, 2*, 13–16.

Bornstein, B. H., Deffenbacher, K. A., Penrod, S. D., & McGorty, E. K. (2012). Effects of exposure time and cognitive operations on facial identification accuracy: A meta-analysis of two variables associated with initial memory strength. *Psychology, Crime & Law, 18*, 473–490. doi:10.1080/1068316X.2010.508458

Bornstein, B. H., Laub, C. E., Meissner, C. A., & Susa, K. J. (2013). The cross-race effect: Resistant to instructions. *Journal of Criminology.* doi:10.1155/2013/745836

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36).

Brigham, J. C., Bennett, L., Meissner, C. A., & Mitchell, T. (2007). The influence of race on eyewitness memory. In R. Lindsay, D. Ross, J. Read, & M. Toglia (Eds.), *Handbook of eyewitness psychology* (pp. 257–281). Mahwah, NJ: Lawrence Erlbaum & Associates.

Brigham, J. C., & Brandt, C. C. (1992). Measuring lineup fairness: Mock witness responses versus direct evaluations of lineups. *Law and Human Behavior*, *16*, 475–489. doi:10.1007/BF01044619

Brigham, J. C., Maass, A., Snyder, L. D., & Spaulding, K. (1982). Accuracy of eyewitness identification in a field setting. Journal of Personality and Social Psychology, 42, 673. doi:10.1037/0022-3514.42.4.673

Brigham, J. C., & Ready, D. J. (1985). Own-race bias in lineup construction. *Law and Human Behavior, 9*, 415–424. doi:10.1007/BF01044480

Bukach, C. M., Cottle, J., Ubiwa, J., & Miller, J. (2012). Individuation experience predicts other-race effects in holistic processing for both Caucasian and Black participants. *Cognition, 123*, 319–324. doi:10.1016/j.cognition.2012.02.007

Caharel, S., Montalan, B., Fromager, E., Bernard, C., Lalonde, R., & Mohamed, R. (2011). Other-race and inversion effects during the structural encoding stage of face processing in a race categorization task: An event-related brain potential study. *International Journal of Psychophysiology, 79*, 266–271. doi:10.1016/j.ijpsycho.2010.10.018

Carlson, C. A., Gronlund, S. D., & Clark, S. E. (2008). Lineup composition, suspect position, and the sequential lineup advantage. *Journal of Experimental Psychology: Applied, 14*, 118–128. doi:10.1037/1076-898X.14.2.118

Cates, P., & Trigg, J. O. (2014). Louisiana man exonerated after serving nearly 17 years for an attempted rape that DNA evidence shows he didn't commit. *Innocence Project*. Retrieved from http://www.innocenceproject.org/news-events-exonerations/2014/louisiana-man-exonerated-after-serving-nearly-17-years-for-an-attempted-rape-that-dna-evidence-shows-he-didnt-commit

Charman, S., & Wells, G. L. (2013). Applied lineup theory. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *The handbook of eyewitness psychology, Vol II: Memory for people* (pp. 219–254).

Clifford, B. R., & Richards, V. J. (1977). Comparison of recall by policemen and civilians under conditions of long and short durations of exposure. *Perceptual and Motor Skills, 45*, 503–512. doi:10.2466/pms.1977.45.2.503

Corneille, O., Huart, J., Becquart, E., & Bredart, S. (2004). When memory shifts toward more typical category exemplars: Accentuation effects in the recollection of ethnically ambiguous faces. *Journal of Personality and Social Psychology, 86*, 236–250. doi:10.1037/0022-3514.86.2.236

Cutler, B. L., Dexter, H. R., & Penrod, S. D. (1990a). Nonadversarial methods for sensitizing jurors to eyewitness evidence. *Journal of Applied Social Psychology, 20*, 1197–1207. doi:10.1111/j.1559-1816.1990.tb00400.x

Cutler, B. L., Penrod, S. D., & Dexter, H. R. (1990b). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior, 14*, 185–191. doi:10.1007/BF01062972

Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987). Improving the reliability of eyewitness identification: Putting context into context. *Journal of Applied Psychology, 72*, 629–637. doi:10.1037/0021-9010.72.4.629

DeLozier, S., & Rhodes, M. G. (2014). The impact of value-directed remembering on the own-race bias. *Acta Psychologica, 154*, 62–68. doi:10.1016/j.actpsy.2014.11.009

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115*, 107–117. doi:10.1037/0096-3445.115.2.107

Doleac, J. L., & Sanders, N. J. (2015). Under the cover of darkness: How ambient light influences criminal activity. *Review of Economics and Statitics, 97*, 1093–1103. doi:10.1162/REST_a_00547

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia, 44*, 576–585. doi:10.1016/j.neuropsychologia.2005.07.001

Elliott, E. S., Wills, E. J., & Goldstein, A. G. (1973). The effects of discrimination training on the recognition of white and oriental faces. *Bulletin of the Psychonomic Society, 2*, 71–73. doi:10.3758/BF03327717

Evans, J. R., Marcon, J. L., & Meissner, C. A. (2009). Cross-racial lineup identification: Assessing the potential benefits of context reinstatement. *Psychology, Crime & Law, 15*, 19–28. doi:10.1080/10683160802047030

Fawcett, J. M., Russell, E. J., Peace, K. A., & Christie, J. (2013). Of guns and geese: A meta-analytic review of the 'weapon focus' literature. *Psychology, Crime & Law, 19*, 35–66.

Felson, M., & Poulsen, E. (2003). Simple indicators of crime by time of day. *International Journal of Forecasting, 19*, 595–601. doi:10.1080/1068316X.2011.599325

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 44, pp. 1–74). doi:10.1016/S0065-2601(08)60317-2

Fradella, H. F. (2006). Why judges should admit expert testimony on the unreliability of eyewitness testimony. *Federal Courts Law Review*, 3-93. Retrieved from http://www.fclr.org/fclr/articles/html/2006/fedctslrev3.pdf

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review, 118*, 247–279. doi:10.1037/a0022327

Furl, N., Garrido, L., Dolan, R. J., Driver, J., & Duchaine, B. (2011). Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. *Journal of Cognitive Neuroscience, 23*(7), 1723–1740. doi:10.1162/jocn.2010.21545

Gaither, S. E., Schultz, J. R., Pauker, K., Sommers, S. R., Maddox, K. B., & Ambady, N. (2014). Essentialist thinking predicts decrements in children's memory for racially ambiguous faces. *Developmental Psychology, 50*, 482–488. doi:10.1037/a0033493

Goldsmith, W. W., & Blakely, E. J. (2010). *Separate societies: Poverty and inequality in U.S. Cities*. Philadelphia: Temple.

Goldstein, A. G., & Chance, J. E. (1985). Effects of training on Japanese face recognition: Reduction of the other-race effect. *Bulletin of the Psychonomic Society, 23*, 211–214. doi:10.3758/BF03329829

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1). New York: Wiley.

Gronlund, S. D., Carlson, C. A., Neuschatz, J. S., Goodsell, C. A., Wetmore, S. A., Wooten, A., et al. (2012). Showups versus lineups: An evaluation using ROC analysis. *Journal of Applied Research in Memory and Cognition, 1*, 221–228. doi:10.1016/j.jarmac.2012.09.003

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*(6), 223–233. doi:10.1016/S1364-6613(00)01482-0

Hancock, P. J., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences, 4*(9), 330–337. doi:10.1016/S1364-6613(00)01519-9

Hancock, K., J., & Rhodes, G. (2008). Contact, configural coding, and the other-race effect in face recognition. *British Journal of Psychology*, *99*, 45–56. doi:10.1348/000712607X199981

Harrison, S. A., Gauthier, I., Hayward, W. G., & Richler, J. J. (2014). Other-race effects manifest in overall performance, not qualitative processing style. *Visual Cognition*, *22*, 843–864. doi:10.1080/13506285.2014.918912

Hehman, E., Mania, E. W., & Gaertner, S. L. (2010). Where the division lies: Common ingroup identity moderates the cross-race facial recognition effect. *Journal of Experimental Social Psychology, 46*, 445–448. doi:10.1016/j.jesp.2009.11.008

Hugenberg, K., Miller, J., & Claypool, H. M. (2007). Categorization and individuation in the cross-race recognition deficit: Toward a solution to an insidious problem. *Journal of Experimental Social Psychology, 2007*, 334–340. doi:10.1016/j.jesp.2006.02.010

Hugenberg, K., Wilson, J. P., See, P. E., & Young, S. G. (2013). Toward a synthetic model of Own Group Biases in face memory. *Visual Cognition, 21*, 1392–1417. doi:10.1080/13506285.2013.821429

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review, 117*, 1168–1187. doi:10.1037/a0020463

Hulse, L. M., & Memon, A. (2006). Fatal impact? The effects of emotional arousal and weapon presence on police officers' memories for a simulated crime. *Legal and Criminological Psychology, 11*, 313–325. doi:10.1348/135532505X58062

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology, 85*, 616–626. doi:10.1037/0022-3514.85.4.616

Jackiw, L. B., Arbuthnott, K. D., Pfeifer, J. E., Marcon, J. L., & Meissner, C. A. (2008). Examining the cross-race effect in lineup identification using Caucasian and First Nations samples. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 40*, 52–57. doi:10.1037/0008-400x.40.1.52

Jonsson, P. (2007). The police lineup is becoming suspect practice. *Christian Science Monitor*. Retrieved June 15, 2011, from http://www.csmonitor.com/2007/0206/p01s02-usju.html

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience, 17*, 4302–4311.

Kawakami, K., Williams, A., Sidhu, D., Choma, B. L., Rodriguez-Bailón, R., Cañadas, E., et al. (2014). An eye for the I: Preferential attention to the eyes of ingroup members. *Journal of Personality and Social Psychology, 107*, 1–20. doi:10.1037/a0036838

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science, 18*, 1084–1089. doi:10.1111/j.1467-9280.2007.02029.x

Kloth, N., Shields, S., & Rhodes, G. (2014). On the other side of the fence: The effects of social categorisation and spatial arrangement on memory for own-race and other-race faces. *Journal of Vision, 14*, 1261. doi:10.1371/journal.pone.0105979

Lane, S. M., & Meissner, C. A. (2008). A 'middle road' approach to bridging the basic–applied divide in eyewitness identification research. *Applied Cognitive Psychology, 22*, 779–787. doi:10.1002/acp.1482

Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 22*, 1364–1382. doi:10.1037/0278-7393.22.6.1364

Levin, D. T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology: General, 129*, 559–574. doi:10.1037/0096-3445.129.4.559

Light, L. L., Kayra-Stuart, F., & Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 212–228. doi:10.1037/0278-7393.5.3.212

Lindsay, R. C. L. (1994). Expectations of eyewitness performance: Jurors' verdicts do not follow from their beliefs. In D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 362–384). New York: Cambridge University Press.

Lindsay, R. C. L., & Wells, G. L. (1983). What do we really know about cross-race eyewitness identification. *Evaluating witness evidence: Recent psychological research and new perspectives*, 219–233.

Lindsay, R. C. L., & Wells, G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology, 70*, 556–564. doi:10.1037/0021-9010.70.3.556

Lipton, J. P. (1996). Legal aspects of eyewitness testimony. In S. L. Sporer, R. S. Malpass, & G. Koehnken (Eds.), *Psychological issues in eyewitness identification* (pp. 7–22). Mahwah, NJ: Lawrence Erlbaum Associates.

Loftus, E. F. (1974). Reconstructing memory: The incredible eyewitness. *Psychology Today, 8*, 116–119. Retrieved from http://www.jstor.org/stable/29761487

Loftus, E. F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.

Loftus, E. F., Loftus, G. R., & Messo, J. (1987). Some facts about "weapon focus". *Law and Human Behavior, 11*, 55–62. doi:10.1007/BF01044839

MacLin, O. H., MacLin, K. M., & Malpass, R. S. (2001). Race, arousal, attention, exposure, and delay: An examination of factors moderating face recognition. *Psychology, Public Policy, and Law, 7*, 134–152. doi:10.1037/1076-8971.7.1.134

MacLin, O. H., & Malpass, R. S. (2001). Racial categorization of faces: The ambiguous race face effect. *Psychology, Public Policy, & Law, 7*, 98–118. doi:10.1037/1076-8971.7.1.98

MacLin, O. H., & Malpass, R. S. (2003). The ambiguous-race face illusion. *Perception, 32*, 249–252. doi:10.1068/p5046

Macmillan, N. A., & Creelman, C. D. (2004). Detection theory: A user's guide. *Psychology Press,*. doi:10.4324/9781410611147

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology, 51*, 93–120. doi:10.1146/annurev.psych.51.1.93

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Calvini, G. (1999). Seeing more than we can know: Visual attention and category activation. *Journal of Experimental Social Psychology, 35*, 590–602. doi:10.1006/jesp.1999.1396

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., Thorn, T. M. J., & Castelli, L. (1997). On the activation of social stereotypes: The moderating role of processing objectives. *Journal of Experimental Social Psychology, 33*, 471–489. doi:10.1006/jesp.1997.1328

Macrae, C. N., Quinn, K. A., Mason, M. F., & Quadflieg, S. (2005). Understanding others: The face and person construal. *Journal of Personality and Social Psychology, 89*, 686–695. doi:10.1037/0022-3514.89.5.686

Malpass, R. S., & Devine, P. G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology, 66*, 482–489. doi:10.1037/0021-9010.66.4.482

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology, 13*, 330–334. doi:10.1037/h0028434

Malpass, R. S., Lavigueur, H., & Weldon, D. E. (1973). Verbal and visual training in face recognition. *Attention, Perception, & Psychophysics, 14*, 285–292. doi:10.3758/BF03212392

Malpass, R. S., Tredoux, C. G., & McQuiston-Surrett, D. (2009). Public policy and sequential lineups. *Legal and Criminological Psychology, 14*, 1–12. doi:10.1348/135532508X384102

Martire, K. A., & Kemp, R. I. (2009). The impact of eyewitness expert evidence and judicial instruction on juror ability to evaluate eyewitness testimony. *Law and Human Behavior, 33*(3), 225–236. doi:10.1007/s10979-008-9134-z

Mason, M. F., Cloutier, J., & Macrae, C. N. (2006). On construing others: Category and stereotype activation from facial cues. *Social Cognition, 24*, 540–562. doi:10.1521/soco.2006.24.5.540

McDonnell, G. P., Bornstein, B. H., Laub, C. E., Mills, M., & Dodd, M. D. (2014). Perceptual processes in the cross-race effect: Evidence from eyetracking. *Basic and Applied Social Psychology, 36*(6), 478–493. doi:10.1080/01973533.2014.958227

McGugin, R. W., Tanaka, J. W., Lebrecht, S., Tarr, M. J., & Gauthier, I. (2011). Race-specific perceptual discrimination improvement following short individuation training with faces. *Cognitive Science, 35*, 330–347. doi:10.1111/j.1551-6709.2010.01148.x

McKone, E., Stokes, S., Liu, J., Cohan, S., Fiorentini, C., Pidcock, M., et al. (2012). A robust method of measuring other-race and other-ethnicity effects: The Cambridge Face Memory Test format. *PLoS ONE, 7*, 1–6. doi:10.1371/journal.pone.0047956

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7*, 3–35. doi:10.1037/1076-8971.7.1.3

Meissner, C. A., Brigham, J. C., & Butz, D. A. (2005a). Memory for own- and other-race faces: A dual-process approach. *Applied Cognitive Psychology, 19*, 545–567. doi:10.1002/acp.1097

Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005b). Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. *Memory and Cognition, 33*, 783–792. doi:10.3758/BF03193074

Michel, C., Rossion, B., Han, J., Chung, C., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science, 17*, 608–615. doi:10.1111/j.1467-9280.2006.01752.x

Nakashima, S. F., Morimoto, Y., Takano, Y., Yoshikawa, S., & Hugenberg, K. (2014). Faces in the dark: Interactive effects of darkness and anxiety on the memory for threatening faces. *Frontiers in Psychology, 5*, 1091. doi:10.3389/fpsyg.2014.01091

Neuschatz, J. S., Wetmore, K. K., Cash, D., Gronlund, S. D., & Goodsell, C. A. (in press). Comprehensive evaluation of showups. In M. Miller & B. Bornstein (Eds.), *Advances in psychology and law*. New York: Springer.

New Jersey v. Cromedy, 727 A.2d 456, 467 (N. J. 1999).

New Jersey v. Henderson 2011, 27 A. 3d 872 (N.J. 2011).

Ng, W., & Lindsay, R. C. L. (1994). Cross-race facial recognition: Failure of the contact hypothesis. *Journal of Cross-Cultural Psychology, 25*, 217–232. doi:10.1177/0022022194252004

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115*, 39–57. doi:10.1037/0096-3445.115.1.39

Palmer, M. A., Brewer, N., Weber, N., & Nagesh, A. (2013). The confidence-accuracy relationship for eyewitness identification decisions: Effects of exposure duration, retention interval, and divided attention. *Journal of Experimental Psychology: Applied, 19*, 55–71. doi:10.1037/a0031602

Papailiou, A. P., Yokum, D. V., & Robertson, C. T. (2014). The novel New Jersey eyewitness instruction induces skepticism but not sensitivity. *Arizona Legal Studies Discussion Paper*, 14–17. doi:10.1371/journal.pone.0142695

Pauker, K., Ambady, N., & Freeman, J. B. (2013). The power of identity to motivate face memory in biracial individuals. *Social Cognition, 31*(6), 780–791. doi:10.1521/soco.2013.31.6.780

Pauker, K., Weisbuch, M., Ambady, N., Sommers, S. R., Adams, R. B., & Ivcevic, Z. (2009). Not so black and white: Memory for ambiguous group members. *Journal of Personality and Social Psychology*, *96*, 795–810. doi:10.1037/a0013265

Pawlenko, N. B., Safer, M. A., Wise, R. A., & Holfeld, B. (2013). A teaching aid for improving jurors' assessments of eyewitness accuracy. *Applied Cognitive Psychology, 27*, 190–197. doi:10.1002/acp.2895

Pickel, K. L. (2009). The weapon focus effect on memory for female versus male perpetrators. *Memory*, *17*(6), 664–678. doi:10.1080/09658210903029412

Quinn, K. A., Mason, M. F., & Macrae, C. N. (2009). Familiarity and person construal: Individuating knowledge moderates the automaticity of category activation. *European Journal of Social Psychology, 39*, 852–861. doi:10.1002/ejsp.596

Quinn, K. A., Mason, M. F., & Macrae, C. N. (2010). When Arnold is "The Terminator", we no longer see him as a man: The temporal determinants of person perception. *Experimental Psychology, 57*, 27–35. doi:10.1027/1618-3169/a000004

Rangarajan, V., Hermes, D., Foster, B. L., Weiner, K. S., Jacques, C., Grill-Spector, K., et al. (2014). Electrical stimulation of the left and right human fusiform gyrus causes different effects in conscious face perception. *The Journal of Neuroscience*, *34*, 12828–12836. doi:10.1523/JNEUROSCI.0527-14.2014

Ratner, K. G., & Amodio, D. M. (2013). Seeing "us vs. them": Minimal group effects on the neural encoding of faces. *Journal of Experimental Social Psychology, 49*, 298–301. doi:10.1016/j.jesp.2012.10.017

Rhodes, G., Brake, S., Taylor, K., & Tan, S. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology*, *80*, 313–331. doi:10.1111/j.2044-8295.1989.tb02323.x

Rhodes, G., Locke, V., Ewing, L., & Evangelista, E. (2009). Race coding and the other-race effect in face recognition. *Perception*, *38*, 232–241. doi:10.1068/p6110

Rodin, M. J. (1987). Who is memorable to whom: A study of cognitive disregard. *Social Cognition*, *5*, 144–165. doi:10.1521/soco.1987.5.2.144

Rossion, B., & Michel, C. (2011). An experience-based holistic account of the other-race face effect. In A. Calder, G. Rhodes, M.H. Johnson, & J.V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 215–243). Oxford: Oxford University Press.

Rule, N. O., Ambady, N., Adams, R. B., & Macrae, C. N. (2007). Us and them: Memory advantages in perceptually ambiguous groups. *Psychonomic Bulletin & Review, 14*, 687–692. doi:10.3758/BF03196822

Rule, N. O., Garrett, J. V., & Ambady, N. (2010). Places and faces: Geographic environment influences the ingroup memory advantage. *Journal of Personality and Social Psychology, 98*, 343–355. doi:10.1037/a0018589

Sangrigoli, S. & de Schonen, S. (2004). Recognition of own-race and other-race faces by three-month-old infants. *The Journal of Child Psychology and Psychiatry*, *45*, 1219–1227. doi:10.1111/j.1469-7610.2004.00319.x

Sangrigoli, S., Pallier, C., Argenti, A. M., Ventureyra, V. A. G., & de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science.* doi:10.1111/j.0956-7976.2005.01554.x

Scheck, B., Neufeld, P., & Dwyer, J. (2000). *Actual innocence*. New York, NY: Signet.

Schmechel, R. S., O'Toole, T. P., Easterly, C., & Loftus, E. F. (2006). Beyond the ken? Testing jurors' understanding of eyewitness reliability evidence. *Jurimetrics*, 177–214.

Semmler, C., Brewer, N., & Douglass, A. B. (2012). Jurors believe eyewitnesses. In B. L. Cutler (Ed.), *Conviction of the innocent: Lessons from psychological research*. Washington, DC: APA. doi:10.1037/13085-009

Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100*, 139–156. doi:10.1037/0033-2909.100.2.139

Sim, J. J., Correll, J., & Sadler, M. S. (2013). Understanding police and expert performance when training attenuates (vs. exacerbates) stereotypic bias in the decision to shoot. *Personality and Social Psychology Bulletin*, *39*, 291–304. doi:10.1177/0146167212473157

Smith, S. M., Lindsay, R. C., Pryke, S., & Dysart, J. E. (2001). Postdictors of eyewitness errors: Can false identifications be diagnosed in the cross-race situation? *Psychology, Public Policy, and Law, 7*, 153–169. doi:10.1037/1076-8971.7.1.153

Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin, 118*, 315–327. doi:10.1037/0033-2909.118.3.315

Sporer, S. L. (2001). The own-race bias: Beyond recognition of faces in the laboratory. *Psychology, Public Policy, and the Law, 7*, 170–200. doi:10.1037/1076-8971.7.1.170

Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62*, 207–218. doi:10.1037/0022-3514.62.2.207

Susa, K. J., Meissner, C. A., & de Heer, H. (2010). Modeling the role of social-cognitive processes in the recognition of own- vs. other-race faces. *Social Cognition, 28*, 523–537. doi:10.1521/soco.2010.28.4.523

Tanaka, J. W., & Pierce, L. J. (2009). The neural plasticity of other-race face recognition. *Cognitive, Affective, and Behavioral Neuroscience, 9*, 122–131. doi:10.3758/CABN.9.1.122

Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience*, *3*, 764–770. doi:10.1038/77666

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36*, 778–793. doi:10.1037/0022-3514.36.7.778

Tong, Y., Yu, Y., Chen, L., Han, D., Sun, H., Qiu, X., et al. (2014). Configural processing advantage for Mongoloid than Caucasian faces during the structure coding stage. *Journal of Integrative Neuroscience*, *13*, 693–705. doi:10.1142/S0219635214500277

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology A*, *43*, 161–204. doi:10.1080/14640749108400966

Valentine, T., & Endo, M. (1992). Towards an exemplar model of face processing: The effects of race and distinctiveness. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *44A*, 671–703. doi:10.1080/14640749208401305

Van Bavel, J. J., Packer, D., & Cunningham, W. A. (2011). Modulation of the fusiform face area following minimal exposure to motivationally relevant faces: Evidence of in-group enhancement (not out-group disregard). *Journal of Cognitive Neuroscience, 23*, 3343–3354. doi:10.1162/jocn_a_00016

Wan, L., Crookes, K., Reynolds, K. J., Irons, J. L., & McKone, E. (2015). A cultural setting where the other-race effect on face recognition has no social–motivational component and derives entirely from lifetime perceptual experience. *Cognition*, *144*, 91–115. doi:10.1016/j.cognition.2015.07.011

Wells, G. L., & Olson, E. A. (2001). The other-race effect in eyewitness identification: What do we do about it? *Psychology, Public Policy, and the Law, 7*, 230–246. doi:10.1037/1076-8971.7.1.230

Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology*, *54*, 277–295. doi:10.1146/annurev.psych.54.101601.145028

Wilson, J. P., See, P. E., Bernstein, M. J., Hugenberg, K., & Chartier, C. (2014). Differences in anticipated interaction drive own group biases in face memory. *PloS one*, *9*(3), e90668. doi:10.1371/journal.pone.0090668

Wolff, N., Kemter, K., Schweinberger, S. R., & Wiese, H. (2014). What drives social in-group biases in face recognition memory? ERP evidence from the own-gender bias. *Social Cognitive and Affective Neuroscience*, *9*, 580–590. doi:10.1093/scan/nst024

Wylie, L. E., Bergt, S., Haby, J., Brank, E. M., & Bornstein, B. H. (2015). Age and lineup type differences in the own-race bias. *Psychology, Crime & Law*, *21*, 490–506. doi:10.1080/1068316X.2014.989173

Young, S. G., Bernstein, M. J., & Hugenberg, K. (2010). When do own-group biases in face recognition occur? Encoding versus post-encoding. *Social Cognition, 28*, 240–250. doi:10.1521/soco.2010.28.2.240

Young, S. G., & Hugenberg, K. (2012). Individuation motivation and face experience can operate jointly to produce the Own-Race Bias. *Social Psychological and Personality Science, 3*, 80–87. doi:10.1177/1948550611409759

Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2009). Interracial contexts debilitate same-race face recognition. *Journal of Experimental Social Psychology, 45*, 1123–1126. doi:10.1016/j.jesp.2009.05.009

Risk-Reducing Interventions for Justice-Involved Individuals: A Critical Review

Kirk Heilbrun, David DeMatteo, Chris King, Alice Thornewill and Sarah Phillips

The importance of services to justice-involved individuals that reduce the risk of re-offense seems clear. Both individuals receiving such services and the larger society in which they live benefit from responsible, prosocial behavior rather than criminal offending. However, there are important questions about the purpose, nature, and empirical support for rehabilitative services provided to individuals who become involved with the criminal justice system. These questions are addressed in this chapter, which also reviews the existing evidence supporting various kinds of interventions.

Scope and Definitional Considerations

In this chapter, we define "justice-involved" as, at minimum, having been arrested for criminal charges. We define this broadly to include arrests that may result in diversion or specialized intervention at first booking, jail, prison reentry, or parole —four of the five intercepts on the Sequential Intercept Model[1] (Griffin et al. 2015; Munetz and Griffin 2006). Our focus is adults, so we do not review evidence related to the rehabilitation of juvenile offenders. Although much of the evidence discussed

[1]The Sequential Intercept Model is a conceptual model that identifies five points at which standard criminal processing can be interrupted to provide diversion to treatment-oriented alternatives: (1) law enforcement/emergency services, (2) booking/initial court hearings, (3) jails/courts, (4) reentry, and (5) community corrections/community support.

K. Heilbrun (✉) · D. DeMatteo · C. King · A. Thornewill · S. Phillips
Department of Psychology, Drexel University, 119 Stratton Hall, 3141 Chestnut St., Philadelphia, PA 19104, USA
e-mail: kirk.heilbrun@drexel.edu; kh33@drexel.edu

D. DeMatteo
e-mail: dsd25@drexel.edu

B.H. Bornstein and M.K. Miller (eds.), *Advances in Psychology and Law*, Advances in Psychology and Law 2, DOI 10.1007/978-3-319-43083-6_9

in this chapter relates to the rehabilitation of offenders with behavioral health disorders,[2] we also cover rehabilitative interventions for individuals without such disorders. "Risk reduction" refers to lowering the risk of criminal reoffending, although we do not see this as mutually exclusive from clinical symptomatic improvement and recovery (as discussed in the section on the Clinical vs. Criminogenic Debate, later in this chapter). We address evidence regarding interventions in the community, in nonsecure residential settings (e.g., halfway houses), and in prisons and secure hospitals. We exclude only jails, as a substantial proportion of individuals in jails are detained pending disposition of charges, making it impossible to meaningfully separate the predisposition detention mission of jails for these individuals from the post-sentence secure rehabilitative mission of jails regarding sentenced individuals. We do not cover evidence regarding psychotropic medication. Remaining interventions, however—therapy/counseling, case management, and milieu—are addressed in the present chapter.

Prevalence of Behavioral Health Disorders in Justice-Involved Individuals

The estimated prevalence of behavioral health disorders among individuals involved with the criminal justice system depends on several considerations. One is how broadly "behavioral health" is defined, with some investigators using only severe mental illness and others broadening that definition to include nonpsychotic depression, anxiety, trauma, substance abuse, and intellectual disability. A second consideration is how this information is gathered; this ranges from self-report of specific symptoms to a very detailed structured diagnostic interview such as the Structured Clinical Interview for DSM-5 (First et al. 2015) supported by documented history. A third variable involves where the individual is placed: jails, state and federal prisons, and community-based placements all appear in the literature (see Sarteschi 2013, for a summary).

For present purposes, it is less important to identify precise estimates than to make a larger point: the prevalence of broadly defined behavioral health disorders in those involved in the criminal justice system is substantial and apparently increasing. It was estimated in 1999 that 7.4 % of federal offenders, 16.2 % of state offenders, and 16 % of those on probation had "mental health problems" (Ditton 1999). In 2006, by comparison, a Bureau of Justice Statistics report described 45 % of federal offenders and 56 % of state offenders as experiencing mental health problems (James and Glaze 2006). Using a structured clinical interview, Steadman et al. (2009) gathered Connecticut jail data that they used to estimate the national

[2]This is defined broadly to include serious mental illness, substance abuse, anxiety disorders, personality disorders, intellectual disability, brain dysfunction, learning disorders, trauma-related disorders, and disorders of attention and concentration.

prevalence of severe mental illness in jail populations as 17.1 % for males and 34.3 % for females. Other estimates of rates of significant behavioral health problems include 76 % (jail or probation; Castillo and Fiftal Alarid 2011), 18 % (New Jersey prisons; Blitz et al. 2005), and 78 % (probation or parole; Solomon et al. 2002). But severe mental illness is only a part of the broader spectrum of behavioral health disorders experienced by justice-involved individuals; substance use disorders, anxiety, and PTSD are also estimated to be common within this population (NCCHC 2002; Sarteschi 2013).

Considering this evidence, it seems justifiable to conclude that at least half (and perhaps more) of justice-involved individuals experience at least one behavioral health disorder when defined as broadly as is done in this chapter. This means that rehabilitative services delivered to such individuals could focus only on clinical outcomes such as symptom remission and still be insufficient to treat all individuals with significant clinical needs.

Despite the apparent breadth of such clinical need, however, we suggest that a primary focus on clinical treatment and traditional clinical outcomes is not appropriate for a justice-involved population. Services to justice-involved individuals are almost entirely publicly funded, and society has a substantial interest in supporting rehabilitative services that reduce the risk of further offending. But how can we reconcile this overwhelming need for behavioral health services in contemporary justice-involved populations with the legitimate societal interest in public safety? The answer has two parts. First, rehabilitative interventions for both offense risk reduction *and* behavioral health needs must be identified; second, we must employ those interventions that provide the greatest impact for shared treatment goals—those that both reduce re-offense risk *and* improve behavioral health functioning. This is discussed in more detail in the next two sections.

Risk-Need-Responsivity as an Organizing Structure

The Risk-Need-Responsivity (RNR) model (Andrews and Bonta 2010) was developed to help structure assessment and intervention planning for those under correctional supervision, with the goal of reducing re-offense risk. It has three principles. The risk principle requires that re-offense risk be measured accurately, and those at high risk be treated more intensively and longer. The need principle requires targeting criminogenic needs (those associated with the likelihood of re-offending); through improving criminogenic deficits and strengthening protective factors for an individual, the risk of re-offending is reduced. The responsivity principle has two components: tailoring of risk-reducing services to individual capacities (specific responsivity) and providing such services in modalities such as cognitive behavioral or social learning that have been empirically supported with this population (general responsivity). Research has yielded support for seven major dynamic risk factors associated with re-offense risk: antisocial personality pattern, procriminal attitudes, antisocial associates, poor use of leisure time, substance

abuse, problematic marital or family circumstances, and problematic circumstances at school or work (Andrews and Bonta 2010). Several meta-analyses support the applicability of this model to various justice-involved populations (Andrews 2012 (*k* number of tests of treatment = 374); Andrews and Bonta 2010 (*k* controlled experimental tests of the effects of various judicial/correctional treatment interventions on recidivism = 374); Dowden and Andrews 1999 (*k* unique studies of correctional treatment programs for female offenders = 26, with 45 effect sizes), 2000 (*k* primary studies of correctional treatment programs for violent offenders = 35, with 52 tests); Hanson et al. 2009 (*k* studies of sex offender treatment = 23). In this chapter, we specify the number of *relevant* comparisons or studies, *k,* upon which cited meta-analyses were based; this figure is not always the same as the total number of comparisons or studies upon which a meta-analysis was based.

The question of the RNR model's applicability to individuals with behavioral health problems has been addressed partly by a meta-analysis incorporating offenders with and without mental illness (Bonta et al. 1998), which identified primarily the same risk factors for both groups (*k* unique samples = 64). It has also been addressed through an updated meta-analysis (Bonta et al. 2013, using comparisons/"studies" *k* = 126, drawn from 96 unique samples of mentally disordered offenders). In the 1998 study, clinical diagnosis was not associated with re-offense risk. In the 2013 study, the authors concluded that RNR risk/need factors are more predictive of general and violent recidivism than clinical factors (with the exception of antisocial personality/psychopathy). It is reasonable to conclude, therefore, that many offenders (whatever their behavioral health needs) are influenced in their propensity for offending by a set of common influences that are described by RNR. Two other questions arise, however. Is there a small subgroup of justice-involved individuals who are influenced primarily by mental health needs—and on whom RNR-identified criminogenic needs have little influence? Is there a difference in how an intervention should be delivered to two groups when one group has substantial behavioral health needs (e.g., major mental illness) and the second group does not, but both share the risk factor of substance abuse? These questions are discussed further in subsequent sections.

Clinical Versus Criminogenic Interventions: Toward a Partial Resolution

Separate teams of investigators have addressed the question of whether justice-involved individuals with severe mental illness commit offenses that are directly linked to their symptoms. Considering jail detainees in a diversion program (Junginger et al. 2006) and those with mental illness on parole (Peterson et al. 2010), both teams concluded that in only about 10 % of cases was the experience of symptoms of mental illness (e.g., auditory hallucinations, paranoid delusions) directly related to the offending behavior. A review of 18 relevant studies (Skeem et al. 2015b) indicated strong support for using general risk assessment tools to

assess recidivism risk for mentally disordered offenders, and preliminary evidence suggested that cognitive-behavioral interventions are more effective in targeting general risk factors than psychiatric treatment alone. However, they also found no direct support for the applicability of the three core RNR principles in treating mentally disordered offenders. Among high-risk mentally disordered offenders, data from the MacArthur Risk Study indicate that psychosis immediately preceded 12 % of serious acts of violence, but the great majority of the 100 individuals studied (80 %) had exclusively "nonpsychosis-preceded" violence (Skeem et al. 2015a). Both criminogenic and psychotic factors have predictive efficacy in mentally disordered offenders (Prins et al. 2015). However, a detailed investigation of 143 mentally disordered offenders using interviews and record reviews revealed that symptoms rarely directly motivated the offenses (of the 429 crimes coded, 4 % related directly to psychosis, 3 % related directly to depression, and 10 % related directly to bipolar disorder) (Peterson et al. 2014).

These are important findings. They are particularly relevant to the assumption that treatment of only mental health symptoms will reduce re-offense risk among justice-involved individuals with serious mental illness. It led to the recent proposal that the great majority of this group should be rehabilitated using RNR to guide a focus upon criminogenic needs (Eno Louden et al. 2015).

In this chapter, we propose a modified version of this argument. Acknowledging the strong evidence that it is criminogenic needs, not traditional mental health symptoms, that are largely linked to offending behavior within this population, we also note that outpatient treatment (primarily psychotropic medication) significantly reduces the risk of arrest among such individuals (Van Dorn et al. 2013). This suggests (consistent with Prins et al. 2015) that psychotropic medication may have a favorable impact on *both* symptom reduction *and* rearrest risk, and hence should be prominent among the interventions used with this population.

We continue this analysis with a careful review of criminogenic needs identified by RNR. By considering which of these may provide the greatest benefit on both criminogenic and clinical grounds, we approach this criminogenic versus clinical question by removing the dichotomy. What are the criminogenic needs that are also important for behavioral health and recovery? The following sections address this question in the context of cognitive interventions, problem-solving interventions, and behavioral health interventions.

Cognitive Interventions

Cognitive interventions, in contrast to more traditional behavioral approaches, focus on the importance of internal thought processes in regulating and changing behavior (Bandura 1969). The foundation of cognitive therapy was established in 1961 by Albert Ellis, a psychologist who created "rational emotive therapy" based on the theory that thoughts control feelings (Ellis and Harper 1975). Following Ellis' work, psychiatrist Aaron Beck founded and developed modern-day cognitive

therapy. Under Beck's cognitive model, therapists help individuals overcome difficult emotions by teaching them to change unproductive or irrational thoughts that lead to upsetting emotional reactions. By identifying and changing distorted thinking through cognitive therapy, an individual develops skills to modify beliefs and thinking patterns that cause distress (Beck 1976).

In correctional settings, the most common intervention that utilizes cognitive therapy techniques is Cognitive Behavioral Therapy (CBT). CBT interventions involve the use of both cognitive and behavioral methods in rehabilitating offenders. Under the CBT model, an offender is taught to identify and restructure distorted cognitions and develop social skills. To accomplish these goals, a therapist will encourage cognitions that result in positive actions (a cognitive approach) and reinforce positive actions (a behavioral approach) (Milkman and Wanberg 2007). This section of the chapter will focus on the cognitive aspects of CBT.

CBT: Cognitive Components

Practitioners delivering CBT interventions generally focus on two key cognitive elements: automatic thoughts and core beliefs (e.g., Beck 1976; Beck 1979; Burns 1989; Ellis and Harper 1975). Automatic thoughts, also described as "thought habits" (Wanberg and Milkman 1998, 2006), are defined as cognitions that arise instantaneously and unconsciously in response to an outside event (e.g., Beck 1976, 1995, 1996; Freeman et al. 1990; Milkman and Wanberg 2007). Cognitive theorists have identified three types of automatic thoughts: expectations, appraisals, and attributions.

Expectations are cognitions that specific actions will bring about certain results. One type of expectation—efficacy expectation—is particularly relevant to offenders. Efficacy expectations are cognitions about behavioral choices that occur in the face of impending circumstances (Milkman and Wanberg 2007). For instance, if an offender escapes arrest by running from the police, her efficacy expectation in that behavior will be reinforced and she will likely choose to run from police in future similar situations.

An appraisal occurs when an individual evaluates the value of his or her current experiences and reactions to those experiences. Appraisals can often be a source of biased thinking and tend to cause subsequent emotions (Beck 1996). For example, a person with major depressive disorder who is turned down for a job might experience a distorted appraisal such as "I'm worthless," which in turn leads to a feeling of sadness.

Attributions are cognitions about why events occur or why certain actions led to certain results. An individual's belief about the source of his or her own achievements or difficulties arises from attributions (Rotter 1966). For instance, after committing a crime, an offender might think, "I am responsible for robbing that store," demonstrating an internal locus of control. Alternatively, an offender might harbor an external locus of control and think, "If the store had better security, I would not have tried to rob it" (Rotter 1966).

Core beliefs, or underlying assumptions, refer to an individual's long-term thought processes (Seligman et al. 2001). Individuals are generally less conscious of core beliefs than automatic thoughts, and core beliefs tend to be more constant and lasting than automatic thoughts (Seligman et al. 2001). Beck defined core beliefs as schemas that form an individual's automatic cognitions (Beck 1996). Core beliefs tend to guide how an individual interprets and appraises external situations (Milkman and Wanberg 2007). Cognitive therapy is intended to help an individual restructure core beliefs that are distorted or irrational.

Criminal Thinking

Research has shown that cognitive interventions are particularly appropriate for justice-involved individuals. Indeed, CBT is considered one of the most effective treatment options for criminal offenders. Repeat offenders are more prone to experience irrational cognition and engage in what is known as "criminal thinking." According to some (e.g., Walters 2002; Yochelson and Samenow 1976), criminal thinking consists of distorted cognitions that include self-justification, displacement of blame, poor moral reasoning, dominance and entitlement, incorrect interpretations of social cues, and overly optimistic perceptions of reality (see Beck 1999; Lipsey et al. 2007; Walters and White 1989; Yochelson and Samenow 1976). These biased thinking habits may lead an offender to perceive harmless situations as threatening. For instance, an offender may interpret an entirely benign comment as rude or intentionally aggressive (Lipsey et al. 2007). Additionally, offenders engaging in these maladaptive thinking patterns often feel a need for instant gratification, confusing wants with needs. For instance, an offender carrying no money might see a CD in a store and think to himself "I want that CD, so I need to have it right now."

Another tenet of criminal thinking involves the "victim stance," which is a belief that one is being unfairly blamed and rejected by society without reason. This involves a failure to understand that the rejection and blame experienced may be due to one's own antisocial behavior (Lipsey et al. 2007). Many offenders may feel unfairly rejected and adopt a "victim stance" because they were raised within a subculture where such antisocial thought patterns were in fact adaptive. For instance, in a prison or street culture environment, cognitions that would otherwise seem distorted ("if someone is rude to me, I must physically punish him or her, or I will not be respected") may indeed be the best way to survive (Lipsey et al. 2007).

Furthermore, research highlights that offenders often commit "thinking errors" (Barriga et al. 2000; Elliot and Verdeyen 2002; Samenow 1984) and maintain antisocial core beliefs and values (Jennings et al. 1983; Kohlberg 1976). Moreover, the cognitive skill set of offenders is often limited (Andrews and Bonta 2010; Gendreau et al. 1996; Ross and Fabiano 1985). Research has shown that cognitive interventions, and CBT treatments in particular, have been effective in rehabilitating offenders who display such cognitive patterns.

Key Elements of Cognitive Interventions for Offenders

An important assumption of the CBT interventions delivered in a correctional setting is that an offender's distorted thinking is learned (Lipsey et al. 2007). As such, the goal of most CBT interventions employed in correctional contexts is to teach offenders how to recognize the thought processes that led to their criminal behavior (Lipsey et al. 2007).

With an aim of helping offenders hold themselves accountable, the cognitive intervention is generally delivered in a series of steps. First, the therapist helps the offender recognize and self-monitor automatic thoughts. Second, offenders are taught to identify distorted, biased, or irrational thinking patterns and core beliefs. Third, the therapist helps the offender develop techniques to restructure those distorted, biased, or irrational automatic thoughts and core beliefs, through cognitive restructuring (Beck 1995; Dozois and Dobson 2001; Freeman et al. 1990; Leahy and Neary 1997). To successfully deliver such cognitive interventions in a correctional setting, therapists rely on several techniques, including cognitive skills training, cognitive restructuring, and anger management. In addition, therapists will employ supplementary techniques such as relapse prevention, social skills development, and morality training (Lipsey et al. 2007).

Cognitive Skills Training. Cognitive skills training consists of teaching offenders various thinking skills, including interpersonal problem solving, abstract thinking, critical reasoning, causal thinking, goal setting, long-term planning, and perspective taking (Lipsey et al. 2007; Robinson 1995). Therapists often instill these techniques in offenders by engaging in role-play scenarios. Additionally, offenders will be asked to use these cognitive skills in real-life situations that would ordinarily yield distorted thought processes and criminal behavior (Lipsey et al. 2007). This training is aimed at providing offenders with decision-making strategies such as to stop and think prior to acting, come up with alternative resolutions, and consider the consequences of their behavior (Lipsey et al. 2007; Porporino et al. 1991; Robinson 1995).

Cognitive Restructuring. Cognitive restructuring is used to modify distorted or inaccurate cognitive processes that contribute to criminal thinking. Therapists may deliver cognitive restructuring in a variety of forms (Milkman and Wanberg 2007). The "self-talk" cognitive restructuring method consists of stopping distorted thought processes and restructuring maladaptive thinking by planting positive thoughts (McMullin 2000). Other cognitive restructuring interventions focus on problem solving (D'Zurilla and Goldfried 1971; D'Zurilla and Nezu 2001), mood management skills (Beck 1976; Monti et al. 1995), critical reasoning (Ross et al. 1986), and rationally responding to, de-catastrophizing, and scaling emotions (Reinecke and Freeman 2003).

Anger Management. Anger management training entails teaching offenders specific cognitive skills for circumstances likely to cause an angry reaction. To teach these skills, a therapist helps the offender identify automatic thoughts that arise in these situations. Then the therapist helps the offender assess whether those

thoughts are rational. Once irrational thoughts are identified, the therapist and offender work together to replace them with more adaptive ones (Lipsey et al. 2007).

Anger management may also be taught through stress inoculation training (Meichenbaum 1975). In stress inoculation training, offenders are taught to deal with stressful situations by first identifying environmental "triggers" that result in an angry response. Offenders then develop self-statements to use when facing these cues. Such statements (e.g., "Relax, this is not that big of a deal" or "This is not worth losing my temper over") help offenders reframe the situation. Relaxation techniques are also taught as part of stress inoculation training so offenders learn to calm themselves mentally and physically in stressful situations. Stress inoculation training also involves role-play between the therapist and offender so offenders can practice these techniques (Beck and Fernandez 1998; Lochman and Lenhart 1993).

Supplementary Techniques. In addition to the primary cognitive skills training and anger management techniques, CBT interventions often incorporate various supplemental components. These auxiliary techniques may include social development training, moral reasoning education, and relapse prevention skills instruction (Lipsey et al. 2007).

To help offenders develop social skills, therapists will focus on replacing antisocial attitudes and feelings with more prosocial alternatives. Additionally, as part of the social skills training, offenders will often be encouraged to reduce or eliminate any antisocial relationships or associations they may have (Andrews and Bonta 2010).

Moral reasoning education may consist of training offenders to be honest, develop moral accountability in terms of their beliefs, attitudes and behavior, and cultivate respect for the well-being of others and societal rules (Little and Robinson 1986). Additionally, to help an offender develop moral reasoning skills, he or she may be exposed to a moral dilemma and then asked to engage in a discussion about how to handle that dilemma (Goldstein and Glick 1987, 1994). Additionally, some interventions strive to help offenders understand how their previous actions affected their victims (Lipsey et al. 2007).

Relapse prevention training consists of teaching offenders to develop self-management and self-control skills to use in situations that could lead to criminal behavior (Milkman and Wanberg 2007; Marlatt et al. 2002). After developing these skills, offenders should be able to identify high-risk circumstances and use coping skills to prevent criminal relapses (Little and Robinson 1986).

Primary CBT Interventions Used in Correctional Settings

Multiple empirically based CBT programs are used to rehabilitate justice-involved individuals. Below are descriptions of five prototypical CBT programs commonly used in correctional settings.

Reasoning and Rehabilitation. Reasoning and Rehabilitation (R&R) (Ross and Fabiano 1985) is an evidence-based CBT intervention designed to help youths and adults who have antisocial tendencies and exhibit criminal behavior. R&R consists of exercises that change offender thinking that is impulsive, illogical or rigid, and teach them to think of alternative ways of responding that lessens the adverse impact on others, particularly victims. The program consists of 35 structured 2 h sessions for groups consisting of between 6 and 12 individuals (Ross et al. 1988).

Aggression Replacement Training. Aggression Replacement Training (Goldstein and Glick 1987, 1994) is an empirically supported CBT intervention geared toward justice-involved youth with aggression control problems. Aggression Replacement Training consists of three components: Social Skills Training, Anger Control, and Moral Reasoning Education. Social Skills Training, generally delivered through role playing and modeling, teaches offenders to replace antisocial tendencies with prosocial behaviors. Through such training, offenders learn self-control by recording anger-arousing experiences, identifying "triggers," and applying anger control techniques. The Moral Education component consists of exposing offenders to ethical dilemmas and leading a discussion aimed at developing moral reasoning (Goldstein and Glick 1987, 1994).

Thinking for a Change. Thinking for a Change (T4C) (Bush et al. 1997) is a CBT intervention designed for small groups of individuals in a variety of correctional settings, including prisons, jails, probation, and parole supervision. T4C consists of 22 sessions aimed to help participants learn cognitive restructuring skills, problem-solving techniques, and social skills. Each session is comprised of group exercises and homework designed to assist participants develop these skills (Bush et al. 1997).

Cognitive Interventions Program. The Cognitive Interventions Program (CIP) (National Institute of Corrections 1996) consists of 30 sessions designed to deliver a cognitive restructuring curriculum. The program aims to teach offenders to recognize that their behaviors result directly from the choices they make. The lessons help participants recognize distortions and errors in their cognitions, such as victim stance, overoptimism, and a lack of empathy. The program also helps offenders identify antisocial attitudes that may influence their choices and teaches prosocial thinking styles (National Institute of Corrections 1996).

Effectiveness of CBT Interventions for Offenders

A review of the literature on offenders' rehabilitation reveals many beneficial effects of cognitive approaches (Cullen and Gendreau 1989; Gendreau and Ross 1987; Husband and Platt 1993; Landenberger and Lipsey 2005; Wilson et al. 2005). Indeed, research has shown that most effective interventions involve treatment that focuses on the improvement of cognitive functions (Gendreau and Andrews 1990).

One meta-analysis revealed that CBT interventions are critical in delivering effective correctional treatment to adult and juvenile offenders (Andrews et al. 1990 (*k* "intervention appropriateness" versus sanctions comparisons = 154, drawn from 80 studies). Reviews have also shown that cognitive-behavioral programs in particular tend to be the most effective form of rehabilitation for offenders (Allen et al. 2001; Andrews et al. 1990; Cullen and Gendreau 1989; Gendreau and Andrews 1990; Gendreau and Ross 1987; Husband and Platt 1993; Kenne 2011; Losel 1995; MacKenzie and Hickman 1998).

Multiple meta-analyses have demonstrated the effectiveness of CBT interventions in criminal justice settings for both adults and juveniles (e.g., Lipsey and Cullen 2007). In one review of group-structured CBT programs for offenders (*k* studies = 20), researchers found that CBT interventions effectively reduced the criminal behavior of participating offenders (Wilson et al. 2005). Wilson et al. (2005) also identified specific aspects of CBT interventions that were particularly effective, finding that therapy focused on moral reconation, cognitive restructuring, and cognitive skills led to lower recidivism rates. Another meta-analysis of research studies comparing behavioral and cognitive-behavioral interventions (*k* behavioral programs and cognitive-behavioral programs combined = 69) found cognitive-behavioral treatments more effective in terms of reducing recidivism rates (Pearson et al. 2002).

Several meta-analyses dedicated solely to CBT interventions for offenders have also demonstrated the overall effectiveness of such programs in terms of reducing recidivism (Landenberger and Lipsey 2005 (*k* CBT studies = 58); Lipsey et al. 2001 (*k* cognitive-behavioral-program studies = 14); Pearson et al. 2002; Tong and Farrington 2006 (*k* Reasoning and Rehabilitation evaluation studies = 16, from 26 comparisons). One of these meta-analyses found that the CBT interventions reduced offender recidivism—and that the decrease in recidivism was primarily due to the CBT interventions rather than the solely behavioral techniques (Pearson et al. 2002). Another found that, after controlling for various methodological variables, CBT treatment programs that included any one-on-one treatment, anger control, or cognitive restructuring were significantly associated with better recidivism outcomes in bivariate analyses (Landenberger and Lipsey 2005). Other CBT elements (cognitive skills, interpersonal problem solving, social skills, moral reasoning, victim impact, substance abuse, behavioral modification, and relapse prevention) did not show a significant bivariate relationship with effect size. However, in a multivariate analysis (also comprised of methodological controls, participant risk level, amount of treatment, quality of treatment, and whether a program emphasized CBT or not), only anger control and interpersonal problem-solving elements were significantly and positively associated with outcome, while victim impact and behavioral modification were both significantly negatively related to effect size. Landenberger and Lipsey (2005) also examined their sample of studies in terms of the name-brand CBT that was used (Reasoning and Rehabilitation, Moral Reconation Therapy, Aggression Replacement Therapy, Interpersonal Problem-Solving Therapy, Thinking for a Change, programs focused on substance abuse, other manualized programs, all others programs). They found that none of

the name-brand programs stood out from the pack—i.e., none were significantly associated with effect size in bivariate or multivariate analyses. To summarize, Landenberger and Lipsey's (2005) primary findings relevant to this section of the chapter were that (a) anger management training, in particular, showed a robust relationship with better recidivism outcomes, and (b) no name-brand programs were distinguishable from the others in terms of larger reductions in recidivism.

Future Directions

With the growing incarceration rate, continued problem of criminal recidivism, and public funding of rehabilitation for offenders, the importance of analyzing the effectiveness of offender rehabilitation interventions is clear. Current research supports the effectiveness of cognitive interventions, but additional study is needed to clarify what works in correctional contexts. Such research will allow practitioners to tailor interventions to meet the needs of specific offender populations.

Some studies have examined the effects of CBT interventions in relation to various factors, such as offender risk level and correctional setting (e.g., Tong and Farrington 2006), but additional research in these areas would be useful. To deliver an intervention most effectively, a clinician must know the techniques that work best for high-risk versus low-risk offenders. Additionally, interventions need to be considered in correctional context, as offenders may be in maximum security prisons, rehabilitation centers, or under supervision in the community. Moreover, studies should explore the relationship between the effectiveness of a cognitive intervention and other variables, such as previous recidivism rates, socioeconomic status, age, and IQ level.

Some evidence has shown that race and gender may influence the effectiveness of a CBT intervention (Kenne 2011), but additional research is needed to clarify the extent of these relationships. The findings of such research would likely be useful in developing and disseminating interventions tailored to specific offender populations. Such a nuanced approach would help practitioners develop more effective interventions for offenders and potentially further reduce recidivism rates (Kenne 2011).

Finally, some have criticized the quality of the current research examining the extent to which certain interventions lead to positive change among offenders (Telander 2005). These criticisms note the lack of random sampling and inadequate sample sizes in many of the published studies. As such, future researchers should aim to obtain adequately sized samples, and use appropriate treatment and control groups.

Problem-Solving Interventions

In this section, we review problem-solving deficits and related interventions among offender populations. Our coverage is limited to deficiencies and remedies of offenders' *personal* problem-solving skills, at the level of cognitions, affect, and behavior. We do not address problem-solving courts (also known as specialty or treatment/therapeutic courts), such as drug courts, mental health courts, and veteran courts. These types of courts seek to apply legal leverage to engage preselected defendants in interdisciplinary treatment for psychosocial problems related to a court's specialty area, which are thought to underlie eligible court participants' criminal behavior (Daicoff and Wexler 2003; Griffin et al. 2015; Wiener and Brank 2013). Although these courts fall outside the scope of this section, it should be noted that specialty court referrals to treatment often involve interventions that target defendants' personal problem-solving deficits.

Nature of the Intervention

There are two major strands in the offender problem-solving literature (McGuire 2005a; McMurran 2005). The first is the *training* strand from correctional psychology, characterized by "name-brand" programs often based on the interpersonal cognitive problem-solving model (e.g., Spivack et al. 1976). One such program is *Reasoning and Rehabilitation* (R&R), a widely disseminated, manualized, multifaceted intervention program for offenders that includes training in social problem solving (Ross and Ross 1995; D'Zurilla and Nezu 2007). The second is the *therapy* strand, which arises out of the social problem-solving model from clinical psychology (e.g., D'Zurilla and Goldfried 1971). Examples include Problem-Solving Therapy for sex offenders (e.g., Nezu et al. 2005a) and the *Stop & Think* program for offenders with personality disorders (e.g., McMurran et al. 2001b).

Different problem-solving programs, and problem-solving components of more comprehensive intervention programs, have been developed to address the needs of different offender populations. These programs/components vary somewhat in their duration, intensity, format, level of emphasis and elaboration, and content. McGuire (2005a) discusses some of these differences. Generally, however, all of these programs share a common presumption that poor problem solving is associated with psychosocial adjustment problems, including offending behavior. All problem-solving training approaches therefore seek to help offenders develop skills for responding more adaptively to problems in their lives and, hopefully, avoid reoffending as a result.

Regarding treatment targets, Problem-Solving Therapy and the social problem-solving model, for instance, conceptualize a positive versus negative overall orientation to problems, and different problem-solving styles (rational/planful, impulsive/careless, and avoidant) (D'Zurilla and Nezu 2007; Nezu

et al. 2009). More adaptive functioning is associated with a positive problem orientation—a general tendency to perceive problems as personally solvable challenges, and to commit to expending the time and effort to address versus avoid problems. Obstacles to adopting a positive outlook on problems include low self-efficacy beliefs, negative thinking, and negative emotions. Furthermore, both greater adaptive functioning and a positive problem orientation are associated with a rational/planful problem-solving style—the effective problem-solving skills of accurately defining/formulating problems, generating numerous potential solutions, making decisions based on pros and cons of various alternative solutions, and implementing and monitoring a solution. As for intervention techniques, therapeutic/training programs in problem solving incorporate a variety of cognitive-behavioral change strategies, including functional analysis, psychoeducation, practice and feedback, self-monitoring, training in affect regulation, and use of cognitive restructuring techniques (McGuire 2005a). Problem-solving therapy/training appears to occupy an intermediate position on a continuum ranging from cognitively oriented therapies (cognitive therapy, schema-focused therapy) to behaviorally oriented therapies (self-instructional training, social skills training, behavior therapy, behavior modification) (McGuire 2005a).

Empirical Research

In their review of the empirical literature on social problem solving and adjustment, D'Zurilla and Nezu (2007) included studies on offending behavior. They summarized the research findings among adolescents as follows:

- First, adolescents demonstrated less effective problem solving than their parents;
- Second, a negative problem orientation among adolescents was associated with aggression, an impulsive problem-solving style with delinquency, and an avoidant style with aggression and delinquency;
- Third, poorer social problem-solving skills were associated with increased problem behaviors (drug use, delinquency, running away, and unprotected sex) among high-risk (delinquent, runaway, or disadvantaged) teenagers; and
- Fourth, youthful offenders with heroin use disorders performed less effectively on a problem-solving measure than nonaddicted youth offenders (Greening 1997; Jaffe and D'Zurilla 2003; Leadbeater et al. 1989; Platt et al. 1973; for a review of a few additional studies of delinquent youths' problem solving, see Antonowicz and Ross 2005 and McGuire 2005a).

Researchers have also examined social problem-solving abilities in aggressive versus nonaggressive men with mild intellectual disabilities, concluding that more aggressive individuals had greater problem-solving deficits (Basquill et al. 2004). The aggressive participants also identified interpersonal intent less accurately and generated aggressive solutions more frequently.

With respect to adult offenders, high neuroticism has been related to poor problem solving in a sample of offenders with mental illnesses (including personality disorders) (McMurran et al. 2001a). Other researchers investigated social problem solving in a sample of offenders convicted of child sex offenses (Nezu et al. 2005b, c). Child sexual offenders differed from the general population with respect to negative problem orientation and impulsive/careless problem-solving style, both of which were also related to self-reported sexual deviance. In contrast, only scores on a scale of avoidant problem-solving style added unique predictive validity (beyond demographics and prior abuse as a child) to a clinician-rated measure of sexual aggression.

Antonowicz and Ross (2005) noted two studies that found limited problem-solving skills among inmates with poor adjustment to prison (Higgins and Thies 1981; Zamble and Porporino 1988). Consistent with this finding, other researchers have observed that among young offenders, poorer social problem-solving abilities in young offenders were associated with increased distress and suicidality, and more victimization by bullies (Biggam and Powers 1999a–c). Adult prisoners on a special unit for emotionally distressed persons evidenced poorer social problem solving relative to a comparison group drawn from the general offender population (Hayward et al. 2008). A negative problem orientation has also been found to predict self-reported depression and anxiety in adult prisoners (McMurran and Christopher 2009). Finally, McGuire (2005a) interpreted the results of a study by Zamble and Quinsey (1997) as indicating that poor problem solving often preceded recidivism in offenders with extensive criminal histories.

Since the McMurran and McGuire (2005) and D'Zurilla and Nezu (2007) reviews, a small number of new nonintervention studies have been published on offenders' problem-solving abilities. The studies examined the psychometrics of short and long versions of the Social Problem-Solving Inventory–Revised (SPSI–R; D'Zurilla et al. 2002). Both found the measure to be sufficiently reliable and valid to support its use with the respective populations studied—sexual offenders (Wakeling 2007) and offenders with intellectual disabilities (Lindsay et al. 2011). In a sample of untreated sexual offenders, Barnett and Wood (2008) found that participants reported an overall average current problem-solving ability on the SPSI–R. They described a need for future studies to evaluate whether sexual offenders use their problem-solving skills toward prosocial ends (to guard against antisocial tendencies).

D'Zurilla and Nezu (2007) also reviewed outcome studies of problem-solving interventions for offenders, frequently summarizing relevant chapters from an earlier review (McMurran and McGuire 2005). They discussed four intervention programs: *R&R*, *Think First*, *Stop & Think*, and *Project CBT/STOP*. *R&R* has been the most extensively studied problem-solving program for offenders; Antonowicz (2005) identified 22 studies of varying methodological quality that evaluated *R&R* implementations in institutional and community settings. Excluding studies with single-group designs, the majority of between-subjects studies (11 out of 18) lent support to *R&R*'s effectiveness for reducing recidivism, while the remaining 7 produced mixed results. Quantitative reviews of *R&R* have produced consistent

results. Based on a sample of seven *R&R* studies, Pearson et al. (2002) reported a 26 % reduction in recidivism for program participants relative to controls. Subsequently, Tong and Farrington (2006) reported a weighted mean odds ratio of 1.16 for reconvictions based on 16 controlled or quasi-controlled studies (yielding 26 comparisons), or a 14 % overall decrease in recidivism among *R&R* participants compared to controls. Additional moderator analyses supported the effectiveness of the program in institutional and community settings, as well as for low- and high-risk offenders. Another meta-analysis of CBT interventions for offenders, which included name-brand programs as a moderator, found that *R&R*'s effectiveness was comparable to other name-brand and generic CBT programs [Landenberger and Lipsey 2005; Lipsey et al. 2007 (*k* studies of name-brand programs = 58)].

The next program reviewed by D'Zurilla and Nezu (2007), *Think First*, had not been subjected to a randomized controlled trial at the time of their review. However, the few repeated measures studies, summarized by McGuire (2005b), were promising with respect to short-term change on many relevant psychological measures, including measures of social problem solving, criminal thinking, and impulsivity (McGuire and Hatcher 2001; Ong et al. 2003; Steele 2002a). A more recent study that used a pre-post design found that completion of *Think First* improved impulsivity levels, criminal thinking styles, locus of control, and some aspects of social problem-solving ability, although an unexplained increase in impulsive/careless problem-solving style was also observed at the end of treatment (Burgoyne and Tyson 2013).

McGuire (2005b) also reviewed follow-up recidivism studies that used matched-group designs. The first follow-up study found that program-completing probationers performed poorly compared to nontreated prisoners (Debidin and Lovbakke 2005; Stewart-Ong et al. 2004). However, subsequent research has produced more promising results. One study found that program-completing probationers had lower reconviction rates than non-completers and probationer controls, especially completers who originally met selection criteria for the program (Steele 2002b). Another study found that the program had a large effect on reconviction rates for program-completing probationers versus both non-completers and those who were referred to the program but never started it (Stewart-Ong et al. 2004). Subsequent work that considered risk for recidivism found group differences on reconviction rates at up to 18-months follow-up, and also that risk moderated the association between program completion and reconviction (Roberts 2004). A large retrospective evaluation of the program found that, controlling for demographics, risk, and offense type, completing the program was associated with a 29 % reduction in reconviction compared to non-completers and probationer controls, while failing to complete the program was associated with a 22 % increase in reconvictions relative to program completers and probationer controls (Hollin et al. 2002). Research on attrition from *Think First* has suggested that the program may require an overly demanding reading level (Davies et al. 2004), and that compared to program completers, non-completers tended to have shorter criminal histories, were less likely to accept responsibility for their offending, reported stress as a

problem less often, and more frequently acknowledged problems with self-esteem (Westmarland et al. 2002).

The third program reviewed by D'Zurilla and Nezu (2007), *Stop & Think*, had yielded little empirical data by 2007. McMurran et al. (2005) reported both pilot data and preliminary data from a larger study that showed improvements in overall social problem solving and certain problem-solving styles (e.g., impulsivity) among offenders with mental illnesses or personality disorders participating in the program. More recent research on *Stop & Think* has found that treatment completers tended to have greater rational problem-solving styles and lower impulsive/careless styles compared to treatment non-completers (McMurran et al. 2008). Another study found pre-post differences supporting the effectiveness of the program for reducing emotional distress among adult prisoners on a special mental health unit (Hayward et al. 2008).

Finally, D'Zurilla and Nezu (2007) described a demonstration program for sexual offenders with intellectual disabilities called Project CBT/STOP, which consisted of Problem-Solving Therapy combined with other CBT approaches (Nezu et al. 2005a, b). A single-group pre-post evaluation of the program (Nezu et al. 2005c, 2006) found an increase in adaptive behavior and motivation for participants. Positive trends were also seen for pre-post changes in clinical target behaviors, and the posttreatment recidivism rate over a 3-year outcome period was 4 %.

Research conducted since the McMurran and McGuire (2005) and D'Zurilla and Nezu (2007) reviews has found that completing either *R&R*, an abbreviated and modified version of *R&R* (*Enhanced Thinking Skills*), or *Think First* was not associated with a reduction in reconviction compared to program non-completers, referrals who did not start the programs, and offenders who were not assigned to any programs (McGuire et al. 2008). However, in stratifying the groups by risk level for non-participation/non-completion, a small treatment effect was found for moderate- and high-risk cases. Furthermore, a larger evaluation of the three programs found that each was associated with a similar reduction in recidivism for program completers versus nonstarters, non-completers, and non-referrals (Hollin et al. 2008). Another analysis of these programs found that risk of recidivism was associated with non-participation/non-completion (Palmer et al. 2009). Although very high-risk individuals were more likely to fail to follow up on a referral or drop out of treatment, a large recidivism treatment effect was observed when these offenders completed one of the programs—while no treatment effect was observed for low-risk individuals. Similar findings had been reported earlier with a larger sample (Palmer et al. 2008). Finally, among female offenders, completion of *Think First* or *Enhanced Thinking Skills* was not associated with a reduction in recidivism compared to a non-referral comparison group, although non-completers had a higher rate of recidivism relative to the comparison group (Palmer et al. 2015).

In addition to the programs reviewed thus far, Black et al. (2008) reported the results of a small pilot study of a 20-week supplemental group treatment for incarcerated female offenders diagnosed with borderline personality disorder, which included didactic sessions on problem-solving skills. The program was associated with medium to large effect sizes for relevant mental health outcomes (disorder

severity, negative affectivity, and depression), although the researchers did not isolate the effects of the program's problem-solving components. Similar results were obtained in a subsequent study using a similar but larger sample, in addition to new findings that the program was associated with reduced suicidal behavior and disciplinary infractions—and that baseline symptom severity was negatively related to improvement in the program (Black et al. 2013). In another small pilot study of a treatment program based on the social problem-solving model for offenders with intellectual disabilities, Lindsay et al. (2011) reported positive treatment effects on problem orientation and problem-solving style using a multiwave study design. A brief problem-solving intervention was also evaluated in male and female diverted and incarcerated samples (Spiropoulos et al. 2005). Relative to corresponding comparison groups, many of the treated groups showed reductions in correctional misconducts and depression scores, although no effects were observed on measures of conflict and employment.

Rather than focusing only on specific name-brand CBT programs for offenders, Landenberger and Lipsey (2005) and Lipsey et al. (2007) meta-analyzed the effects of specific CBT elements included in various treatment programs. In a multivariate model that included methodological controls, risk level, treatment duration and quality, whether CBT was emphasized, and other CBT elements (cognitive skills, cognitive restructuring, social skills, moral reasoning, victim impact, substance abuse, behavior modification, relapse prevention, and individual attention), they observed that, of the CBT elements, larger reductions in recidivism were associated only with programs that included attention to interpersonal problem solving or anger control. In a similar multivariate model that examined interventions at the level of the named type of CBT that was used, *R&R* and "Interpersonal Problem Solving" performed comparably to all of the other CBT protocols that were examined (Moral Reconation Therapy, Aggression Replacement Therapy, Thinking for Change, and substance abuse focused programs).

Research Gaps/Areas for Future Research

McMurran (2005) has identified a number of future directions for research on problem solving and offending. Her suggestions included focusing on individual differences in social problem-solving styles, and how these differences develop and relate to personality traits, information processing, and offending. She also recommended that researchers explore the links between social problem solving and emotion, examining constructs such as affect regulation and emotional intelligence. In addition, she suggested that concepts from cognitive therapy, such as automatic thoughts and cognitive schemas, might influence social problem solving. As such, it might be beneficial to examine these constructs as part of research on social problem solving. Likewise, she suggested investigating the influence of cognitive therapies for personality disorders, such as schema therapy, on social problem solving. Finally, she recommended improving the measurement of social problem

solving in offenders, given the problems with the reading level and self-awareness required for self-report process measures like the SPSI–R, and with the subjectivity and limited external validity of outcome measures such as the Means–End Problem-Solving procedure.

Since 2005, research has (1) investigated measures such as the SPSI–R in offender samples; (2) documented the social problem-solving profiles of different types of offenders; (3) investigated the links between social problem-solving skills and constructs from offender rehabilitation theories (e.g., primary good/needs from the Good Lives Model; e.g., Ward et al. 2012); (4) examined the relationship between social problem solving and emotional adjustment of offenders while incarcerated and evaluated outcomes for preexisting problem-solving training programs; and (5) developed and evaluated new training programs for social problem solving. Nevertheless, the amount of new research that has been generated since McMurran's (2005) proposed research agenda has been modest. More is needed, particularly on the development of more reliable and valid assessment methods–as well as research that investigates the developmental pathways of social problem solving, and the interactions among affect, cognition, and neurocognitive functioning, personality, and social problem solving. Component and dismantling studies would also be useful for isolating the effect of training in social problem solving versus other treatment targets and strategies.

Behavioral Health Interventions

The following section will consider evidence-based interventions for substance abuse, trauma-related disorders, and serious mental illness (SMI) that have demonstrated effectiveness in reducing associated clinical symptoms and recidivism. As a thorough review of the evidence base for all behavioral health interventions is beyond the scope of this chapter, readers are referred elsewhere (e.g., McMurran 2009; Young and Thome 2011) for such a review.

Substance Abuse

The prevalence of substance abuse among offenders is substantially higher than that of the general population. Current estimates indicate that the prevalence for offenders is between 10 and 48 % for males and between 30 and 60 % for females (Fazel et al. 2006; Lynch et al. 2014). Substance abuse has also been shown to be associated with criminal recidivism (Dowden and Brown 2002). Although there are numerous interventions designed to reduce substance use problems that also have demonstrated criminal justice applications (e.g., motivational interviewing; McMurran 2009; case management; Siegal et al. 2002), the three treatment approaches that will be described here are therapeutic communities, cognitive-behavioral interventions, and "blended approaches" that combine therapeutic community and cognitive-behavioral techniques.

Therapeutic Communities. Derived from the self-help tradition, therapeutic communities (TCs) were initially developed for use with non-correctional and community-based populations (Wexler 1995). As support for prison-based treatment grew, the original TC model was adapted for use in secure settings and is currently one of the most successful in-prison treatment approaches (Prendergast et al. 2003; U. S. Department of Health and Human Services 2005). Although implemented in a residential setting, a number of characteristics distinguish TC programs from those that identify simply as "residential." First, TCs are a "total treatment environment," implemented in isolation from the general prison population (Inciardi et al. 2004, p. 90). Separation from the anxiety, fear, and violence that tend to permeate prison culture provides individuals with an opportunity to begin creating positive change toward a drug free lifestyle (U.S. DHHS 2005). TC programs are highly structured and hierarchical in nature, with an emphasis on improving one's standing in the world, having participants help one another, and actively participating in all community activities. Peers are stratified by level of seniority and are supervised by clinical staff, many of whom are former substance users. Of paramount importance is the recognition that drug use is a disorder of the whole person, and that addiction is one component of a much larger issue (De Leon 1989).

Implementation typically occurs in three stages. Following orientation to the TC rules, residents engage in prison-based individual and group therapy for at least 12 months to further their recovery. The active treatment phase is first followed by work release in a transitional TC setting and then by an aftercare phase in which individuals participate in community-based counseling programs to maintain the progress made in treatment (Inciardi et al. 2004).

The success of TC programming in reducing rates of drug relapse and criminal reoffending has been examined. Meta-analyses [Holloway et al. 2006 (k TC program studies = 7); Mitchell et al. 2012 (k TC program studies = 35)] have supported the effectiveness of TC programs in reducing recidivism and relapse in comparison to other prison-based drug treatment programs (e.g., boot camp, group counseling, narcotic maintenance) and community-based programs such as post-release supervision and drug courts. Results from the Correctional Drug Abuse Treatment Effectiveness (CDATE) project, an extensive systematic review and meta-analysis that examined the effects of programs within 20 intervention categories (k TC program studies = 7), supported the effectiveness of TC programs in reducing recidivism for incarcerated offenders with histories of substance abuse (Pearson and Lipton 1999). Likewise, outcome data from recent controlled trials of TC interventions are consistent with previous meta-analytic findings (Olson and Lurigio 2014). However, the success of TC programming may depend on the outcome variable under consideration. It has been suggested that prison-based drug treatment may exert stronger effects on reducing reincarceration than either rearrest or drug relapse (Welsh and Zajac 2013).

Behavioral and Cognitive-Behavioral Approaches. The underlying theory of cognitive and behavioral interventions and their use with offenders has been described earlier in this chapter. Substance abuse treatment programs in correctional

settings that employ cognitive and behavioral interventions assist offenders in changing criminal beliefs and values, and may utilize systems of reward and punishment to promote prosocial behavior. Such interventions include social skills training, criminal thinking programs, relapse prevention, and contingency contracting (U.S DHHS 2005). Results from the CDATE meta-analysis support the use of cognitive-behavioral interventions with justice-involved substance abusers (Pearson et al. 2002; *k* behavioral and cognitive-behavioral programs combined = 69). However, the effect of combined cognitive-behavioral approaches on recidivism was greater than the effect of solely behavioral interventions.

"Blended" Approaches. TC programs regard substance abuse as the result of disturbed social functioning and personality development, whereas cognitive-behavioral approaches consider substance abuse to be the outcome of prior learning experiences, thoughts, and emotions (Malinowski 2003). However, the two approaches are not incompatible. The Residential Substance Abuse Treatment (RSAT) program, located at the South Idaho Correctional Institution (SICI), has successfully incorporated both cognitive-behavioral and 12-step programming within the context of a TC environment. Cognitive-behavioral components of the program include group process, psychoeducation regarding the nature of the relationship between thinking and behavior, and cognitive restructuring of thinking errors via self-monitoring (Stohr et al. 2001). Since its inception at SICI in 1997, RSAT programming has been implemented within the majority of Idaho prisons (U.S. DOJ 2005) and has been shown to significantly delay time to first rearrest following release from prison (Jensen and Kane 2010).

Trauma-Related Disorders

Trauma has been defined as "any form of interpersonal or domestic physical, sexual, or emotional abuse or neglect, which is sufficiently detrimental to cause prolonged physical, psychological, or social distress to the individual" (Moloney et al. 2009, p. 427). Rates of exposure to trauma are higher among incarcerated populations than in the general population (Messina and Grella 2006). This is compounded by the reality that prison itself can also be traumatizing (Arrigo and Bullock 2008; Wallace et al. 2011). Estimates place the prevalence of PTSD among male and female offenders between 4 and 21 %, with female offenders disproportionately affected (Goff et al. 2007). Indeed, the current prevalence of PTSD within exclusively female samples (48.2 %) is much higher (Zlotnick 1997). The high prevalence of PTSD symptoms is concerning not only from a mental health perspective but from a public safety perspective as well, considering that PTSD is related to an increased risk of reoffending (Ardino et al. 2013). Thus, trauma-informed interventions such as Seeking Safety (see next section) designed to address the complex needs of incarcerated men and women (e.g., high comorbidity of PTSD and substance abuse; Lynch et al. 2014) might be helpful in reducing risk. Recent research has yielded promising results, as discussed next.

Trauma-Informed and Gender-Responsive Therapeutic Communities. At its core, trauma-informed care involves the provision of services that "do no harm" (i.e., that do not inflict further trauma or reactivate past traumatic experiences). Other facets of this approach include training staff to be aware of the impact of trauma and accurately identifying trauma symptoms (Hodas 2006). Utilizing these principles, in conjunction with a gender-responsive curriculum that emphasizes the role that relationships and intimate partners play in women's addiction and recovery (Helping Women Recover; Covington 1999), Messina et al. (2010) examined the effectiveness of a trauma-informed and gender-responsive TC in reducing drug use and rates of reincarceration. In comparison to a standard prison-based TC, women in the gender-responsive condition had greater reductions in drug use and were less likely to have been reincarcerated within 1 year after being released on parole.

Seeking Safety. Seeking Safety (SS) is a present-focused, manualized cognitive-behavioral intervention for individuals with comorbid PTSD and substance abuse (Najavits 2009). The program can be implemented in any clinical setting (e.g., inpatient, outpatient, criminal justice) and is based upon five principles (National Center for Trauma-Informed Care 2008). The first principle involves establishing one's safety, which is defined broadly to include discontinuing drug use, decreasing the frequency of self-harm, and ending abusive relationships. Principles 2 and 3 emphasize the importance of an integrated approach to treating co-occurring PTSD and substance use. Principles 4 and 5 highlight the program's four content areas—cognitive, behavioral, interpersonal, and case management—and clinician processes such as compassion and ensuring that the client has as much control as possible (Najavits n.d.). Treatment components include psychoeducation regarding the relationship between trauma and drug use and 25 cognitive-behavioral and/or interpersonally oriented treatment topics (e.g., Setting Boundaries in Relationships, Asking for Help) that introduce new coping skills (NCTIC 2008; Lynch et al. 2012). Preliminary data support the effectiveness of SS in reducing substance use and symptoms of PTSD in samples of incarcerated women (Lynch et al. 2012; Zlotnick et al. 2003) and a trend for women in SS to be less likely to return to prison over 6 months post-release (Zlotnick et al. 2009). To date, one study has examined the feasibility and preliminary efficacy of implementing SS among male offenders, and it found promising results for substance use and mental health outcomes (Barrett et al. 2015).

Serious Mental Illness

Considering previously noted evidence suggesting that offending is influenced by the presence of mental illness for a small subgroup of offenders, the provision of mental health services for these individuals is warranted on both clinical and risk reduction grounds. Interventions in this section were selected for inclusion in an effort to represent alternatives from both the criminal justice and mental health

models of treatment (Skeem et al. 2011). These interventions, to be discussed next, include Modified Therapeutic Community and Assertive Community Treatment.

Modified Therapeutic Community. Based on the theoretical framework of the TC model, modified therapeutic community (MTC) programs were adapted to meet the needs of individuals with co-occurring substance use and mental illness. MTC programs retain an emphasis on personal responsibility, mutual self-help, and using peers as role models (Sacks et al. 2010). Relative to traditional TCs, MTCs are less intense (i.e., shortened and simplified meetings, fewer activities), allow for greater individualization, and involve smaller caseloads (Sacks et al. 1999). In addition, the MTC contains a medication education component (U.S. DHHS 2005). Results from controlled trials and a recent meta-analysis favor MTC programs over comparison groups in reducing substance use and criminal activity (Sacks et al. 2004, 2010; k studies = 3, with 4 comparisons).

Forensic Assertive Community Treatment. Derived from Assertive Community Treatment (ACT), an intensive community-based treatment that provides care to individuals with mental illness, Forensic Assertive Community Treatment (FACT) is designed to prevent jail detention and reduce criminal recidivism (Morrissey et al. 2007). Program components such as targeting of individuals with prior arrests, use of court sanctions to incentivize participation, and inclusion of probation officers on treatment teams differentiate FACT from ACT (Lamberti et al. 2004). Although the evidence for FACT effectiveness has historically been weak (Morrissey et al. 2007), as it had been limited to a small number of pre-post designs, a recent randomized clinical trial that compared FACT with treatment as usual (TAU) yielded encouraging results (Cusack et al. 2010). At 12 months, FACT participants had fewer jail bookings and hospital stays than did individuals in the TAU condition. This same pattern was seen at 24 month follow-up, highlighting the potential value of providing targeted behavioral health services to criminal justice populations.

Future Directions

Evidence supporting the utility of a variety of behavioral health interventions in reducing clinical symptoms related to substance use, trauma, and severe mental illness, and the impact of these interventions on criminal recidivism, has been discussed in this section. However, our current knowledge on what constitutes best practice is limited by the difficulty in implementing randomized controlled trials (RCTs) in research, and by the absence of sufficient recidivism data.

RCTs are undertaken relatively infrequently within the criminal justice system, and researchers tend to use quasi-experimental designs (Perry et al. 2009). Although there are a variety of ethical, legal, and practical challenges associated with using RCTs with justice-involved individuals, the very limited use of this research design within criminal justice is unfortunate, given the susceptibility of quasi-experimental designs to sources of bias (Farrington and Welsh 2005; Perry

et al. 2009). Increased use of RCTs would strengthen the research base underlying the interventions discussed in this chapter, and improve our ability to draw causal conclusions.

Finally, the empirical investigations described in this section reported accompanying recidivism data, but many studies examining the effectiveness of behavioral health interventions do not. Although it may seem unnecessary to collect recidivism data when the primary goal of most behavioral health interventions is to reduce the number and/or severity of clinical symptoms associated with the disorder in question, we have strongly argued that it is not only necessary but essential when working with justice-involved populations. Collecting such outcome data would allow investigators to determine the clinical and the risk-reducing impacts of various interventions, and facilitate the move toward an integrated model of service delivery that prioritizes both but funds them simultaneously.

Conclusion

This chapter has reviewed the evidence regarding interventions for justice-involved individuals. The discussion concerning the most appropriate interventions for such individuals has insufficiently considered the primary goal of the intervention. In a justice-involved population, it is very important to focus in particular on interventions that have been shown to reduce re-offense risk. Given the apparently high prevalence of behavioral health disorders in this population, it is particularly important to consider the efficiency and value of interventions that can both improve behavioral health and reduce re-offense risk. Future research that improves knowledge regarding such "dually valuable" interventions is clearly indicated. Such research may be guided by the RNR model, which has been shown to be effective in identifying risk-relevant needs and gauging responsivity for the larger population of justice-involved individuals.

References

Allen, L. C., MacKenzie, D. L., & Hickman, L. J. (2001). The effectiveness of cognitive behavioral treatment for adult offenders: A methodological, quality-based review. *International Journal of Offender Therapy and Comparative Criminology, 45*, 498–514.

Andrews, D. (2012). The Risk-Need-Responsivity (RNR) model of correctional assessment and treatment. In J. Dvoskin, J. Skeem, R. Novoaco, & K. Douglas (Eds.), *Using social science to reduce violent offending* (pp. 127–156). New York: Oxford University Press. doi:10.1093/acprof:oso/9780195384642.001.0001

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does treatment work? A psychologically informed meta-analysis. *Criminology, 28*, 369–404.

Andrews, D., & Bonta, J. (2010). *The psychology of criminal conduct* (5th ed.). New Providence, NJ: LexisNexis.

Antonowicz, D. H. (2005). The reasoning and rehabilitation program: Outcome evaluations with offenders. In M. McMurran & J. McGuire (Eds.), *Social problem solving and offenders: Evidence, evaluation and evolution* (pp. 163–182). Chichester, England: Wiley. doi:10.1002/9780470713488.ch9

Antonowicz, D. H., & Ross, R. R. (2005). Social problem solving deficits in offenders. In M. McMurran & J. McGuire (Eds.), *Social problem solving and offenders: Evidence, evaluation and evolution* (pp. 91–102). Chichester, England: Wiley. doi:10.1002/9780470713488.ch5

Ardino, V., Milani, L., & Di Blasio, P. (2013). PTSD and re-offending risk: The mediating role of worry and a negative perception of other people's support. *European Journal of Psychotraumatology, 4*, N. P. doi:10.3402/ejpt.v4i0.21382

Arrigo, B. A., & Bullock, J. L. (2008). The psychological effects of solitary confinement on prisoners in supermax units: Reviewing what we should know and recommending what should change. *International Journal of Offender Therapy and Comparative Criminology, 52*, 622–640. doi:10.1177/0306624X07309720

Bandura, A. (1969). *Principles of behavior modification*. New York: Holt, Rinehart & Winston.

Barnett, G., & Wood, J. L. (2008). Agency, relatedness, inner peace, and problem solving in sexual offending: How sexual offenders prioritize and operationalize their good lives conceptions. *Sexual Abuse, 20*, 444–465. doi:10.1177/1079063208325202

Barrett, E. L., Indig, D., Sunjic, S., Sannibale, C., Sindicich, N., & Mills, K. (2015). Treating comorbid substance use and traumatic stress among male prisoners: A pilot study of the acceptability, feasibility, and preliminary efficacy of Seeking Safety. *International Journal of Forensic Mental Health, 14,* 45–55. doi:10.1080/14999013.2015.1014527

Barriga, A. Q., Landau, J. R., Stinson, B. L., Liau, A. K., & Gibbs, J. C. (2000). Cognitive distortion and problem behaviors in adolescents. *Criminal Justice and Behavior, 27*, 36–56.

Basquill, M., Nezu, C. M., Nezu, A. M., & Klein, T. L. (2004). Aggression-related hostility bias and social problem solving deficits in adult males with mental retardation. *American Journal on Mental Retardation, 109*, 255–263. doi:10.1352/0895-8017(2004)109<255:AHBASP>2.0.CO;2

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.

Beck, A. T. (Ed.). (1979). *Cognitive therapy of depression*. Guilford Press.

Beck, A. T. (1996). Beyond belief: A theory of modes, personality, and psychopathology. In P. M. Salkovskis (Ed.), *Frontiers of cognitive therapy* (pp. 1–25). New York: Guilford.

Beck, A. T. (1999). *Prisoners of hate: The cognitive basis of anger, hostility, and violence*. New York: Harper Collins.

Beck, J. S. (1995). *Cognitive therapy: Basics and beyond*. New York: Guilford.

Beck, R., & Fernandez, E. (1998). Cognitive-behavioral therapy in the treatment of anger: A meta-analysis. *Cognitive Therapy and Research, 22*, 63–74.

Biggam, F., & Power, K. (1999a). A comparison of the problem solving abilities and psychological distress of suicidal, bullied, and protected prisoners. *Criminal Justice and Behavior, 26*, 196–216. doi:10.1177/0093854899026002003

Biggam, F., & Power, K. (1999b). Social problem solving skills and psychological distress among incarcerated young offenders: The issue of bullying and victimization. *Cognitive Therapy and Research, 23*, 307–326. doi:10.1023/A:1018795632067

Biggam, F., & Power, K. (1999c). Suicidality and the state-trait debate on problem solving deficits: A re-examination with incarcerated young offenders. *Archives of Suicide Research, 5*, 27–42. doi:10.1080/13811119908258313

Black, D. W., Blum, N., Eichinger, L., McCormick, B., Allen, J., & Sieleni, B. (2008). STEPPS: Systems training for emotional predictability and problem solving in women offenders with borderline personality disorder in prison—A pilot study. *CNS Spectrums, 13*, 881–886. doi:10.1017/S1092852900016989

Black, D. W., Blum, N., McCormick, B., & Allen, J. (2013). Systems Training for Emotional Predictability and Problem Solving (STEPPS) group treatment for offenders with borderline

personality disorder. *Journal of Nervous and Mental Disease, 201*, 124–129. doi:10.1097/NMD.0b013e31827f6435

Blitz, C., Wolff, N., Pan, K., & Pogorzeiski, W. (2005). Gender specific behavioral health and community release patterns among New Jersey prison inmates: Implications for treatment and community reentry. *American Journal of Public Health, 95*, 1741–1746. doi:10.2105/AJPH.2004.059733

Bonta, J., Blais, J., & Wilson, H. (2013). *The prediction of risk for mentally disordered offenders: A quantitative synthesis*. Public Safety Canada, User Report 2013:01. Retrieved July 14, 2015 from http://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/prdctn-rsk-mntlly-dsrdrd/prdctn-rsk-mntlly-dsrdrd-eng.pdf

Bonta, J., Law, M., & Hanson, K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin, 123*, 123–142. doi:10.1037/0033-2909.123.2.123

Burgoyne, B., & Tyson, G. (2013). An evaluation of the 'Think First' program. *Australian & New Zealand Journal of Criminology, 46*, 88–100. doi:10.1177/0004865812469973

Burns, D. D. (1989). *The feeling good handbook*. New York: William Morrow.

Bush, J., Glick, B., & Taymans, J. (1997, Revised 1998). *Thinking for a change: Integrated cognitive behavior change program*. National Institute of Corrections. Washington D.C.: U.S. Department of Justice.

Castillo, E., & Fiftal Alarid, L. (2011). Factors associated with recidivism among offenders with mental illness. *International Journal of Offender Therapy and Comparative Criminology, 55*, 98–117. doi:10.1177/0306624X09359502

Covington, S. S. (1999). *Helping women recover: A program for treating addiction*. San Francisco: Jossey-Bass.

Cullen, F. T, & Gendreau, P. (1989). The effectiveness of correctional rehabilitation. In L. Goodstein & D.L. MacKenzie (Eds.), *The American prison: Issues in research policy* (pp. 23–44). New York: Plenum. doi:10.1300/J020v18n03_09

Cusack, K. J., Morrissey, J. P., Cuddeback, G. S., Prins, A., & Williams, D. M. (2010). Criminal justice involvement, behavioral health service use, and costs of forensic assertive community treatment: A randomized trial. *Community Mental Health Journal, 46*, 356–363. doi:10.1007/s10597-010-9299-z

D'Zurilla, T. J., & Goldfried, M. R. (1971). Problem solving and behavior modification. *Journal of Abnormal Psychology, 78*, 107–126. doi:10.1037/h0031360

D'Zurilla, T. J., & Nezu, A. M. (2001). Problem solving therapies. In K. S. Dobson (Ed.), *Handbook of cognitive-behavioral therapies* (2d ed., pp. 211–245). New York: Guilford.

D'Zurilla, T. J., Nezu, A. M., & Maydeu-Olivares, A. (2002). *Social Problem solving Inventory-Revised (SPSI–R)*. North Tonawanda, NY: Multi-Health Systems.

D'Zurilla, T. J., & Nezu, A. M. (2007). *Problem-solving therapy: A positive approach to clinical intervention* (3rd ed.). New York, NY: Springer.

Daicoff, S., & Wexler, D. B. (2003). Therapeutic jurisprudence. In I. B. Weiner (Editor-in-Chief) & A. M. Goldstein (Vol. Ed.), *Handbook of psychology: Vol. 11. Forensic psychology* (pp. 561–580). doi:10.1002/0471264385.wei1128

Davies, K., Lewis, J., Byatt, J., Purvis, E., & Cole, B. (2004). *An evaluation of the literacy demands of general offending behaviour programmes*. Findings 233. London, England: Home Office Research, Development and Statistics Directorate.

De Leon, G. (1989). Therapeutic communities for substance abuse: Overview of approach and effectiveness. *Psychology of Addictive Behaviors, 3*, 140–147. doi:10.1037/h0080571

Debidin, M., & Lovbakke, J. (2005). Offending behavior programmes in prison and probation. In G. Harper & C. Chitty (eds.), *The Impact of corrections on reoffending: A review of 'what works'* (pp. 31–54). Home Office Research Study 291. London, England: Home Office Research, Development and Statistics Directorate.

Ditton, P. (1999). *Mental health and treatment of inmates and probationers (Bureau of Justice Statistics Special Report NCJ 174463)*. Washington, D.C.: U.S. Department of Justice.

Dowden, C., & Andrews, D. (1999). What works for female offenders: A meta-analytic review. *Crime & Delinquency, 45*, 438–452. doi:10.1177/0011128799045004002

Dowden, C., & Brown, S. L. (2002). The role of substance abuse factors in predicting recidivism: A meta-analysis. *Psychology, Crime, & Law, 8*, 243–264. doi:10.1080/10683160208401818

Dozois, D. J., & Dobson, K. S. (2001). Information processing and cognitive organization in unipolar depression: Specificity and comorbidity issues. *Journal of Abnormal Psychology, 110*, 236–246.

Elliot, W., & Verdeyen, V. (2002). *Game over! Strategies for managing inmate deception*. Lanham, MD: American Correctional Association.

Ellis, A., & Harper, R. A. (1975). *A new guide to rational living*. Englewood Cliffs, NJ: Prentice-Hall.

Eno Louden, J., Manchak, S., O'Connor, M., & Skeem, J. (2015). Applying the sequential intercept model to reduce recidivism among probationers and parolees with mental illness. In P. Griffin, K. Heilbrun, E. Mulvey, D. DeMatteo, & C. Schubert (Eds.), *The sequential intercept model and criminal justice: Promoting community alternatives for individuals with serious mental illness* (pp. 118–136). New York: Oxford.

Farrington, D. P., & Welsh, B. C. (2005). Randomized experiments in criminology: What have we learned in the last two decades? *Journal of Experimental Criminology, 1*, 9–38. doi:10.1007/s11292-004-6460-0

Fazel, S., Bains, P., & Doll, H. (2006). Substance abuse and dependence in prisoners: A systematic review. *Addiction, 101*, 181–191. doi:10.1111/j.1360-0443.2006.01316.x

First, M., Williams, J., Karg, R., & Spitzer, R. (2015). *Structured clinical interview for DSM-5—Research version*. Arlington, VA: American Psychiatric Association.

Freeman, A., Pretzer, J., Fleming, B., & Simon, K. M. (1990). *Clinical applications of cognitive therapy*. New York: Plenum. doi:10.1007/978-1-4684-0007-6

Gendreau, P., & Andrews, D. A. (1990). Tertiary prevention: What the meta-analyses of the offender treatment literature tell us about what works. *Canadian Journal of Criminology, 32*, 173–184.

Gendreau, P., & Ross, R. R. (1987). Revivification of rehabilitation: Evidence from the 1980's. *Justice Quarterly, 4*, 349–407. doi:10.1080/07418828700089411

Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offenders recidivism: What works! *Criminology, 34*, 575–607. doi:10.1111/j.1745-9125.1966.tb01220.x

Goff, A., Rose, E., Rose, S., & Purves, D. (2007). Does PTSD occur in sentenced prison populations? A systematic literature review. *Criminal Behaviour and Mental Health, 17*, 152–162. doi:10.1002/cbm.653

Goldstein, A. P., & Glick, B. (1987). *Aggression replacement training: A comprehensive intervention for aggressive youth*. Champaign, IL: Research Press.

Goldstein, A. P., & Glick, B. (1994). *The prosocial gang: Implementing aggression replacement training*. Thousand Oaks, CA: Sage.

Greening, L. (1997). Adolescent stealers' and nonstealers' social problem solving skills. *Adolescence, 32*, 51–55.

Griffin, P., Heilbrun, K., Mulvey, E., DeMatteo, D., & Schubert, C. (Eds.) (2015). *Criminal justice and the sequential intercept model: Promoting community alternatives for people with severe mental illness.* New York: Oxford. doi:10.1093/med:psych/9780199826759.001.0001

Hanson, R., Bourgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior, 36*, 865–891. doi:10.1177/0093854809338545

Hayward, J., McMurran, M., & Sellen, J. (2008). Social problem solving in vulnerable adult prisoners: Profile and intervention. *The Journal of Forensic Psychiatry & Psychology, 19*, 243–248. doi:10.1080/14789940701752193

Higgins, J. P., & Thies, A. P. (1981). Social effectiveness and problem solving thinking of reformatory inmates. *Journal of Offender Services, Counseling and Rehabilitation, 5*, 93–98. doi:10.1080/10509674.1981.9963670

Hodas, G. R. (2006). *Responding to childhood trauma: The promise and practice of trauma informed care*. Retrieved from http://www.nasmhpd.org/sites/default/files/Responding%20to%20Childhood%20Trauma%20-%20Hodas.pdf

Hollin, C. R., McGuire, J., Hounsome, J. C., Hatcher, R. M., Bilby, C. A. L., & Palmer, E. J. (2008). Cognitive skills behavior programs for offenders in the community: A reconviction analysis. *Criminal Justice and Behavior, 35*, 269–283. doi:10.1177/0093854807312234

Hollin, C. R., McGuire, J., Palmer, E. J., Bilby, C., Hatcher, R., & Holmes, A. (2002). *Introducing pathfinder programmes into the Probation Service: An interim report (Home Office Research Study 247)*. London, England: Home Office.

Holloway, K. R., Bennett, T. H., & Farrington, D. P. (2006). The effectiveness of drug treatment programs in reducing criminal behavior: A meta-analysis. *Psicothema, 18*, 620–629.

Husband, S. D., & Platt, J. J. (1993). The cognitive skills component in substance abuse treatment in correctional settings: A brief review. *Journal of Drug Issues, 23*, 31–42.

Inciardi, J. A., Martin, S. S., & Butzin, C. A. (2004). Five-year outcomes of therapeutic community treatment of drug-involved offenders after release from prison. *Crime & Delinquency, 50*, 88–107. doi:10.1177/0011128703258874

Jaffe, W. B., & D'Zurilla, T. J. (2003). Adolescent problem solving, parent problem solving, and externalizing behavior in adolescents. *Behavior Therapy, 34*, 295–311. doi:10.1016/S0005-7894(03)80002-3

James, D., & Glaze, L. (2006). *Mental health problems of prison and jail inmates (Bureau of Justice Statistics Special Report NCJ 213600)*. Washington, D.C.: U.S. Department of Justice.

Jennings, W. S., Kilkenny, R., & Kohlberg, L. (1983). Moral-development theory and practice for youthful and adult offenders. In W. S. Laufer & J. M. Day (Eds.), *Personality theory, moral development & criminal behavior* (pp. 281–355). New York: Lexington.

Jensen, E. L., & Kane, S. L. (2010). The effect of therapeutic community on time to first re-arrest: A survival analysis. *Journal of Offender Rehabilitation, 49*, 200–209. doi:10.1080/10509671003666594

Junginger, J., Claypoole, K., Laygo, R., & Crisanti, A. (2006). Effects of serious mental illness and substance use on criminal offenses. *Psychiatric Services, 57*, 879–882. doi:10.1176/appi.ps.57.6.879

Kenne, D. R. (2011). Examination of thinking error and the responsivity principle in a cognitive-behavioral intervention for offenders: Implications for criminal justice policy. *Dissertation Abstracts International Section A, 72*, 744.

Lamberti, J. S., Weisman, R., & Faden, D. I. (2004). Forensic assertive community treatment: Preventing incarceration of adults with severe mental illness. *Psychiatric Services, 55*, 1285–1293. doi:10.1176/appi.ps.55.11.1285

Landenberger, N. A., & Lipsey, M. W. (2005). The positive effects of cognitive–behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology, 1*, 451–476. doi:10.1007/s11292-005-3541-7

Leadbeater, B. J., Hellner, I., Allen, J. P., & Aber, J. L. (1989). Assessment of interpersonal negotiation strategies in youth engaged in problem behaviors. *Developmental Psychology, 25*, 465–472. doi:10.1037/0012-1649.25.3.465

Leahy, D., & Neary, J. P. (1997). Public policy towards R&D in oligopolistic industries. *American Economic Review, 87,* 642–662.

Lindsay, W. R., Hamilton, C., Moulton, S., Scott, S., Doyle, M., & McMurran, M. (2011). Assessment and treatment of social problem solving in offenders with intellectual disability. *Psychology, Crime & Law, 17*, 181–197. doi:10.1080/10683160903392756

Lipsey, M. W., & Cullen, F. T. (2007). The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annual Review of Law and Social Science, 3*, 297–320. doi:10.1146/annurev.lawsocsci.3.081806.112833

Lipsey, M. W., Chapman, G., & Landenberger, N. A. (2001). Cognitive-behavioral programs for offenders. *The Annals of the American Academy of Political and Social Science, 578*, 144–157. doi:10.1177/0002716201578001009

Lipsey, M. W., Landenberger, N. A., & Wilson, S. J. (2007). Effects of cognitive-behavioral programs for criminal offenders. *Campbell Systematic Reviews, 6*, 1–27.

Little, G., & Robinson, K. (1986). *How to escape your prison: A moral reconation therapy workbook*. Memphis: Eagle Wing Books.

Lochman, K. E., & Lenhart, L. A. (1993). Anger coping intervention for aggressive children: Conceptual models and outcome effects. *Clinical Psychology Review, 10*, 146–148. doi:10.1016/S0272-7358(05)80006-6

Losel, F. (1995). The efficacy of correctional treatment: A review and synthesis of meta-evaluations. In J. McGuire (Ed.), *What works: Reducing reoffending: Guidelines from research and practice* (pp. 79–111). New York: John Wiley.

Lynch, S. M., DeHart, D. D., Belknap, J. E., Green, B. L., Dass-Brailsford, P., Johnson, K. A., et al. (2014). A multisite study of the prevalence of serious mental illness, PTSD, and substance use disorders of women in jail. *Psychiatric Services, 65*, 670–674. doi:10.1176/appi.ps.201300172

Lynch, S. M., Heath, N. M., Mathews, K. C., & Cepeda, G. J. (2012). Seeking Safety: An intervention for trauma-exposed incarcerated women. *Journal of Trauma & Dissociation, 13*, 88–101. doi:10.1080/15299732.2011.608780

MacKenzie, D. L., & Hickman, L. J. (1998). *What works in corrections? An examination of the effectiveness of the type of rehabilitation programs offered by Washington State Department of Corrections (Report to the state of Washington legislature joint audit and review committee)*. College Park, MD: University of Maryland, Department of Criminology and Criminal Justice.

Malinowski, A. (2003). "What works" with substance users in prison? *Journal of Substance Use, 8*, 223–233. doi:10.1080/14659890310001636107

Marlatt, G. A., Parks, G. A., & Witkiewitz, K. (2002). *Clinical guidelines for implementing relapse prevention therapy: A guideline developed for the behavioral health recovery management project*. Seattle, WA: University of Washington Addictive Behaviors Research Center.

McGuire, J. (2005a). Social problem solving: Basic concepts, research, and applications. In M. McMurran & J. McGuire (Eds.), *Social problem solving and offenders: Evidence, evaluation and evolution* (pp. 3–30). Chichester, England: Wiley. doi:10.1002/9780470713488.ch1

McGuire, J. (2005b). The *Think First* Programme. In M. McMurran & J. McGuire (Eds.), *Social problem solving and offenders: Evidence, evaluation and evolution* (pp. 183–206). Chichester, England: Wiley. doi:10.1002/9780470713488.ch10

McGuire, J., Bilby, C. A. L., Hatcher, R. M., Hollin, C. R., Hounsome, J., & Palmer, E. J. (2008). Evaluation of structured cognitive—behavioural treatment programmes in reducing criminal recidivism. *Journal of Experimental Criminology, 4*, 21–40. doi:10.1007/s11292-007-9047-8

McGuire, M., & Hatcher, R. (2001). Offence-focused problem solving: Preliminary evaluation of a cognitive skills program. *Criminal Justice and Behavior, 28*, 564–587. doi:10.1177/009385480102800502

McMullin, R. E. (2000). *The new handbook of cognitive therapy techniques*. New York: W.W. Norton.

McMurran, M. (2005). Social problem solving and offending: Reflections and directions. In M. McMurran & J. McGuire (Eds.), *Social problem solving and offending: Evidence, evaluation and evolution* (pp. 297–307). Chichester, England: Wiley. doi:10.1002/9780470713488

McMurran, M. (2009). Motivation interviewing with offenders: A systematic review. *Legal and Criminological Psychology, 14*, 83–100. doi:10.1348/135532508X278326

McMurran, M., & Christopher, G. (2009). Social problem solving, anxiety, and depression in adult male prisoners. *Legal and Criminological Psychology, 14*, 101–110. doi:10.1348/135532507X267031

McMurran, M., & McGuire, J. (Eds.). (2005). *Social problem solving and offending: Evidence, evaluation and evolution*. Chichester, England: Wiley. doi:10.1002/9780470713488

McMurran, M., Egan, V., & Duggan, C. (2005). Stop & Think! Social problem-solving therapy with personality-disordered offenders. In M. McMurran & J. McGuire (Eds)., *Social problem*

solving and offending: Evidence, evaluation and evolution (pp. 207–220). Chichester, England: Wiley. doi:10.1002/9780470713488.ch6

McMurran, M., Egan, V., Blair, M., & Richardson, C. (2001a). The relationship between social problem solving and personality in mentally disordered offenders. *Personality and Individual Differences, 30*, 517–524. doi:10.1016/S0191-8869(00)00050-7

McMurran, M., Fyffe, S., McCarthy, L., Duggan, C., & Latham, A. (2001b). Stop & Think! Social problem solving therapy with personality-disordered offenders. *Criminal Behaviour and Mental Health, 11*, 273–285. doi:10.1002/cbm.401

McMurran, M., Huband, N., & Duggan, C. (2008). A comparison of treatment completers and non-completers of an in-patient treatment programme for male personality-disordered offenders. *Psychology and Psychotherapy: Theory, Research and Practice, 81*, 193–198. doi:10.1348/147608308X288762

Meichenbaum, D. H. (1975). *Stress inoculation training*. New York: Pergamon Press.

Messina, N., & Grella, C. (2006). Childhood trauma and women's health outcomes in a California prison population. *American Journal of Public Health, 96*, 1842–1848. doi:10.2105/AJPH.2005.082016

Messina, N., Grella, C. E., Cartier, J., & Torres, S. (2010). A randomized experimental study of gender-responsive substance abuse treatment for women in prison. *Journal of Substance Abuse Treatment, 38*, 97–107. doi:10.1016/j.jsat.2009.09.004

Milkman, H., & Wanberg, K. (2007). *A review and discussion for corrections professionals*. Washington, DC: National Institute of Corrections.

Mitchell, O., Wilson, D. B., & MacKenzie, D. L. (2012). The effectiveness of incarceration-based drug treatment on criminal behavior: A systematic review. *Campbell Systematic Reviews, 18*, 1–75. doi:10.4073/csr.2012.18

Moloney, K. P., van den Bergh, B. J., & Moller, L. F. (2009). Women in prison: The central issues of gender characteristics and trauma. *Public Health, 123*, 426–430. doi:10.1016/j.puhe.2009.04.002

Monti, P. M., Rohsenow, D. J., Colby, S. M., &, Abrams, D. B. (1995). Coping and social skills training. In R. K. Hester & W. R. Miller (Eds.), *Handbook of alcoholism treatment approaches: Effective alternatives* (pp. 221–241). Boston: Allyn & Bacon.

Morrissey, J., Meyer, P., & Cuddeback, G. (2007). Extending assertive community treatment to criminal justice settings: Origins, current evidence, and future directions. *Community Mental Health Journal, 43*, 527–544. doi:10.1007/s10597-007-9092-9

Munetz, M., & Griffin, P. (2006). Use of the Sequential Intercept Model as an approach to decriminalization of people with severe mental illness. *Psychiatric Services, 57*, 544–549. doi:10.1176/appi.ps.57.4.544

Najavits, L. M. (2009). Psychotherapies for trauma and substance abuse in women. *Trauma, Violence, & Abuse, 10*, 290–298. doi:10.1177/1524838009334455

Najavits, L. M. (n.d.). *Implementing seeking safety therapy for PTSD and substance abuse: Clinical guidelines*. Retrieved from http://www.bhrm.org/guidelines/PTSD.pdf

National Center for Trauma-Informed Care. (2008). *Models for developing trauma-informed behavioral health systems and trauma-specific services*. Retrieved from http://www.ct.gov/dmhas/lib/dmhas/trauma/TraumaModels.pdf

National Commission on Correctional Health Care. (2002). *The health status of soon-to-be-released inmates: A report to Congress* (Vol. I). Washington, DC: U.S. Department of Justice.

Nezu, A. M., Nezu, C. M., & McMurran, M. (2009). Problem solving therapy. In W. T. O'Donohue & J. E. Fisher (Eds.), *General principles and empirically supported techniques of cognitive behavior therapy* (pp. 500–505). Hoboken, NJ: John Wiley & Sons.

Nezu, C. M., D'Zurilla, T. J., & Nezu, A. M. (2005a). Problem solving therapy: Theory, practice, and application to sex offenders. In M. McMurran & J. McGuire (Eds.), *Social problem solving and offenders: Evidence, evaluation and evolution* (pp. 103–123). Chichester, England: Wiley. doi:10.1002/9780470713488.ch6

Nezu, C. M., Greenberg, J., & Nezu, A. M. (2005b). Project STOP: Cognitive behavioral assessment and treatment for sex offenders with intellectual disabilities. *International Journal of Behavioral Consultation and Therapy, 1*, 191–203. doi:10.1037/h0100744

Nezu, C. M., Nezu, A. M., Dudek, J. A., Peacock, M., & Stroll, J. (2005c). Social problem solving correlates of sexual deviancy and aggression among adult child molesters. *Journal of Sexual Aggression, 11*, 27–36. doi:10.1080/13552600512331329540

Nezu, C. M., Fiore, A. A., & Nezu, A. M. (2006). Problem solving treatment for intellectually disabled sex offenders. *International Journal of Behavioral Consultation and Therapy, 2*, 266–276. doi:10.1037/h0100781

NIC (National Institute of Corrections) (1996). *Cognitive interventions program: Think.* National Institute of Corrections Information Center. Washington D.C.: U.S. Department of Justice.

Olson, D. A., & Lurigio, A. J. (2014). The long-term effects of prison-based drug treatment and aftercare services on recidivism. *Journal of Offender Rehabilitation, 53*, 600–619. doi:10.1080/10509674.2014.956965

Ong, G., Al-Attar, Z., Roberts, C., & Harsent, L. (2003). *Think First: An accredited community-based cognitive-behavioural programme in England and Wales. Findings from the prospective evaluation in three Probation Areas*. Oxford, England: Probation Studies Unit, Centre for Criminological Research, University of Oxford.

Palmer, E. J., Hatcher, R. M., McGuire, J., & Hollin, R. C. (2015). Cognitive skills programs for female offenders in the community: Effect on reconviction. *Criminal Justice and Behavior, 42*, 345–360. doi:10.1177/0093854814552099

Palmer, E. J., McGuire, J., Hatcher, R. M., Hounsome, J. C., Bilby, C. A. L., & Hollin, C. R. (2008). The importance of appropriate allocation to offending behavior programs. *International Journal of Offender Therapy and Comparative Criminology, 52*, 206–221. doi:10.1177/0306624X07303877

Palmer, E. J., McGuire, J., Hatcher, R. M., Hounsome, J. C., Bilby, C. A. L., & Hollin, C. R. (2009). Allocation to offending behavior programs in the English and Welsh Probation Service. *Criminal Justice and Behavior, 36*, 909–922. doi:10.1177/0093854809338889

Pearson, F. S., & Lipton, D. S. (1999). A meta-analytic review of the effectiveness of corrections-based treatments for drug abuse. *The Prison Journal, 79*, 384–410. doi:10.1177/0032885599079004003

Pearson, F. S., Lipton, D. S., Cleland, C. M., & Yee, D. S. (2002). The effects of behavioural/cognitive–behavioural programmes on recidivism. *Crime and Delinquency, 48*, 476–496. doi:10.1177/0011128702048003006

Perry, A. E., Darwin, Z., Godfrey, C., McDougall, C., Lunn, J., Glanville, J., et al. (2009). *Substance Use and Misuse, 44*, 374–400. doi:10.1080/10826080802347560

Peterson, J., Skeem, J., Hart, E., Vidal, S., & Keith, F. (2010). Typology of offenders with mental disorder: Exploring the criminalization hypothesis. *Psychiatric Services, 61*, 1217–1222. doi:10.1176/appi.ps.61.12.1217

Peterson, J., Skeem, J., Kennnealy, P., Bray, B., & Zvonkovic, A. (2014). How often and how consistently do symptoms directly precede criminal behavior among offenders with mental illness? *Law and Human Behavior, 38*, 439–449. doi:10.1037/lhb0000075

Platt, J. J., Scura, W. C., & Hannon, J. R. (1973). Problem solving thinking of youthful incarcerated heroin addicts. *Journal of Community Psychology, 1*, 278–281. doi:10.1002/1520-6629(197307)1:3<278:AID-JCOP2290010307>3.0.CO;2-I

Porporino, F. J., Fabiano, E. A., & Robinson, D. (1991). *Focusing on successful reintegration: Cognitive skills training for offenders* (Report No. R-19). Retrieved from Government of Canada Publications website: http://publications.gc.ca/site/eng/9.568196/publication.html

Prendergast, M. L., Hall, E. A., & Wexler, H. K. (2003). Multiple measures of outcome in assessing a prison-based drug treatment program. *Journal of Offender Rehabilitation, 37*, 65–94. doi:10.1300/J076v37n03_04

Prins, S., Skeem, J., Mauro, C., & Link, B. (2015). Criminogenic factors, psychotic symptoms, and incident arrests among people with serious mental illnesses under intensive outpatient treatment. *Law and Human Behavior, 39*, 177–188. doi:10.1037/lhb0000104

Reinecke, M. A., & Freeman, A. (2003). Cognitive therapy. In A. S. Gurman & S. B. Messer (Eds.), *Essential psychotherapies: Theory and practice* (2nd ed., pp. 224–271). New York: Guilford.

Roberts, C. (2004). An early evaluation of a cognitive offending behavior programme ('Think First') in probation areas. *Vista: Perspective on probation, 8*, 130–136.

Robinson, D. (1995). *The impact of cognitive skills training on post-release recidivism among Canadian federal offenders* (Report No. R-41). Retrieved from Government of Canada Publications website http://publications.gc.ca/site/eng/379318/publication.html

Ross, R. R., & Fabiano, E. A. (1985). *Time to think: A cognitive model of delinquency prevention and offender rehabilitation.* University of Tennessee: Institute of Social Sciences and Art.

Ross, R., & Ross, R. (1995). *Thinking straight: The reasoning and rehabilitation program for delinquency prevention and offender rehabilitation*. Ottawa, Ontario, Canada: Air Training.

Ross, R. R., Fabiano, E. A., & Ewles, C. D. (1988). Reasoning and rehabilitation. *International Journal of Offender Therapy and Comparative Criminology, 32*, 29–35. doi:10.1177/0306624X8803200104

Ross, R. R., Fabiano, E. A., & Ross, R. D. (1986). *Reasoning and rehabilitation: A handbook for teaching cognitive skills.* Ottawa, Ontario: T3 Associates.

Rotter, J. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*, 1–28. doi:10.1037/h0092976

Sacks, S., McKendrick, K., Sacks, J. Y., & Cleland, C. M. (2010). Modified therapeutic community for co-occurring disorders: Single investigator meta-analysis. *Substance Abuse, 31*, 146–161. doi:10.1080/08897077.2010.495662

Sacks, S., Sacks, J. Y., & De Leon, G. (1999). Treatment for MICAs: Design and implementation of the modified TC. *Journal of Psychoactive Drugs, 31*, 19–30. doi:10.1080/02791072.1999.10471722

Sacks, S., Sacks, J. Y., McKendrick, K., Banks, S., & Stommel, J. (2004). Modified TC for MICA offenders: Crime outcomes. *Behavioral Sciences and the Law, 22*, 477–501. doi:10.1002/bsl.599

Samenow, S. E. (1984). *Inside the criminal mind*. New York: Times Books.

Sarteschi, C. (2013). Mentally ill offenders involved with the U.S. criminal justice system: A synthesis. *Sage Open*, July–September, 1–11. doi:10.1177/2158244013497029

Seligman, M. E. P., Walker, E. F., & Rosenhan, D. L. (2001). *Abnormal psychology* (4th ed.). New York: W.W. Norton.

Siegal, H., Li, L., & Rapp, R. (2002). Case management as a therapeutic enhancement: Impact on posttreatment criminality. *Journal of Addictive Diseases, 21*, 37–47. doi:10.1300/J069v21n04_04

Skeem, J. L., Manchak, S., & Peterson, J. K. (2011). Correctional policy for offenders with mental illness: Creating a new paradigm for recidivism reduction. *Law and Human Behavior, 35*, 110–126. doi:10.1007/s10979-010-9223-7

Skeem, J., Kennealy, P., Monahan, J., Peterson, J., & Appelbaum, P. (2015). Psychosis uncommonly and inconsistently precedes violence among high-risk individuals. *Clinical Psychological Science*. Published online before print April 24, 2015. doi:10.1177/2167702615575879

Skeem, J., Steadman, H., & Manchak, S. (2015). Applicability of the Risk-Need-Responsivity model to persons with mental illness involved in the criminal justice system. *Psychiatric Services in Advance*. Published online May 1, 2015. http://dx.doi.org/10.1176/appi.ps.201400448

Solomon, P., Draine, J., & Marcus, S. (2002). Predicting incarceration of clients of the psychiatric probation and parole service. *Psychiatric Services, 53*, 50–56. doi:10.1176/appi.ps.53.1.50

Spiropoulos, G. V., Spruance, L., Voorhis, P. V., & Schmitt, M. M. (2005). Pathfinders and Problem Solving: Comparative effects of two cognitive-behavioral programs among men and women offenders in community and prison. *Journal of Offender Rehabilitation, 42*, 69–94. doi:10.1300/J076v42n02_05

Spivack, G., Platt, J. J., & Shure, M. B. (1976). *The problem solving approach to adjustment*. San Francisco, CA: Jossey-Bass.

Steadman, H., Osher, F., Clark Robbins, P., Case, B., & Samuels, S. (2009). Prevalence of serious mental illness among jail inmates. *Psychiatric Services, 60*, 761–765. doi:10.1176/appi.ps.60.6.761

Steele, R. (2002a). *Psychometric features of think first participants pre and post programme*. Merseyside, England: Research and Information Section, National Probation Service.

Steele, R. (2002b). *Reconviction of offenders on think first*. Merseyside, England: Research and Information Section, National Probation Service.

Stewart-Ong, G., Harsent, L., Roberts, C., Burnett, R., & Al-Attar, Z. (2004). *Think First prospective research study: Effectiveness and reducing attrition*. London, England: National Probation Service.

Stohr, M. K., Hemmens, C., Buane, D., Dayley, J., Gornik, M., & Kjaer, K., et al. (2001). *Residential substance abuse treatment for state prisoners (RSAT) partnership process evaluation, final report* (NCJ publication No. 187352). Retrieved from https://www.ncjrs.gov/pdffiles1/nij/grants/187352.pdf

Telander, R. L. (2005). Offender treatment: An evaluation of a cognitive restructuring program. *Dissertation Abstracts International, 66*, 1187.

Tong, L. S. J., & Farrington, D. P. (2006). How effective is the "Reasoning and Rehabilitation" programme in reducing reoffending? A meta-analysis of evaluations in four countries. *Psychology, Crime & Law, 12*, 3–24. doi:10.1080/10683160512331316253

U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Substance Abuse Treatment. (2005). *Substance abuse treatment for adults in the criminal justice system treatment improvement protocol (TIP) series 44* (DHHS publication No. 05-4056). Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK64137/pdf/Bookshelf_NBK64137.pdf

US. Department of Justice, Office of Justice Programs, Bureau of Justice Assistance. (2005). *Residential substance abuse treatment for state prisoners (RSAT) program* (NCJ publication No. 206269). Retrieved from https://www.ncjrs.gov/pdffiles1/bja/206269.pdf

Van Dorn, R., Desmarais, S., Petrila, J., Haynes, S., & Singh, J. (2013). Effects of outpatient treatment on risk of arrest of adults with serious mental illness and associated costs. *Psychiatric Services, 64*, 856–862. doi:10.1176/appi.ps.201200406

Wakeling, H. C. (2007). The psychometric validation of the Social Problem solving Inventory—Revised with UK incarcerated sexual offenders. *Sexual Abuse, 19*, 217–236. doi:10.1007/s11194-007-9038-3

Wallace, B. C., Conner, L. C., & Dass-Brailsfor, P. (2011). Integrated trauma treatment in correctional health care and community-based treatment upon reentry. *Journal of Correctional Health Care, 17*, 329–343. doi:10.1177/1078345811413091

Walters, G. D. (2002). *Criminal belief systems: An Integrated-Interactive Theory of Lifestyles*. Westport, CT: Praeger.

Walters, G. D., & Elliott, W. N. (1999). Predicting release and disciplinary outcome with the psychological inventory of criminal thinking styles: Female data. *Legal and Criminological Psychology, 4*, 15–21.

Walters, G. D., & White, T. W. (1989). Thinking criminal: A cognitive model of lifestyle criminality. *Women Lawyers Journal, 76*, 4.

Wanberg, K. W., & Milkman, H. B. (1998). *Criminal Conduct and Substance Abuse Treatment: Strategies for Self-Improvement and Change; The provider's guide*. Thousand Oaks, CA: Sage Publications.

Wanberg, K. W., & Milkman, H. B. (2006). *Criminal conduct and substance abuse treatment: Strategies for self-improvement and change: The participant's workbook*. Thousand Oaks, CA: Sage Publications.

Ward, T., Yates, P. M., & Willis, G. M. (2012). The Good Lives Model and the Risk Need Responsivity Model: A critical response to Andrews, Bonta, and Wormith (2011). *Criminal Justice and Behavior, 39*, 94–110. doi:10.1177/0093854811426085

Weiner, R. L., & Brank, E. M. (Eds.). (2013). *Problem solving courts: Social science and legal perspectives*. New York: Springer. doi:10.1007/978-1-4614-7403-6

Welsh, W. N., & Zajac, G. (2013). A multisite evaluation of prison-based drug treatment: Four-year follow-up results. *The Prison Journal, 93*, 251–271. doi:10.1177/0032885513490265

Westmarland, N., Reid, P., Coulson, S., Hughes, J., & Hester, M. (2002). *An investigation into the factors associated with attrition in the Northumbria Probation Think First Programme*. Sunderland, England: International Centre for the Study of Violence and Abuse, University of Sunderland.

Wexler, H. K. (1995). The success of therapeutic communities for substance users in American prisons. *Journal of Psychoactive Drugs, 27*, 57–66. doi:10.1080/02791072.1995.10471673

Wilson, D. B., Bouffard, L. A., & MacKenzie, D. L. (2005). A quantitative review of structured, group-oriented, cognitive-behavioral programs for offenders. *Criminal Justice and Behavior, 32*, 172–204.

Yochelson, S., & Samenow, S. E. (1976). *The criminal personality: A profile for change* (Vol. l). New York: Jason Aronson.

Young, S., & Thome, J. (2011). ADHD and offenders. *World Journal of Biological Psychiatry, 12*, 124–128. doi:10.3109/15622975.2011.600319

Zamble, E., & Porporino, F. J. (1988). *Coping behavior and adaption in prison inmates*. New York: Springer. doi:10.1007/978-1-4613-8757-2

Zamble, E., & Quinsey, V. L. (1997). *The criminal recidivism process*. Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511527579

Zlotnick, C. (1997). Postraumatic stress disorder (PTSD), PTSD comorbidity, and childhood abuse among incarcerated women. *Journal of Nervous and Mental Disease, 185*, 761–763. doi:10.1097/00005053-199712000-00007

Zlotnick, C., Johnson, J., & Najavits, L. M. (2009). Randomized controlled pilot study of cognitive-behavioral therapy in a sample of incarcerated women with substance use disorder and PTSD. *Behavior Therapy, 40*, 325–336. doi:10.1016/j.beth.2008.09.004

Zlotnick, C., Najavits, L. M., Rohsenow, D. J., & Johnson, D. M. (2003). A cognitive-behavioral treatment for incarcerated women with substance abuse disorder and posttraumatic stress disorder: Findings from a pilot study. *Journal of Substance Abuse Treatment, 25*, 99–105. doi:http://dx.doi.org/10.1016/S0740-5472(03)00106-5

Deterring Gangs: Criminal Justice Approaches and Psychological Perspectives

Jane L. Wood, Emma Alleyne and Hayley Beresford

The issue of exactly what constitutes a gang is fraught with debate and lack of consensus across researchers and policy makers (see Wood and Alleyne 2010 for a fuller discussion). However, the Eurogang network has captured the elements broadly agreed by many researchers and policy makers by stating that, "a street gang (or troublesome youth group corresponding to a street gang elsewhere) is any durable, street-oriented youth group whose identity includes involvement in illegal activity" (Weerman et al. 2009, p. 20). The important elements of this definition are that gangs generally consist of young people who are collectively involved in illegal activities on the streets which create problems for communities and authorities. To address the problems that gangs create, a whole host of strategies have been developed to focus on, for example, prevention, intervention and suppression of gangs and their activities. The problem remains, however, that despite the concerted efforts of justice systems, gangs appear to be flourishing. This leaves criminal justice systems struggling to devise effective programs that will contain the disturbing menace that gangs pose in communities, schools, and prisons and the threat that they pose to individual gang members—on both a physical and psychological level. The current discussion considers programs that aim to deter existing gang activity and the political policies that underpin them: Gang prevention programs, although perhaps including elements of deterrence, will not be included since deterring current gang membership is not their key feature. The aim of this chapter is to provide an overall evaluation of deterrence strategies and to examine more closely the psychological effects that group processes may have in insulating gang members against the intended effects of legal deterrence efforts. Although gang membership is a worldwide phenomenon, the focus in this chapter will be on programs and policies that have been devised and implemented in the U.S. and the UK.

J.L. Wood (✉) · E. Alleyne · H. Beresford
School of Psychology, Centre of Research and Education in Forensic Psychology (CORE-FP), University of Kent, Canterbury, Kent CT2 7NP, UK
e-mail: J.L.Wood@kent.ac.uk

B.H. Bornstein and M.K. Miller (eds.), *Advances in Psychology and Law*, Advances in Psychology and Law 2, DOI 10.1007/978-3-319-43083-6_10

Deterrence

Deterrence has two forms: general and specific. *General* deterrence refers to the idea that nonoffending populations will be deterred from committing offenses because of the punishments that may follow committing crime. *Specific* deterrence refers to punishments designed to dissuade individual offenders from committing further offenses (Stafford and Warr 1993). However, for deterrence methods to be effective, they rely on the offenders being rational actors who make rational choices about their involvement in crime (i.e., rational choice theory; see Clarke and Cornish 1985). They make these choices by weighing up the pros [which may be material (e.g., financial), mental (e.g., thrill) and/or reputational (e.g., peer approval, status enhancement; see Matsueda 1992)] and the cons [which may be legal (e.g., imprisonment), financial (loss of earnings—fines), emotional (e.g., feeling shame) and/or reputational (e.g., loss of respect from significant others)]. If more benefits than costs from crime are anticipated then offenses will be repeated. For instance, research shows that if young people anticipate that they will experience thrills from committing acts of theft they are likely to commit more thefts, and if they perceive that a valued status such as being seen as 'cool' can be gained via theft and/or violence then they maintain an intention to commit further acts of theft and/or violence (Matsueda et al. 2006). On the other hand, if the *costs* of crime are perceived as outweighing any anticipated benefits then both youth and adults are likely to make a rational choice not to commit further crimes (Zimring and Hawkins 1973). In short, it is the anticipated costs of crime that should act as a deterrent to future offending.

In order to quantify these rational choices, the Expected Utility Model of crime explains that a calculation of expected costs of crime may be defined as the probability of getting caught multiplied by the anticipated punishment, whilst the expected benefits are defined as the probability of getting away with the offense multiplied by its anticipated gains (Becker 1968). Injected into this equation is the probability parameter, which is a subjective evaluation of an individual's perceived chances of being caught and punished for the offense. In other words, an essential element of an expected utility calculation is what the individual *perceives* to be the potential costs and rewards of criminal activity. For this calculation to effectively reduce crime by deterring offending it is necessary that people's rational choice calculi favor the *perceived costs* rather than the *perceived benefits* side of the equation, and that they anticipate a real likelihood of being caught and punished for an offense. In short, the expected risks associated with crime must be perceived to outweigh the expected returns (Becker 1968).

The rational choice calculus used for decision making may include several estimations of offense outcomes. Research findings support a rational choice perspective of offending by showing that, for youth, if they believe that they will be arrested for committing acts of violence or theft then they will commit fewer such offenses—provided that they view arrest as a bad outcome (Matsueda et al. 2006). However, this same piece of research also showed how it is the perceived benefits

of crime that are the ultimate drivers of rational choices. As such, potential offenders may pay little attention to the anticipated costs of offending. Nonetheless, individuals' risk perceptions are not static; they are fluid and malleable and updated when new information comes to light. For example, research shows that juvenile offenders judge their risks of arrest according to their previous experiences. That is, if they have previously been arrested they are more likely to increase their risk perceptions to levels higher than they had before arrest (Anwar and Loughran 2011). However, this updating process may have limitations. As Anwar and Loughran note, updating risk perceptions seems to relate primarily to the crimes for which they had been arrested (e.g., acts of violence) rather than to criminal activity generally or crimes for which they had not been arrested (e.g., income-generating offenses).

Exactly why this is so, is not clear. It could be that offenders believe that because they are already known to the police for committing a specific type of offense, they are likely to be obvious suspects following similar offenses and as a result consider it wise to avoid committing such offenses. Equally, it could be that offenders know the penalties for committing offenses for which they have already been convicted but have little knowledge of the punishments associated with other offenses. What does seem clear is that policies that aim to deter offenders from *any* form of offending seem far less likely to hit the mark than policies that target specific offense types (Anwar and Loughran 2011). Also, as the number of crimes committed by an offender increases, a corresponding drop occurs in the effect of arrest on risk perceptions even though it remains positive. As a result, compared to less experienced offenders, more experienced offenders are likely to have weaker perceptions of risk across time and experience.

The idea that risk perceptions emerge from a systematic cognitive model bounded by rationality is intuitively appealing but it is not without its problems. For instance, it is well documented in psychology that people are cognitive misers (Fiske and Taylor 1991) who frequently resort to heuristics (Tversky and Kahneman 1974) when making decisions. People are also prey to the gambler's fallacy (Matsueda et al. 2006) in which they assume that if something has happened many times recently then it is less likely to occur as frequently in the future. So, offenders using a gambler's fallacy as a decision base for committing crime may erroneously judge that if, for instance, they have escaped arrest then they are due to be caught, or if they have been arrested several times then they are due to get away with an offense (Pogarsky and Piquero 2003).

The source of information on which risk perceptions are founded is also an important feature in decisions to offend. Some may gain an idea of arrest certainty from popular media which depicts the restoration of justice via the consistent arrest, prosecution, and punishment of offenders (Parker and Grasmick 1979). Such deterrent effects stemming from a diet of media consumption are, however, more likely to be effective with naïve individuals with no familiarity of criminal justice systems than it is with more seasoned offenders who factor in personal experience (Matsueda et al. 2006).

Deterrence effects may also be influenced by a trade-off between immediate rewards and delayed punishments (Apel 2013). For example, the rewards of crime are relatively immediate whereas criminal sanctions are likely to follow lengthy criminal proceedings and fall far in the future. Deterrence may also be negatively influenced by 'dimensional preferences' which strongly push individuals towards the potential benefits of crime than towards its potential costs (Carroll 1978). Even without the above potential influences, a rational choice perspective assumes that people base their decisions on an *accurate knowledge* of facts such as knowing which punishments follow which crimes. In reality people are not good at judging the likelihood of being caught and equally poor at knowing the penalties associated with offending (Kleck et al. 2005).

Anti-gang Programs

Findings overall indicate that the strongest deterrent effects develop more from the *certainty of apprehension* rather than from the *severity of punishment* (e.g., Klepper and Nagin 1989), even when people are aware of which penalties follow which crimes. Indeed, research shows that police presence can be effective in preventing crime (Braga 2005). Some claim that high police presence and visibility increases offenders' perceptions of risk of apprehension and has such a substantial deterrent effect that it justifies shifting resources from imprisonment to policing (Durlauf and Nagin 2011). Consequently, it could/should be expected that high police presence together with an increased likelihood of punishment would have particularly high deterrent effects on those involved in high levels of offending, such as gang members. The remainder of this section first outlines and then evaluates two of the main approaches used to deter gang membership and activity and continues by considering the development of anti-gang policies that enhance the criminality of gang membership.

Suppression Strategies

High police presence and swift criminal justice responses provide the bases for suppression strategies aimed directly at deterring gangs from criminal activity. Suppression projects began in the USA during the 1970s and developed over the next few decades to include a wide range of programs such as: Operation Hardcore, Operation Hammer, and Operation Safe Streets (Los Angeles); the Gang Congregation Ordinance (Chicago); the Anti-gang Initiative (Dallas); and the multi-site Youth Firearms Violence Initiative (Los Angeles, Milwaukee, Seattle, and San Antonio). The goal of suppression tactics is to respond swiftly and forcibly to gang activities in order to deter gang members from continued criminal activity. To achieve prosecutorial success, some programs (e.g., Operation Hardcore)

included components such as special gang prosecution units and vertical prosecution. Vertical prosecution consists of the allocation of a special gang prosecutor to a case from its outset to conviction and results in an increase in prosecution and conviction rates (OJJDP 1994). Other tactics included suppressing gang members' ability to associate with one another by threatening them with arrest if they were seen together in public places. For instance, the Chicago Gang Congregation Ordinance stated that,

> Whenever a police officer observes a person whom he reasonably believes to be a criminal street gang member loitering in any public place with one or more other persons, he shall order all such persons to disperse and remove themselves from the area. Any person who does not promptly obey such an order is in violation of this section… (Chicago Municipal Code § 8-4-015 June 17, 1992).

Violation of this section was punishable by a fine of up to $500, imprisonment of up to six months and/or up to 120 h of community service. Other programs specifically targeted the most dangerous gangs and most dangerous gang members. For example, the Tri-Agency Resource Gang Enforcement Team (TARGET) in California would identify dangerous gang members, monitor their activities, arrest and prosecute them using vertical prosecution and then closely supervise them whilst on parole to deter them from committing further offenses. Arrest swiftly followed if they did so. In Dallas the Anti-Gang Initiative (1996 and 1997) targeted seven of the city's most prominent gangs. The program's tactics included high visibility of police in gang areas, curfews for suspected gang members, and hard-hitting enforcement of any truancy legislation (Lafontaine et al. 2005).

Cooper and Ward (2008) observed that in some areas, suppression tactics resembled full military campaigns. For example, in Los Angeles, police initiated the Community Resources Against Hoodlums (CRASH) during the 1980s which involved 'gang sweeps' of gang areas and constant pressure on gangs through intense police patrols. The Order-Maintenance Policing (OMP) project, introduced in New York in 1994, adopted the Broken Window Theory (Wilson and Kelling 1982) approach. The Broken Window Theory proposes that if community standards are not maintained (e.g., broken windows are not mended quickly, litter is allowed to accumulate) then a community becomes vulnerable to criminal activity. The solution lies in community members and police cooperating by attending to minor community problems which should then avert an escalation of anti-social behavior. Using these methods, crime should be prevented, or at least confined whilst at a low level. The OMP project initially aimed to operate via collaborations between justice officials and community members and to favor alternatives to arrest such as education and counselling—all of which is in line with Broken Window theory tenets. In reality, OMP employed aggressive tactics of stop-and-frisk actions to detect less serious crime and gun possession (Fagan et al. 1998) and rejected alternatives to arrest and prosecution (Waldeck 2000).

Multifaceted Strategies

Also in the USA, some programs with a focused deterrence perspective adopted a problem-oriented policing approach which was thought to be particularly effective in reducing gang activity (Braga et al. 1999b). Although facets of this approach included suppression strategies, problem-oriented policing has a more multifaceted approach to tackling gangs by identifying a problem, analyzing it, devising an appropriate response and then evaluating that response to see if adjustments are necessary. For example, during the 1980s and 1990s projects such as The Boston Gun Project and Operation Ceasefire took a broad range approach (e.g., also targeting firearm traffickers who supplied gangs with weapons) to increase gang members' perceived risk of apprehension and prosecution. The main targets of these interventions were gang youth well known to the authorities and who generally had a history of convictions (Kennedy et al. 1996) and whose violence was driven by inter-gang disputes (Braga et al. 1999a). The program's aims were outlined to gangs via fliers, street work projects, and interactions with probation department staff. The suppression element of the message was that violence (particularly gun violence) would not be tolerated and a 'pulling all levers' approach (i.e., every criminal justice resource available would be used against those involved in violence) was adopted to suppress and deter gang activity. It was hoped that this level of suppression would encourage gang members to 'police' each other due to the threat that acts of violence committed by one individual would result in a police crackdown on all members of the gang.

However, unlike suppression-only programs, multifaceted programs included support services such as: probation and parole services, police services, prosecution branches (e.g., the Office of the U.S. Attorney), youth services (e.g., Juvenile Corrections), federal agencies (e.g., the Bureau of Alcohol, Tobacco, and Firearms), and community outreach programs/services. These multi-agency approaches encouraged gang members to take advantage of social service support to access: employment opportunities, mentoring schemes, housing support, substance abuse treatment programs, and vocational training opportunities (McGarrell et al. 2006)—all aimed at helping them to relinquish their gang lifestyles.

This 'carrot and stick' (i.e., using persuasion and force simultaneously) approach of multifaceted programs in tackling gangs appeared to be successful, and so other, similar problem-oriented policing programs, functioning on comparable deterrent tenets as Operation Ceasefire, were adopted in Richmond, Chicago, Los Angeles and Indianapolis. For example, the Chicago Alternative Policing Strategy (CAPS 1992) program, just as Operation Ceasefire had done, adopted a democratic police/citizen interactive approach to crime reduction. Police concentrated on their beat areas and got to know residents and the areas' problems. Meetings were held between residents and police to consider urgent crime-related needs and an advisory body of community leaders presented larger concerns to the area commander. In short, a democratic approach, involving all areas of the community, was adopted to tackle gang activity.

Program Efficacy

Zealous suppression approaches within criminal justice responses do appear, on the face of it, to have achieved at least some of their desired deterrent effects. Operation Hardcore increased charge and conviction rates with fewer dismissals of cases involving gang members. It was later judged as having achieved clear improvements in justice responses to gang defendants and their cases (Dahman 1983). However, its effects on gang-related crime was not evaluated specifically so we cannot know if gang-related crime reduced as a result of the program (Cooper and Ward 2008). In California, TARGET identified 570 gang members and 77 gang leaders during its first two years. Of these, two-thirds were held in custody, 99 % were convicted and there was an associated 62 % reduction in gang-related crime and a 47 % decrease in gang crime over seven years (Cooper and Ward 2008). In Dallas, an analysis of its anti-gang initiative showed that gang-related violence decreased significantly during 1996 and 1997 in the targeted areas (Howell 2000). Similarly, OMP in New York was associated with a dramatic reduction in crime and was met with a euphoria of support (Harcourt 1998), heralded as the Holy Grail of the 1990s (Jones 1997) and "…the new path of deterrence" (Kahan 1997, p. 2479).

On the other hand, some programs did not fare so well. Operation Hammer deployed 1000 police officers to conduct nightly gang sweeps and carry out mass arrests of gang members, but of the 1435 people arrested, nearly half were not gang members and 1350 were released without charge (Klein 1995). The operation, which was never formally evaluated, was subsequently abandoned in the face of public criticism (Klein and Maxson 2006). In Chicago, a lack of clarity regarding what constituted loitering resulted in the Gang Congregation Ordinance being accused of enabling arbitrary and discriminatory targeting by the police. It was eventually ruled by the U.S. Supreme Court (*City of Chicago v. Morales* 1999) as unconstitutional because its lack of clarity regarding what citizens may or may not do could curtail harmless behavior and thus violate people's constitutional rights. More recent research also shows that even if laws successfully target gang members they do not necessarily have the desired gang reduction effects, because arresting gang leaders can increase violence by triggering competition among adjacent gangs who aim to take over the territory of leaderless gangs (Vargas 2014).

Multi-faceted programs also appear to have had some success. For instance, the Chicago Alternative Policing Strategy (CAPS 1992) programme appeared to be promising, as there was a drop in crime rates. However, crime rates that showed a downward trend up until 2007 showed an increase in 2008 (Murphy 2008). An evaluation of Operation Ceasefire conducted by Braga et al. (2001) was lauded as demonstrating that the program was associated with a 63 % reduction in youth homicide in Boston (Wellford et al. 2005). However, as Wellford et al. (2005) also note, in the absence of a random controlled trial, it is difficult to attribute the reduction in youth homicide to the program. Also, other authors argue that the downward trend in gun homicides noted during Operation Ceasefire continued only until 1997; in 1998 it began to rise again (Fagan 2002). Fagan argues that since gun

homicides had similarly begun to decline during 1995 in other, non-program cities in Massachusetts, the trends noted in Boston cannot be attributed specifically to Operation Ceasefire. Further, attempts to replicate Operation Ceasefire in other jurisdictions (e.g., Los Angeles and Indianapolis) offered no evidence that the deterrence message embedded in programs had any influence on gang members' behavior (Greene and Pranis 2007).

Researchers have pointed out problems with the facilitation of programs. In many areas, even when a multifaceted approach to tackling gangs was adopted, problem-oriented policing was often not practiced and police rarely sought citizen input and rarely shared with, or valued information provided by, non-criminal organizations (Katz and Webb 2003). Instead, an emphasis was put on the suppression elements of programs as police departments adopted aggressive responses as a result of pressure from citizens to address the problem with gangs. However, this backfired, as hard-liner approaches often resulted in accusations that the police used excessive force, subjected citizens to unnecessary stop and search routines, and subjected citizens' children to harsh physical treatment.

Other suppression efforts by the police included the dispersal of gangs, and these were a universal failure (Katz and Webb 2003). For instance, in Las Vegas, the demolition of a public housing development resulted in gang members and their activities being scattered across the city. This led to a worsening of the existing gang problem because gang members continued with their previous behavior in their new residential areas. Katz and Webb (2003) further observed that in the jurisdictions they examined there was little police training, little monitoring, and no police accountability. Suppression tactics generally operated by increasing special gang unit patrols of gang areas. These special units were, however, composed of police officers who knew little about gang members, had little direct contact with gang members, and who made so few arrests that there was little impact on crime in the area. It is understandable that the police, when under the extreme pressure of public outrage about gang activity, may prioritize suppression strategies. However, what is concerning is that although gang unit police and key stakeholders believed that suppression activities justified the existence of gang units, it was only the gang unit police who believed that suppression efforts reduced gang problems (Katz and Webb 2003). Findings such as these have led some researchers to claim that an imbalance exists between the carrot and the stick in multifaceted approaches. For instance, Greene and Pranis (2007) noted that two-thirds of resources used for gang reduction had gone into suppression initiatives in Los Angeles and that "'Balanced' gang control strategies have been plagued by replication problems and imbalances between law enforcement and community stakeholders" (p. 6).

However, the delivery of programs is likely to differ according to jurisdiction and even program content. Braga and Weisburd's (2012) systematic review and meta-analysis of focused deterrence considered the findings of 11 evaluations of programs that employed focused deterrence strategies to reduce crime and observed that although ten programs reported statistically significant crime reductions, a lack of randomized experimental evaluations generated a tendency to overstate positive outcomes. Nonetheless, Braga and Weisburd's review provides overall support for

deterrence principles and notes that taking a multifaceted approach to youth violence generally—and gang violence more specifically—can dramatically hinder violence. Braga and Weisburd add that focused deterrence strategies succeed in adhering to procedural justice principles of mutual respect and consideration by creating a liaison between police and communities to increase social controls, to deflect offenders away from crime and consequently succeed in accomplishing a reduction in crime in the target community. This approach also helps to develop police legitimacy amongst the local community by increasing public awareness that citizens accept the rule of authority (Skogan and Frydl 2004) which, in turn, decreases the likelihood that people will break the law (Paternoster et al. 1997).

However, Braga and Weisburd (2012) observe that focused deterrence programs, which function both on enhancing individuals' risk perceptions through increased police presence *and* on encouraging pro-social behavior, make accurate program evaluation problematic. For instance, when suppression strategies are coupled with support (e.g., from youth workers, probation and parole officers, churches and community groups), evaluations of program efficacy are rather muddy and it is difficult to untangle the effects of complex program elements to identify exactly what effect each component has had. As multifaceted programs encourage gang members to take advantage of social service provision, employment opportunities, mentoring, housing, substance abuse treatment programs, and vocational training opportunities (McGarrell et al. 2006), we cannot draw firm conclusions about which elements of the program had deterrence effects. In their meta-analysis, Pratt et al. (2006) note how the effect sizes of deterrence are substantially reduced—sometimes to zero—when other factors such as peer influence, antisocial attitudes and self-control are controlled for in analyses. These authors go so far as to say that:

> …the clear drop in predictive power of the deterrence variables from bivariate to multivariate models suggests that empirical support for the effect of formal sanctions on individuals' criminal behavior is most likely an artefact of the failure to control for other "known" predictors of crime/deviance (p. 384).

Therefore, we cannot say with any certainty that the positive effects observed in some multifaceted programs occurred because they heightened individuals' risk perceptions or because they provided opportunities that enticed gang members into more prosocial activities, such as jobs. Indeed, the provision of increased support for youth in terms of social service provision, mentoring, housing, substance abuse treatments, and employment and vocational training may indicate that some of the core risk factors that have historically been linked to gang membership (i.e., social deprivation and exclusion—see Klein 1995) are being addressed in focused deterrence programs, or at least they are given more attention. However, it is assumed that it is deterrence, rather than increased opportunities that lead to program success. Either way, it is difficult to effectively gauge exactly what effects lead to reduced gang activity. This is because evaluations of the effects of support and increased opportunities are too often dogged by substantial methodological flaws (e.g., narrative reviews or descriptions of programs). This then prevents even the most rudimentary conclusions being drawn about program effectiveness (see Fisher et al. 2008, for a systematic review).

An evaluation that seems to offer more support for the effectiveness of deterrence principles, albeit with nongang populations, was conducted with a program that included strategies aimed at low-risk offenders on probation who were not substantial drug or alcohol abusers (Project MUSTER; Weisburd et al. 2008). To conduct the evaluation, Weisburd et al. (2008) used a randomized experimental design to compare the effects of the program on three groups of offenders. The first group underwent intensive probation, threats of violation to court, and incarceration *and* community service (Project MUSTER intervention group), the second group underwent normal probation supervision, and those in the third group were threatened with only one part of the MUSTER treatment, violation of probation (VOP—for non-payment of fines). Findings showed that probationers in the MUSTER cohort and those in the VOP group were more likely to pay their fines than the probationers in the normal supervision group. MUSTER and VOP participants did not differ from each other. The authors justifiably claim that their findings show that if threats of sanctions are associated with non-payment of financial obligations then offenders are more likely to be compliant with legal sanctions.

The Weisburd et al. (2008) study suggests that threats of legal sanctions do act as a deterrent with offenders. However, the offending population involved in MUSTER was, as noted, low-risk offenders. This then begs the question of how effective such strategies are likely to be with more criminally entrenched offenders. Gangs are widely accepted as being deeply engrained criminal groups, whose members are disproportionately involved in crime, particularly serious and violent crimes (Thornberry and Krohn 1997), and there is something about gang membership that promotes higher rates of offending whilst in a gang than before or after membership (Thornberry et al. 2003). It seems unlikely, then, that the promise that the MUSTER approach appears to offer would apply as well to gang members who are deeply embroiled in criminal activity as it does to low-risk offenders. In addition, participants in the MUSTER program were not substantial substance users whereas research indicates that compared to nongang youth, youth in gangs have far higher levels of substance use (Gatti et al. 2005). Consequently we may expect that the need for substances and the associated criminal acts employed to acquire those substances are likely to negate the influence of threats embedded in programs such as MUSTER. In short, analogies between participants in MUSTER and gang members are difficult to see.

Anti-gang Policies

Criminal Enhancement of Gang Activity

To support and complement the intended deterrent effect of anti-gang strategies and special gang prosecution units, anti-gang legislation has been introduced in a number of countries. In the US, at a federal level in (2010), the statute specific to

gangs, "Criminal Street Gangs" (18 U.S.C. § 521.), states that anyone participating in a street gang whose members have engaged in a continuing series of offenses such as violence against others, drug offenses, or conspiracy to commit such offenses, may receive sentences of up to 10 years imprisonment. At a state level, as of 2013, all 50 states and the District of Columbia had enacted some form of legislation regarding gangs and gang activity (National Gang Intelligence Center 2013). Of these, 28 had passed gang prevention laws, 30 had laws that provided enhanced penalties for gang-related crime, 26 states and the District of Columbia had legislation relating to gang activity in school settings, and 12 states had laws to deal with gang-related databases. Since so many anti-gang laws have been developed across the US, it is not possible for this chapter to describe them all, so a few illustrative examples will be used. For example, Nevada introduced anti-gang statutes in 1989 and 1991 which made it illegal to: possess weapons on school property; discharge firearms out of vehicles; aim a firearm at a person; discharge a weapon in a way that could endanger people; coerce minors into illegal activity; commit certain violations at or near schools, bus stops, or recreational facilities; and promote criminal gang activity (Lafontaine et al. 2005). In California, Florida, Georgia, Illinois, and Louisiana, the introduction of the Street Terrorism Enforcement and Prevention Act (STEP) enabled police and prosecutors to identify if a gang fits the Act's definitions and notify individuals categorized as gang members that the gang has been identified and that any subsequent law violations will be treated more harshly than usual (Cooper and Ward 2008).

Civil Gang Injunctions

In addition to gang-specific laws, Civil Gang Injunctions (CGIs; i.e., court-issued restraining orders) were introduced to prohibit and deter gangs from specific legal activities (e.g., loitering at schools, carrying pagers, and riding bicycles) or from entering certain areas. CGIs have been gaining in popularity since the 1990s in the U.S. and in the UK where they were introduced as part of the Police and Crime Act (2009). In the UK, between 2011 and 2014, 88 gang injunctions were put in place in 25 areas identified by the government as Ending Gang and Youth Violence priority areas (Home Office 2014). UK gang injunctions aim to "prevent a person from engaging in, encouraging or assisting gang-related violence and to protect them from gang-related violence" (Home Office 2014, p. 3).

In the US, CGIs predominantly target gangs' territoriality (Hennigan and Sloane 2013). They typically prohibit/limit movement within known "gang" areas and the scope of the CGI can cover a single street or an entire city. For instance, one injunction in Los Angeles covered 16 square miles (Los Angeles County District Attorney's Office 2011). The composition of a CGI involves a good deal of flexibility and variation (Maxson et al. 2005). The behaviors listed may be tailored to address the specific behavior of specific gang members in specific areas and/or they may target entire gangs by name. They may also forbid gang members from

associating with each other by standing, sitting, walking, driving, bicycling, gathering, or appearing anywhere in public with any other named individual or any other known gang member (O'Deane 2011). Police have the discretion to name 10–30 specific people in the injunction but hundreds of unidentified 'John Does' may also be cited—and identified and named at a later date (Myers 2009). All individuals identified as members of a gang issued with a CGI are threatened with arrest if they transgress any part of the order and they face enhanced sentences of ten years following conviction. Between 1980 and 2000, 37 CGIs were issued in Southern California and 22 in Los Angeles (Maxson et al. 2005), and by 2003, 47 % of African American men in Los Angeles County aged between 21 and 24 had been included in the Los Angeles County gang database (Siegel 2003). By 2013, Southern California had 46 CGIs targeting 72 neighborhoods in Los Angeles (Office of the City Attorney of Los Angeles 2013).

The use of CGIs has been criticized for the harsh effects that they have on youth identified as gang members. First, there is no consistency in how gang members are identified (Caldwell 2010) and others have accused officials of using injunctions to criminalize ethnic minorities so as to control their movements (Muniz 2014). Other accusations include that CGIs are used to dehumanize gang members to the point where members of the public take little note when they are shot by police (Muniz 2014). Injunctions are also cited as causing already marginalized individuals to be even further marginalized by being prevented from contributing to positive community activities such as family activities outside the home and even hindering employment prospects (Caldwell 2010). Once labelled as a gang member and subject to a CGI, individuals also have few options. Since CGIs are civil orders, those subject to them are not entitled to a public defender to help them appeal the order unless they are already on probation or parole (Muniz 2014). Consequently, once named in an injunction, those labelled as gang members have an uphill, if not impossible struggle to remove their names from an injunction (Crawford 2009) and in practice few have managed to do so (Hennigan and Sloane 2013). It is not surprising then that some authors argue that labelling an individual as a gang member encourages them to adopt the identity (Huff and McBride 1993).

However, if CGIs achieve the desired deterrent effect then they are likely to continue to be supported in criminal justice circles—regardless of the criticism that they attract. In deterrence terms, the knowledge that emanates from a gang injunction (i.e., that *individuals* are being closely monitored) should, according to deterrence principles, prevent gang members from involvement in gang-related activities because they will heighten gang members' risk perceptions. Evaluations of gang injunctions suggest that they do indeed achieve some of the desired reductions in crime. For instance, a comparison of crime data from 14 areas in Los Angeles covered by injunctions with matched comparison neighborhoods found that violent crime fell by 5–10 % in injunction areas during the first year of injunction use. There was also no evidence that crime had been displaced to adjoining areas (Grogger 2002). A Los Angeles Civil Grand Jury examined the effect of gang injunctions and reached similar conclusions in (2004). An additional explanation as to why gang injunctions may reduce crime is that they promote

changes in the community. For instance, they are thought to increase community cohesion, collective efficacy, and community members' motivation to call the police should they feel threatened by gang activity (Maxson et al. 2005). Maxson et al. (2005) also observed a corresponding reduction in residents' fear of, and intimidation by, gangs. However, of the two gang injunction neighborhoods examined by Maxson et al. (2005), only one experienced these improvements and this was the one which had had a gang injunction the longest. The second neighborhood experienced more gang visibility, more anxiety amongst residents, more social disorder, more property victimization, and less belief that the neighborhood could resolve its problems. Maxson et al. (2005) note that the disparity between the two neighborhoods could be due to: (1) gangs from the original injunction neighborhood spilling their activities to the second neighborhood; (2) methodological issues; or (3) police suppression backfiring and generating gang oppositional defiance (see Jankowski 1991) and/or strengthening gang members' cohesion as suggested by Klein (1995) which, in turn, increases violent crime (see Klein 2011; Wilson and Chermak 2011).

In California the American Civil Liberties Union (ACLU 1997) analyzed the Blythe Street gang injunction and concluded that the injunction did not result in a reduction in violent crime and/or drug trafficking. Also in California, of the 80 % of gang members who committed offenses after being named in an injunction, over half committed crimes in the injunction area (Hale 2006). This suggests that injunctions do not deter since gang members simply continue to offend either in the injunction area or in adjoining areas (Myers 2009). It is also proposed that even when gang injunctions are effective in deterring gang activity, their impact may not be enduring and so research needs to examine reductions in crime for more than a year (Grogger 2005). Other authors point out that the lack of evidence of any long-term improvements (i.e., longer than one year) begs the question of whether gang injunctions are responsible at all for *any* crime reduction and that in some communities no significant reduction in crime has occurred, and even where reductions do occur other factors could be responsible for observed changes (Caldwell 2010). Maxson et al. (2005) caution that any gains made by gang injunctions will be continually threatened by the persistence of gang activities and that gang injunctions (i.e., the stick) are an opportunity to bring about change, but positive opportunities (i.e., the carrot) may be necessary before lasting changes to gang activity will be effective.

Grogger (2005), on the other hand, seriously doubts that deterrence effects are responsible for any observed reductions in crime. As Grogger notes, gang members are named in injunctions due to police suspicions that they have been involved in serious crimes such as robberies, drug offenses, or rape, and this means that they have already run the risk of attracting serious and lengthy sanctions. Therefore, Grogger doubts that threatening gang members, who seem willing to run the risk of lengthy sanctions, with lesser threats for transgressing gang injunctions (maximum of 6 months), will deter gang members' activities. More recent findings support Grogger's contentions by showing that deterrence principles (i.e., the threat of arrest and punishment) have less influence on gang members than on nongang

youth living in CGI areas (Hennigan and Sloane 2013). Instead, Hennigan and Sloane found that CGIs seemed to disrupt the strength of members' gang identity and because identity mediated the relationship between a gang's cohesion and its criminal activities, gang members living in a CGI area compared to gang members living in a non-CGI area were involved in lower levels of crime.

The findings regarding the success or failure of CGIs are, to say the least, equivocal. Even if gang injunctions succeed in initially denting gang activity as gang members warily adjust to this new initiative in their communities, there is no evidence so far to suggest that such dents have longevity. Equally, residents in communities where gang injunctions are issued may *initially* perceive the heightened police interest in gang activity in their neighborhood as a motivation to report gangs and crime, but again, there is no evidence so far that this effect is enduring. As Grogger (2005) suggests, more research is needed to examine the perspectives of residents in gang injunction communities before conclusions can be made regarding the effects of this legislation overall. In short, any effects that injunctions might have might simply wear off over time and we need more research before any firm conclusions can be drawn.

The Doctrine of Joint Enterprise or Common Purpose

A further policy that targets gang members' activities is the doctrine of joint enterprise or common purpose which is employed in several countries. For the purpose of this chapter, discussion will focus particularly on the UK where this 300-year-old law has undergone something of a popular revival in recent years. Originally introduced to combat illegal practices such as dueling, the doctrine holds responsible all individuals present at the time for a crime and has been enthusiastically applied to gang members. For example, in the UK, gang members (secondary parties) who are present at the time of an offense (e.g., homicide) but who do not participate in the actual crime, can be held as responsible as the gang member (principal offender) who performs the conduct element of the offense (murders the victim). By being held as responsible as the gang member who committed the offense, secondary parties will also be convicted of murder, which in the UK carries a mandatory life sentence. These secondary liability principles—that is, the shared responsibility of principal and secondary parties—can be applied to any offense but are used principally for offenses involving violence. In short, it means that all those present at the scene of a crime may be equally sentenced even if they all had different roles in the offense. There is a great deal of unease regarding what is often called the 'parasitic liability' component of joint enterprise as this means that if all those involved in a crime (e.g., robbery) are aware that one of their group may, during the original crime, commit another crime (e.g., murder) then all are guilty of that murder; proof of their participation in that murder is not necessary (McClenaghan et al. 2014). In essence, the doctrine functions on the presumption of

group members having full and pre-existing knowledge of each other's intentions and actions.

In the UK, it is estimated that almost 500 people have been convicted of murder as secondary parties between 2005 and 2013 (The Guardian 2014), but no formal records are kept of the actual numbers of prosecutions and convictions brought under the joint enterprise doctrine (House of Commons Justice Committee 2014). Also in the UK, in 2011, a Justice Select Committee set up to examine joint enterprise practice and effects heard from a number of people that joint enterprise was often used to address gang incidents as it enabled the prosecution of several members of one gang at a time for any member's criminal activity (McClenaghan et al. 2014). Given that research shows how gangs may include tens, and in some instances hundreds of members (see Klein and Maxson 2006 for a fuller discussion of gang sizes in the USA), it defies reason to assume that *all* gang members can possibly be *fully aware* of other members' intentions. Even though it is unlikely that such vast numbers would be present at the same event, it seems highly unlikely that all who are present, especially if they happen to be peripheral or fringe gang members, will be fully aware of others' intentions. As Jeremy Corbyn, the Member of Parliament appointed as the Justice Select Committee's special rapporteur (an independent expert) on joint enterprise and a key figure in the governmental inquiry, stated to McClenaghan et al. (2014):

> I came to the issue because of dealing with young people in my constituency who are peripheral members of gangs. These are young people who are peripheral to some, often horrific, incidents. That doesn't make them all guilty. It makes them in the wrong place at the wrong time, often, but it doesn't make them all guilty. Unless there is incontrovertible proof that they took an active part in an attack then we end up prosecuting and ruining the life chances of quite a lot of young people, who are frankly bored and hanging around the streets (p. 29).

Research conducted by Eady (2013) for the organization Joint Enterprise Not Guilty by Association (JENGbA) adds to the argument that joint enterprise prosecutions often involve inconsistencies that are difficult to comprehend. In this study, of the 101 people convicted under the joint enterprise doctrine, 70–80 % of those aged under 22 maintained that during trial the prosecution mislabeled them as gang members. Eady argues that there are potentially four explanations for why the prosecution would allege a defendant's gang involvement: (1) The defendant is a gang member and denying membership is less destructive to their case; (2) Prosecutors perceive gangs are dangerous and responsible for many violent crimes; (3) The prosecution gains a presentational advantage in portraying defendants as gang members; (4) The idea of gang responsibility helps to justify shared responsibility and joint enterprise prosecutions. Interestingly, this study also flagged sentence inconsistencies in which defendants who denied being present at the offense or even knowing the perpetrators received the longest prison sentences (average, 24.54 years) whilst those who admitted being present at the offense, but claimed involvement only in legal activity received an average sentence of 17.65 years, and those who admitted being involved in the offense but to a lesser extent or committing a lesser crime received an average sentence of 18.64 years. As Eady notes,

> In most cases where the Respondent claimed not to have been present, the basis for conviction seems to lie in the Respondent's previous history or associations rather than in the prosecution claiming that they were in fact present. In around 30 % of these cases mobile phone records provided evidence of this association but not, it is argued by Respondents, of involvement with the crime. (p. 11)

In its joint enterprise follow-up report, the House of Commons Justice Committee (December, 2014) examined evidence provided from a number of sources to discern the suitability of applying the joint enterprise doctrine. Amongst the evidence examined in the follow up report was the potential of joint enterprise laws to deter youth from becoming involved in group offenses. Providing oral evidence Simon Natas (a solicitor) argued that:

> The use of punitive law and order policies to combat youth crime is a blunt instrument which runs the risk of making matters worse, not only because young people are much less likely to cooperate with the authorities if they perceive the legal system to be unjust, but also because custodial sentences have been shown to increase, not reduce, reoffending (p. 13).

Giving evidence to the hearing, researcher Dr. Ben Crewe from the Cambridge Institute of Criminology, added to this point by pointing out that for deterrence to be effective people have to be aware of the sanction and that this awareness must shape their risk perception. In their research, Crewe et al. (2015) note how only few of the youth interviewed had any idea what joint enterprise was, even when they had been prosecuted under the law, and even those who were aware of it had limited understanding of its implications. Crewe pointed out that the main deterrent effect of any law functioned on people's belief in the certainty of being caught. Consequently, legislation that holds all gang members responsible for an individual member's actions will have little deterrent impact if the two main principles of deterrence (i.e., awareness of sanctions and a corresponding increase in risk perception) are not understood by the target populations. The Committee concluded that publicly available information about the use of joint enterprise was incomplete and that the use of joint enterprise needs an urgent review since the effects of the doctrine are likely to be negative for the reputation of the justice system, wider society, those convicted and victims of crime.

When reviewing the statistics on gang membership, we can see that stringent police tactics and governmental policies such as those outlined above do not succeed in deterring gang membership. Although the exact numbers of gangs and gang members is largely unknown in the UK, a policy report (Centre for Social Justice 2009) highlighted a host of research findings that suggest that gang culture and its associated violence had increased in the previous decade, with up to 6 % of 10–19-year-olds claiming to be gang members (Sharp et al. 2006) and at least half of the murders of young people in London during 2007 relating to gang activity (Metropolitan Police Authority 2008). A policy paper by the Centre for Social Justice (2012) published after the London riots of 2011 goes on to note how raids leading to the arrest of more senior gang members had left vacuums that younger gang members fought to fill, leading to a corresponding increase in violence as younger gang members "….vie for status and respect using the currency of

violence" (p. 7). Echoing the conclusions of their earlier report (Centre of Social Justice 2009), the policy paper maintains that the police cannot arrest their way out of the gang problem. The report also argues that preventative work that could discourage young people from gang involvement has 'fallen off the radar' (Centre of Social Justice 2012, p. 3).

Unsurprisingly, in an atmosphere of suppressive and punitive prosecution tactics, relations between the police and youth in the UK have experienced repeated blows as police who work on databases of 'permanent suspects' label and target youth as gang accomplices if they are seen to associate with 'known' gang members, even if they have committed no offense (Ralphs et al. 2009). In gang areas where it is unlikely to be possible or wise for non-gang youth to avoid *any* contact with gang members for reasons of personal safety, non-gang youth run the risk of attracting police suspicions that they are gang members. Non-gang youth then view the associated constancy of being stopped and searched as police harassment. In turn, these anti-police attitudes manifest into negative consequences for all because police lose any co-operation they could have had from non-gang youth, and the youth suffer barriers to leading a normal life by being prevented from attending social events and sometimes even school (Ralphs et al. 2009). This may then lead into a further negative effect and encourage non-gang youth to join gangs as noted by Huff and McBride (1993; see above). It is unsurprising then that government policies and the associated police strategies have received scathing criticism that they fail because they do not take an evidence-based approach to gang reduction (Shute and Medina 2014).

In the US, gang figures bode just as badly as they do in the UK for oppressive police and government strategies. The National Gang Intelligence Center (NGIC 2011) notes that gangs are expanding across the U.S. with approximately 1.4 million active gang members belonging to more than 33,000 gangs which pose a threat to communities nationwide. The NGIC goes on to note how gangs are expanding their criminal profiles and that even when incarcerated gang members are not deterred from continuing gang involvement via friends and family; a finding also noted in UK research (see Wood et al. 2010). None of this bodes well for the effects of deterrence on gang membership as it seems that policy, suppression and even threats of punishment fail to deter as intended. As the NGIC noted in 2011:

> Local neighborhood, hybrid and female gang membership is on the rise in many communities. Prison gang members, who exert control over many street gang members, often engage in crime and violence upon their return to the community. Gang members returning to the community from prison have an adverse and lasting impact on neighborhoods, which may experience notable increases in crime, violence, and drug trafficking (p. 11).

Consequently, it appears that police, policy, and even punishment deterrence strategies are ineffective, at least in an overall sense because gangs continue to thrive, adapt, and evolve. However, there are many potential explanations as to why deterrence effects are apparently ineffective in addressing what appears to be a global and flourishing phenomenon.

Why Deterrence Strategies May Not Be Effective with Gang Members: Psychological Explanations

Little research has directly examined gang membership and deterrence effects. What research has shown so far is that gang members are unlikely to be deterred from carrying and using guns (Watkins et al. 2008), unlikely to perceive certainty of arrest, are not influenced by the potential severity of punishment and expect little condemnation from significant others for offending (Maxson et al. 2011). Other research shows that deterrence efforts have less impact on gang members than on other youth (e.g., Hennigan and Sloane 2013). However, what we do not yet fully understand are the reasons why gang members hold such perceptions of chance of arrest, feel a need to carry weapons, and are seemingly immune from the effects of deterrence. Gangs are groups that have the potential to exert a powerful and enduring influence over individual members by shaping the way that they think and behave. This makes a closer consideration of the psychological effects of gang membership on individual gang members vital to the analysis of deterring gang membership. However, to date, a comprehensive psychological exploration of gang membership remains remarkably under-researched. The following section considers issues of identity, cohesion, morality, status enhancement, and oppositional culture, which are a few of the many possible psychological explanations that could help us understand why deterrence strategies may not be effective with gang members.

Social Identity

Research indicates that youth who feel alienated from legitimate social controls such as families, education, and prosocial community endeavors are those most likely to join gangs (e.g. Marshall et al. 2005). As they reject or reduce familiar childhood groups and practices in favor of gang activities, they are likely to feel uncertain about their attitudes, their future, and importantly about who they are (i.e., their identity). *Uncertainty-identity theory* (Hogg 2000) explains that when people feel uncertain about their personal identity they are motivated to affiliate with a group and, in line with social categorization principles (Abrams and Hogg 2010; Tajfel and Turner 1986), they use their group membership to categorize themselves according to the attitudes and behaviors that epitomize that group. Since gangs offer friendship, pride, a sense of identity, heightened self-esteem, excitement, and potentially access to financial benefits (Goldstein 2002) it is not difficult to see why youth struggling with legitimate controls (e.g., school) and personal identity are attracted to gang membership. According to *social identity* (Tajfel and Turner 1986) and *self-categorization theory* (Turner et al. 1987) the shaping of group members' self-views and personal identity continues during membership until harmony between personal and group identities is achieved.

Once the process of intertwining personal identity with the gang's identity is underway, gang members may begin to put the group's aims first. For instance, Hennigan and Spanovic's (2012) examination of social identity and gang membership showed how, compared to nongang youth, gang youth who identify with their group put the group norms of criminal activity ahead of personal fears of punishment for criminal activity. As the authors note, "Since crime and violence are normative among gang-involved youth, personal estimates of getting caught and punished have little or no influence on their criminal and violent behaviors" (Hennigan and Spanovic 2012, p. 143). Hence, police, policy and punishment practices that aim to deter gang involvement are likely to be ineffective with youth who disregard threats of punishment in favor of adhering to group norms.

Pluralistic Ignorance

Of course it is always possible that gang members may not actually agree with many of the group's norms. However, research suggests that even if group members do not fully accept or agree with group norms, they will still adhere to them. The concept of *pluralistic ignorance* suggests that even when individual group members do not agree with their group's activities, they still publicly accept and abide by them because they believe (often wrongly) that the rest of the group does agree with the activities (O'Gorman 1986; Reid et al. 2005b). There is also a great deal of social psychological evidence showing that when people want to be accepted by a group, they will conform to others' decisions—even if they believe those decisions are wrong (Cooper et al. 2004). Research examining gang members' reports also suggests that they experience pluralistic ignorance. For instance, even when gang members express privately that they feel extreme discomfort with some of their criminal activities—they still participate in those activities (Matza 1964). However, social psychological findings suggest that as group members' identification with the group intensifies, they increasingly believe in the group's norms (Reid et al. 2005a). Also, as violence is an intrinsic part of gang life, gang members may be inclined to adhere to ingroup norms due to fears of being on the receiving end of violent sanctions that their gang imposes on norm-violating members (Rimal and Real 2003). This threat is also more immediate and potentially greater than any that the criminal justice system poses. Plus, since the potential rejection of friends or admired others is an especially threatening aspect of norm violation (Baron and Kerr 2003), youth who are vulnerable, socially disenfranchised and whose self-esteem is embodied in the group are likely to feel any threat of group rejection keenly. In short, a gang is likely to wield its own form of deterrence on members (i.e., deterrence of law-abiding behavior) and this may well trump the deterrence efforts of the criminal justice system.

Cognitive Dissonance and Moral Disengagement

Consequently, given the power of ingroup influences and the eagerness with which aspiring gang members accept them, it is likely that gang youth, unlike nongang youth, feel compelled to abide by group norms—such as crime and violence (Viki and Abrams 2012)—even if individual criminal justice deterrence effects on that member are strong and his/her risk perception is high. Nonetheless, any unease that members may feel regarding their involvement in acts of violence is likely to cause them internal turmoil. *Cognitive dissonance* is a state of stress or mental discomfort that results when people's behavior, thoughts and attitudes are in conflict. Because cognitive dissonance is uncomfortable, people are motivated to reduce it (dissonance reduction) by bringing thoughts/attitudes in line with their behavior (Festinger 1962). By not rejecting gang norms and continuing to be involved in gang activity (behavior), gang members are likely to adapt their thoughts and attitudes to bring them in line with their behavior to achieve cognitive consistency.

One strategy that gang members may use to reduce their dissonance is to set aside their existing moral standards (i.e., morally disengage). *Moral disengagement* is a strategy that involves the use of up to eight strategies (e.g., moral justification, euphemistic language, dehumanizing victims) to justify harmful acts and resolve the cognitive dissonance and self-condemnation associated with violating personal moral standards (Bandura et al. 1996). Moral disengagement involves eight socio-cognitive mechanisms which operate on three levels of social processing. The first enables inhumane acts (e.g., violence) to be reinterpreted as worthy. Socio-cognitive mechanisms at this level include: moral justification (behavior is for a good reason—i.e., furthering gang status), euphemistic language (sanitizing descriptions of harm—e.g., violence is just "gang business"), and advantageous comparisons in which personal behavior is favorably compared with others' apparently worse behavior (e.g., our group only assaults—others kill). The second level allows the displacement of responsibility for personal actions onto authority figures (behavior is at the direction of authority figures which then negates personal responsibility), diffusion of responsibility (responsibility for harm is shared by several perpetrators and so individuals are absolved from blame), and distorting the consequences of harm (ignoring, minimizing, or disbelieving that harm has been done). The third level distorts the view of any victims via *dehumanization* processes (the victim is thought of as subhuman and thus devoid of human qualities—see also Alleyne et al. 2014) or via victim *blaming* (they got what they deserved). Research has shown that in order to be accepted by a chosen group, youth may set aside their moral standards (Emler and Reicher 1995), and gang research shows that moral disengagement is higher in street gang members (Alleyne and Wood 2010; Esbensen and Deschenes 1998; Maxson, et al. 2011) and in prisoners involved in prison gang activity (Wood et al. 2009, 2014) than it is in non-gang youth or non-gang prisoners. Moral disengagement strategies have also been empirically linked with increased violence (Bandura et al. 1996), and have been shown to mediate pathways between impoverished neighborhoods strongly associated with

gang membership and anti-social behavior (Hill et al. 2001) and between low levels of empathy and anti-social behavior. As Hyde et al. (Hyde et al. 2010) observe:

> In more modern contexts, urban youth living in impoverished homes and neighborhoods that offer them little hope or opportunity for socially acceptable pathways to success may develop a moral code of behavior that is not bound by mainstream prohibitions against committing antisocial actions, particularly when such actions are associated with the means to obtain financial success (e.g., dealing illicit drugs) or ensuring safety (e.g., joining a gang). (p. 198)

Group Commitment and Cohesion

As strategies such as moral disengagement embed and gang members begin to think, as well as act alike, the group will become more cohesive. Klein (1995) considers cohesion to be "…the quintessential group process" (p. 43; see also Wood 2014 for a fuller discussion) which influences members to become loyal, committed and ready to make sacrifices for a group that they view with pride and respect (Crocker et al. 1994). A meta-analysis examining cohesion concludes that highly cohesive groups are more productive than those that are less cohesive (Evans and Dion 1991) and, as Klein (1995) observes, gangs produce crime. Importantly, this effect is enduring—continuing even after members have left a gang (Pyrooz et al. 2014). Group commitment and the ensuing cohesiveness are not only likely to result in efficient criminal activity; they are also likely to have negative connotations for criminal justice efforts to deter gangs and their activities, since highly cohesive gangs are efficient at mobilizing membership and accessing commodities such as drugs and weapons (Hughes 2013). Such is the effect of cohesion that gang experts such as Klein and Maxson (2006) contend that it is their "….very strong opinion that levels of gang cohesiveness correlate directly with levels of gang crime and with gang responses to our efforts at gang control. Greater cohesion leads to greater crime involvement and greater resistance to gang control" (p. 196).

Criminal justice officials not only have to contend with gang cohesion potentially undermining their deterrence efforts, they also have to consider that gangs may well have their own agendas—primarily those involving other gangs—and these agendas may take priority over individual members' personal risk perceptions. Gang researchers note how a gang's identity is often defined in reference to other gangs—particularly rival gangs (Papachristos et al. 2013). Outgroups may be used as points of reference as a gang assesses its actions and, importantly, its status (Decker and Van Winkle 1996). *Social dominance theory* (*SDT*: Sidanius and Pratto 1999) explains that when group members have high *social dominance orientation* (SDO) they feel compelled to enhance, or reinforce, their group's place in a social hierarchy of groups (such as gangs) which have been arbitrarily constructed to respond to situational factors such as competition for valued resources. So, for example, gangs may compete with each other in an effort to enhance or reinforce

their group's status regarding control over valued illegal resources (e.g., narcotics). Although research examining group processes such as SDO in the context of gangs is still in its infancy, findings so far indicate that, compared to nongang individuals, those involved in gangs and gang activity have higher levels of SDO (Densley et al. 2014; Wood et al. 2014).

Reputation and Status Enhancement

As gangs work to assert/enhance their group's status in this arbitrary-set system they are likely to respond violently to rivals who threaten their aims (Aldridge et al. 2008; Decker and Van Winkle 1996). Densley's (2013) research shows how gang members see violence as the fundamental response to protecting territory and/or gang business. Accordingly, gangs' intergroup enmity is funded by violence which stems from competition for power, domination, reputation, respect, and status (Harding 2012). Intergroup transgressions cannot be ignored if a gang is to maintain its status. This makes status, together with previous conflicts, the most common reasons for intergang violence (Hughes and Short 2005; Papachristos 2009). In turn, violence against rivals helps a gang save face, protects its members by discouraging attacks from rivals, and exacts revenge on opposing groups for any transgressions (Papachristos et al. 2013). Criminal justice efforts to deter gang members' involvement in gang activity are therefore likely to encounter the considerable obstacles of status, reputation, and resource-motivated commitment.

Being a gang member can also enhance individual reputation and status. We know from research that status is highly prized by both street gang members (Alleyne and Wood 2010) and prisoners involved in prison gang activity (South and Wood 2006; Wood et al. 2009, 2014). *Reputation enhancement theory* posits that group membership facilitates individual behavior as members select an image to display in front of specific others (Emler and Reicher 1995). These others then provide positive feedback that reinforces the individual's image within the group. For gang members, crime is likely to be key to a positive personal reputation in a gang world where criminal activity is a prized gang product. Violence in particular will also protect individual members from future victimization (Emler and Reicher 1995). In their study of gangs in the UK, Harris et al. (2011) note:

> Not reacting with often extreme violence was experienced as tantamount to abject failure. There was a sense of being worse than nothing if a once-held status is lost. This was not only due to loss of respect, but also a sense of inevitable attacks and victimization from others. (p. 20)

In a culture of inter-group rivalry where it is important to protect personal and group status and criminal enterprises it is unlikely that criminal justice efforts to deter gang activity via suppression strategies and their accompanying policy infrastructure will have enduring effects. They may well have temporary influences as gangs, wary of new police tactics, effectively lie low—for a while, but the

durability of these effects is likely to be short lived. Gangs have their own agendas, be they social, criminal, or conflict-related, and they are likely to adapt so that they can adhere to these agendas. It also needs to be borne in mind that, even if deterrence strategies are effective in heightening individual gang members' risk perceptions, gang members are also prospective victims—particularly from rival gangs (Sanders 1994). So, even if gang members feel deterred from gang activity and want to rescind their membership to lead law-abiding lives, the probability of continued victimization by rival gangs may compel them to continue their membership to gain the protection that their gang can offer (Vigil 1988).

Oppositional Culture

By employing deterrence strategies, the authorities also run the risk of feeding an "oppositional culture" that has long been acknowledged as endemic in gang norms. *Oppositional culture* means that the group is set in opposition to legitimate authorities (Moore and Vigil 1987). As Short and Strodtbeck (1965) observed, "…. gang members are less favorably disposed toward adult incumbents of legitimate roles such as teacher, religious leader, policeman, businessman, and politician than are their non-gang, lower class counterparts" (pp. 275–276; see also Drury 2010). Research examining gang membership shows how street gang members (Alleyne and Wood 2010), and those involved in prison gang activity (Wood et al. 2014), hold strong anti-authority attitudes. Although many adolescents hold anti-authority attitudes (Drury 2010) which are therefore not unique to gang members, the targeting of gangs in gang suppression programs is likely to bring gang members into frequent and negative contact with authority figures which, in turn, will exacerbate their hostile views of authority. The net effect here is that gangs may come to view themselves as unfairly targeted victims of oppression (Lien 2005), which will help to reinforce their gang identities (McAra and McVie 2005; Ralphs et al. 2009), strengthen their oppositional culture (Klein and Maxson 2006), and may even increase the number of gangs and hence, criminal activities (Hagedorn 2008) as gangs adopt even greater criminal responses in an effort to defend their group identity (Ayling 2011). Members may consider themselves as defenders of their group which they see as being victimized. For instance, identifying this effect in Norwegian gang members, Lien (2005) notes:

> He (*sic*) develops ideas of compassion, love, and sacrifice in relation to his friends, and he explains his acts through a construction of himself as a victim of society. The victimization point is necessary in order to justify the criminal act. He cannot be blamed, the act is heroic rather than evil, and the victims get what they deserve. (p. 121)

As Lien (2005) explains, gang members develop 'heroism' for their group and for what they perceive as their victimization by society. Victimizing opponents is then all the more justified as each gang member strives to demonstrate his/her loyalty to the group and to fulfil his/her perceived obligation to other members. And, in a climate where abject loyalty, cohesion, identity, self-esteem and personal

protection abound it is easy to see why gang members may not be as vulnerable to deterrence strategies as criminal justice agents would wish.

Concluding Remarks

The aim of this chapter was to examine the effectiveness of deterring gang members from their gang lifestyles. To this end, it began by considering some of the myriad of criminal justice initiatives that, at great cost to public purses, have striven to deter gang activity and gang membership. However, so far neither programs nor policies seem to have successfully hit the mark and deterred gang activity in any enduring way. *Some* programs appear to have *some* effects on *some* gangs, leading to euphoric claims that they are effective deterrents (e.g., Kahan 1997). However, this euphoria was premature. A lack of methodologically sound evaluations and accusations that deterrent effects are, at best short-term and at worst counter-productive, has corroded the initial enthusiasm that deterrence programs provide a panacea to the gang problem. Zealous approaches such as suppression tactics and the variety of policy infrastructures supporting them have, all too frequently, backfired amidst accusations of an excessive use of force on citizens—gang and nongang—and unfair/unsound convictions of individuals who are guilty of failing to anticipate others' actions (e.g., Eady 2013). Unsurprisingly, these programs and policies have undergone several incarnations as justice officials and politicians strive for amendments that will uphold justice principles.

Even programs that include carrots as well as sticks and aim not only to suppress gang activity but also to offer gang members viable nongang alternatives, have struggled to keep afloat (e.g., Greene and Pranis 2007). As appealing as these multi-faceted programs are inasmuch as they also help address key gang-related risk factors such as social disadvantage, they have been dogged by accusations that they lack balance between their carrots and their sticks and these have undermined their program integrity. If we feed into this a lack of consistency between and within programs and the methodological problems which have consistently hindered conclusive program evaluations and definitive meta-analyses, we reach a point where, despite vast resource expenditure, we still know little, if anything, about what is truly effective in reducing/preventing gang membership.

In a climate where gang membership appears to be flourishing despite deterrence strategies, a further aim of this chapter was to outline some of the reasons *why* gang members may not be deterred from gang membership and activity. There is little research directly examining deterrence effects with gang members and little research that specifically examines the psychological effects that gangs have on their membership. Although gangs may appear to be homogeneous groups with homogeneous members—they are not. Individual differences mean that gangs and gang members are heterogeneous and that strategies that work with one gang member may fail with another. Consequently we must consider gang members as individuals—and work to devise interventions tailored to their specific needs.

Despite a host of social psychologists (e.g., Abrams and Hogg 2010; Hogg 2000; Tajfel and Turner 1986; Turner et al. 1987) noting the powerful impact that group processes have on individuals, little work has examined the specific nature of those group processes within the gang context (Wood 2014). Effects such as identity development, pluralistic ignorance, resolution of cognitive dissonance via moral disengagement strategies, and the development of group cohesion are all likely to have powerful roles in inoculating gang members from criminal justice efforts to deter their behavior. As noted, gangs have their own agendas and even if those agendas are as simple as wanting to 'hang out' together, gang members will, due to group process influences, stick to their agendas, regardless of the threat of negative consequences (e.g., Hennigan and Spanovic 2012).

Deterrence strategies operating on threat are unlikely to be effective. Gangs are all too familiar with threats, and threats that are more immediate and potentially deadly, such as those radiating from rival gangs, may take precedence over threats that emanate from criminal justice systems. As gangs perceive the multitude of threats leveled at them they are likely, as researchers such as Klein (1995) argue, to become cohesive and the more cohesive a gang is, then the more likely it is to continue, develop and adapt to the cultural climate that undulates around it. We are already seeing gangs adapting to social media methods of functioning (see Patton et al. 2013), and it is reasonable to expect that gangs will continue to adapt and change other fundamental operational methods in response to perceived threats. However, if research can help us to understand more about gang members and identify with robust empirical examinations the group processes that profoundly influence gang membership then we will have the infrastructure to devise effective interventions to tackle gang membership at an individual level. So far, criminal justice systems have attempted to hamstring gangs with suppression strategies, to prosecute gang members collectively via joint enterprise policies, and to offer alternatives to gang life via community support measures; yet, gangs still seem to thrive. Although understanding more about the psychology of gang membership is a research agenda that is in its infancy, more research examining the psychological influences of gangs on gang members is crucial if we are to adequately and consistently deter gang membership and gang activity.

References

Abrams, D., & Hogg, M. A. (2010). Social identity and self-categorization. In J. F. Dovidio, M. Hewstone, P. Glick, & V. M. Esses (Eds.), *The SAGE handbook of prejudice, stereotyping, and discrimination* (pp. 179–193). London: Sage.

Aldridge, J., Medina, J., & Ralphs, R. (2008). Dangers and problems of doing 'gang' research in the UK. In F. van Gemert, D. Peterson, & I. -L. Lien (Eds.), *Street gangs, migration and ethnicity.* Devon, UK: Willan Publishing.

Alleyne, E., Fernandes, I., & Pritchard, E. (2014). Denying humanness to victims: How gang members justify violent behavior. *Group Processes and Intergroup Relations, 17*, 750–762. doi:10.1177/1368430214536064

Alleyne, E., & Wood, J. L. (2010). Gang involvement: Psychological and behavioral characteristics of gang members, peripheral youth, and nongang youth. *Aggressive Behavior, 36*, 423–436. doi:10.1002/ab.20360

American Civil Liberties Union (ACLU) of Southern California. (1997). *False premises, false promises: The Blythe Street gang injunction and its aftermath*. Los Angeles: ACLU Foundation of Southern California.

Anwar, S., & Loughran, T. A. (2011). Testing a Bayesian learning theory of deterrence among serious juvenile offenders. *Criminology, 49*, 667–698. doi:10.1111/j.1745-9125.2011.00233.x

Apel, R. (2013). Sanctions, perceptions, and crime: Implications for criminal deterrence. *Journal of Quantitative Criminology, 29*, 67–101. doi:10.1007/s10940-012-9170-1

Ayling, J. (2011). Gang change and evolutionary theory. *Crime, Law and Social Change, 56*, 1–26. doi:10.1007/s10611-011-9301-x

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology, 71*, 364–374. doi:10.1037/0022-3514.71.2.364

Baron, R. S., & Kerr, N. (2003). *Group process, group decision, group action* (2nd ed.). Buckingham, UK: Open University Press.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy, 76*, 169–217. doi:10.1086/259394

Braga, A. A. (2005). Hot spots policing and crime prevention: A systematic review of randomized controlled trials. *Journal of Experimental Criminology*, *1*, 317–342. doi:10.1007/s11292-005-8133-z

Braga, A. A., Kennedy, D., Waring, E., & Piehl, A. (2001). Problem-oriented policing, deterrence, and youth violence: An evaluation of Boston's Operation Ceasefire. *Journal of Research in Crime and Delinquency, 38*, 195–225. doi:10.1177/0022427801038003001

Braga, A. A., Piehl, A. M., & Kennedy, D. M. (1999a). Youth homicide in Boston: An assessment of supplementary homicide report data. *Homicide Studies*, *3*, 277-299. doi:10.1177/1088767999003004002

Braga, A. A., & Weisburd, D. L. (2012). The effects of "pulling levers" focused deterrence strategies on crime. *Campbell Systematic Reviews, 6*, 1–90. doi:10.4073/csr.2012.6

Braga, A. A., Weisburd, D. L., Waring, E. J., Mazerolle, L. G., Spelman, W., & Gajewski, F. (1999b). Problem-oriented policing in violent crime places: A randomized controlled experiment. *Criminology*, *37*, 541-580. doi:10.1111/j.1745-9125.1999.tb00496.x

Caldwell, B. (2010). Criminalizing day-to-day life: A socio-legal critique of gang injunctions. *American Journal of Criminal Law, 37*(3), 241–290.

Carroll, J. S. (1978). A psychological approach to deterrence: The evaluation of crime opportunities. *Journal of Personality and Social Psychology, 36*, 1512–1520. doi:10.1037/0022-3514.36.12.1512

Centre for Social Justice. (2009). *Dying to belong: An in-depth review of street gangs in Britain*. London: Centre for Social Justice.

Centre for Social Justice. (2012). *Time to wake up: Tackling gangs one year after the riots*. London: Centre for Social Justice.

Chicago Municipal Code § 8-4-015 (1992).

City of Chicago v. Morales, 119 S. Ct. 1849, (1999).

Clarke, R. V., & Cornish, D. B. (1985). Modeling offenders' decisions: A framework for research and policy. *Crime and Justice, 6*, 147–185. doi:10.1086/449106

Cooper, J., Kelly, K. A., & Weaver, K. (2004). Attitudes, norms, and social groups. In M. B. Brewer & M. Hewstone (Eds.), *Social cognition* (pp. 244–267). Oxford, UK: Blackwell.

Cooper, A., & Ward, C. L. (2008). *Prevention, disengagement and suppression: A systematic review of the literature on strategies for addressing young people's involvement in gangs*. Cape Town, South Africa: Resources Aimed at Preventing Child Abuse and Neglect (RAPCAN).

Crawford, L. (2009). No way out: An analysis of exit processes for gang injunctions. *California Law Review*, 161–193.

Crewe, B., Liebling, A., Padfield, N., & Virgo, G. (2015). Joint enterprise: the implications of an unfair and unclear law. *Criminal Law Review, 4*, 252–269.
Criminal Street Gangs. 18 U.S.C. § 521 (2010).
Crocker, J., Luhtanen, R., Blaine, B., & Broadnax, S. (1994). Collective self-esteem and psychological wellbeing among White, Black, and Asian college students. *Personality and Social Psychology Bulletin, 20*, 503–513. doi:10.1177/0146167294205007
Dahmann, J. S. (1983). *Final report evaluation of operation hardcore: A prosecutorial response to violent gang criminality*. Washington, DC: National Institute of Justice, US Department of Justice.
Decker, S. H., & Van Winkle, B. (1996). *Life in the gang: Family, friends, and violence*. Cambridge: Cambridge University Press.
Densley, J. A. (2013). *How gangs work: An ethnography of youth violence*. Oxford, UK: Palgrave Macmillan.
Densley, J. A., Cai, T., & Hilal, S. (2014). Social dominance orientation and trust propensity in street gangs. *Group Processes & Intergroup Relations, 17*(6), 763–779. doi:10.1177/1368430214533161
Drury, J. (2010). Identity dynamics in adolescent-adult communication. In H. Giles, S. A. Reid, & J. Harwood (Eds.), *The dynamics of intergroup communication* (pp. 53–64). New York, NY: Peter Lang.
Durlauf, S. N., & Nagin, D. S. (2011). Imprisonment and crime. *Criminology and Public Policy, 10*, 13–54. doi:10.1111/j.1745-9133.2010.00680
Eady, D., (2013). *Perceptions of people maintaining unjust conviction under joint enterprise law* (Unpublished report compiled for JENGbA). Copy in possession of author.
Emler, N., & Reicher, S. (1995). *Adolescence and delinquency*. Oxford, UK: Blackwell Publishers Ltd.
Esbensen, F.-A., & Deschenes, E. P. (1998). A multisite examination of youth gang membership: Does gender matter? *Criminology, 36*, 799–828. doi:10.1111/j.1745-9125.1998.tb01266.x
Evans, C. R., & Dion, K. L. (1991). Group cohesion and performance: A meta-analysis. *Small Group Research, 22*, 175–186. doi:10.1177/1046496491222002
Fagan, J. (2002). Policing guns and youth violence. *The Future of Children, 12*(2), 133–151.
Fagan, J., Zimring, F. E., & Kim, J. (1998). Declining homicide in New York City: A tale of two trends. *Journal of Criminal Law and Criminology, 88*(4), 1277–1324.
Festinger, L. (1962). *A theory of cognitive dissonance* (Vol. 2). Stanford, CA: Stanford University Press.
Fisher, H., Gardner, F., & Montgomery, F. (2008). Cognitive-behavioural interventions for preventing youth gang involvement for children and young people (7–16). *Cochrane Database of Systematic Reviews*, (2). doi:10.1002/14651858.CD007008.pub2
Fiske, S. T., & Taylor, S. E. (1991). *Social cognition: From brains to culture* (2nd ed.). London, UK: Sage Publications Ltd.
Gatti, E., Tremblay, R. E., Vitaro, F., & McDuff, P. (2005). Youth gangs, delinquency and drug use: a test of the selection, facilitation, and enhancement hypotheses. *Journal of Child Psychology and Psychiatry, 46*, 1178–1190. doi:10.1111/j.1469-7610.2005.00423.x
Goldstein, A. P. (2002). *The psychology of group aggression*. West Sussex, UK: Wiley.
Greene, J., & Pranis, K. (2007). *Gang wars: The failure of enforcement tactics and the need for effective public safety strategies*. Washington, DC: Justice Policy Institute.
Grogger, J. (2002). The effects of civil gang injunctions on reported violent crime: Evidence from Los Angeles County. *Journal of Law and Economics, 45*(1), 69–90.
Grogger, J. (2005). What we know about gang injunctions. *Criminology and Public Policy, 4*, 637–641. doi:10.1111/j.1745-9133.2005.00307.x
Hagedorn, J. M. (2008). *A world of gangs: Armed young men and gangsta culture*. Minneapolis, MN: University of Minnesota Press.
Hale, D. (2006, October 20). Gang injunction: 300 black men targeted [Article from the Bay View]. Retrieved from http://www.streetgangs.com/injunctions/102006_300_black_men_targeted#sthash.gAwCw3vb.dpbs

Harcourt, B. E. (1998). Reflecting on the subject: A critique of the social influence conception of deterrence, the Broken Windows Theory, and Order-Maintenance Policing New York style. *Michigan Law Review, 97*, 291–389. doi:10.2307/1290289

Harding, S. (2012). *Unleashed: The phenomena of status dogs and weapon dogs*. Bristol, UK: The Policy Press.

Harris, D., Turner, R., Garratt, I., & Atkinson, S. Ministry of Justice, National Offender Management Service. (2011). *Understanding the psychology of gang delinquency: Implications for defining effecting violence interventions*. Retrieved from Ministry of Justice website http://www.justice.gov.uk/downloads/publications/research-and-analysis/moj-research/research-gang-violence.pdf

Hennigan, K. M., & Sloane, D. (2013). Improving Civil Gang Injunctions: How implementation can affect gang dynamics, crime, and violence. *Criminology and Public Policy, 12*, 7–41. doi:10.1111/1745-9133.12000

Hennigan, K., & Spanovic, M. (2012). Gang dynamics through the lens of social identity theory. In F. A. Esbensen & C. L. Maxson (Eds.), *Youth gangs in international perspective* (pp. 127–149). New York, NY: Springer.

Hill, G. H., Lui, C., & Hawkins, J. D. (2001). *Early precursors of gang membership: A study of Seattle youth*. Juvenile Justice Bulletin, Washington, DC: U.S. Department of Justice, Office of Justice Program, OJJDP.

Hogg, M. A. (2000). Subjective uncertainty reduction through self-categorization: A motivational theory of social identity processes. *European Review of Social Psychology, 11*, 223–255. doi:10.1080/14792772043000040

Home Office. (2014). *Review of the operation of injunctions to prevent gang-related violence.* London, UK: Home Office.

House of Commons Justice Committee. (2014). *Joint enterprise: Follow up*. London, UK: The Stationery Office Limited.

Howell, J. C. (2000). Youth Gang Programs and Strategies. OJJDP Summary. US—Department Of Education*: Office of Educational Research and improvement Educational Resources Information Center (Eric).*

Huff, C. R., & McBride, W. D. (1993). Gangs and the police. In A. P. Goldstein & C. R. Huff (Eds.), *The gang intervention handbook* (pp. 401–416). Champaign, IL: Research Press.

Hughes, L. A. (2013). Group cohesiveness, gang member prestige, and delinquency and violence in Chicago, 1959–1962. *Criminology, 51*, 795–832. doi:10.1111/1745-9125.12020

Hughes, L. A., & Short, J. F, Jr. (2005). Disputes involving youth street gang members: Micro-social contexts. *Criminology, 43*, 43–76. doi:10.111/j.0011-1348.2005.00002.x

Hyde, L. W., Shaw, D. S., & Moilanen, K. L. (2010). Developmental precursors of moral disengagement and the role of moral disengagement in the development of antisocial behavior. *Journal of Abnormal Child Psychology, 38*, 197–209. doi:10.1007/s10802-009-9358-5

Jankowski, M. S. (1991). *Islands in the street: Gangs and American urban society*. California: Univ of California Press.

Jones, R. A. (1997, August 10). *The puzzle waiting for the new chief* [Article from the Los Angeles Times]. Retrieved from http://articles.latimes.com/1997/aug/10/local/me-21226

Kahan, D. M. (1997). Between economics and sociology: The new path of deterrence. *Michigan Law Review, 95*(8), 2477–2497.

Katz, C. M., & Webb, V. J. (2003). *Police response to gangs: A multi-site study*. Phoenix, Arizona: National Institute of Justice.

Kennedy, D. M., Piehl, A. M., & Braga, A. A. (1996). Youth violence in boston: Gun markets, serious youth offenders, and a use-reduction strategy. *Law and Contemporary Problems, 59*, 147–196. doi:10.2307/1192213

Kleck, G., Sever, B., Spencer, L., & Gertz, M. (2005). The missing link in general deterrence research. *Criminology, 43*, 623–660. doi:10.1111/j.0011-1348.2005.00019

Klein, M. W. (1995). *The American street gang: Its nature, prevalence, and control*. New York, NY: Oxford University Press.

Klein, M. W. (2011). Who can you believe? Complexities of international street gang research. *International Criminal Justice Review, 21*, 197–207. doi:10.1177/1057567711414707

Klein, M. W., & Maxson, C. L. (2006). *Street gang patterns and policies*. New York, NY: Oxford University Press Inc.

Klepper, S., & Nagin, D. (1989). The deterrent effect of perceived certainty and severity of punishment revisited. *Criminology, 27*, 721–746. doi:10.1111/j.1745-9125.1989.tb01052.x

Lafontaine, T., Ferguson, M., & Wormith, J. S. (2005). *Street gangs: A review of the empirical literature on community and corrections-based prevention, intervention and suppression strategies*. Regina, Canada: Government of Saskatchewan.

Lien, I.-L. (2005). The role of crime acts in constituting the gang's mentality. In S. H. Decker & F. M. Weerman (Eds.), *European street gangs and troublesome youth groups: Findings from the Eurogang research program* (pp. 35–62). Oxford, UK: AltaMira Press.

Los Angeles County Civil Grand Jury. (2004). *Management audit of the civil gang injunctions*. In Los Angeles County Civil Grand Jury, 2003–2004 Final Report (pp. 169–335). Available online http://grandjury.co.la.ca.us/gjury03-04/LACGJFR_03-04.pdf

Los Angeles County District Attorney's Office. (2011). Retrieved from http://nixle.com/alert/4680580/

Marshall, B., Webb, B., & Tilley, N. (2005). *Rationalisation of current research on guns, gangs and other weapons: Phase 1*. London, UK: Jill Dando Institute of Crime Science.

Matsueda, R. L. (1992). Reflected appraisals, parental labelling, and delinquency: Specifying a symbolic interactionist theory. *American Journal of Sociology, 97*, 1577–1611. doi:10.1086/229940

Matsueda, R. L., Kreager, D. A., & Huizinga, D. (2006). Deterring delinquents: A rational choice model of theft and violence. *American Sociological Review, 71*, 95–122. doi:10.1177/000312240607100105

Matza, D. (1964). *Delinquency and drift*. New York, NY: Wiley.

Maxson, C. L., Hennigan, K. M., & Sloane, D. C. (2005). "It's getting crazy out there": Can a civil gang injunction change a community? *Criminology and Public Policy, 4*, 577–605. doi:10.1111/j.1745-9133.2005.00305.x

Maxson, C. L., Matsuda, K. N., & Hennigan, K. (2011). "Deterrability" among gang and nongang juvenile offenders: Are gang members more (or less) deterrable than other juvenile offenders? *Crime and Delinquency, 57*, 516–543. doi:10.1177/0011128709343137

McAra, L., & McVie, S. (2005). The usual suspects? Street-life, young people and the police. *Criminology and Criminal Justice, 5*, 5–36. doi:10.1177/1466802505050977

McClenaghan, M., McFadyean, M., & Stevenson, R. (2014). *Joint enterprise: An investigation into the legal doctrine of joint enterprise in criminal convictions*. London, UK: The Bureau of Investigative Journalism.

McGarrell, E. F., Chermak, S., Wilson, J. M., & Corsaro, N. (2006). Reducing homicide through a "lever-pulling" strategy. *Justice Quarterly, 23*, 214–231. doi:10.1080/07418820600688818

Metropolitan Police Authority. (2008). *Seen and heard—young people, policing and crime: An MPA report*. London, UK: Metropolitan Police Authority.

Moore, J., & Vigil, J. D. (1987). Chicano gangs: Group norms and individual factors related to adult criminality. *Aztlan, 18*(2), 27–44.

Muniz, A. (2014). Maintaining racial boundaries: Criminalization, neighborhood context, and the origins of gang injunctions. *Social Problems, 61*, 216–236. doi:10.1525/sp.2014.12095

Murphy, C. (2008). *The Chicago alternative policing strategy: A program analysis.* Retrieved from http://murcoil.com/pdf/CAPS.pdf

Myers, T. A. (2009). The unconstitutionality, ineffectiveness, and alternatives of gang injunctions. *Michigan Journal of Race and Law*, *14*, 285–306. Retrieved from http://heinonline.org/HOL/LandingPage?handle=hein.journals/mjrl14&div=12&id=&page=

National Gang Intelligence Center. (2011). *2011 National gang threat assessment: Emerging Trends*. Washington, DC: National Gang Intelligence Center.

National Gang Intelligence Center. (2013). *2013 National gang report*. Washington, DC: National Gang Intelligence Center.

O'Deane, M. (2011). *Gang injunctions and abatement: Using civil remedies to curb gang related crimes*. Boca Raton, FL: CRC Press.

O'Gorman, H. J. (1986). The discovery of pluralistic ignorance: An ironic lesson. *Journal of the History of the Behavioral Sciences, 22*, 333–347. doi:10.1002/1520-6696(198610)22:4<333:AID-JHBS2300220405>3.0.CO;2-X

Office of Juvenile Justice and Delinquency Prevention. (1994). *Gang suppression and intervention: Problem and response*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Office of the City Attorney of Los Angeles. (2013). *Important facts to know about gang injunctions*. Los Angeles, CA: Office of the City Attorney of Los Angeles.

Papachristos, A. V. (2009). Murder by structure: The social structure of gang homicide. *American Journal of Sociology, 115*, 74–128. doi:10.2139/ssrn.855304

Papachristos, A. V., Hureau, D. M., & Braga, A. A. (2013). Conflict and the corner: The impact of intergroup conflict and geographic turf on gang violence. *Social Science Research Network,*. doi:10.2139/ssrn.1722329

Parker, J., & Grasmick, H. G. (1979). Linking actual and perceived certainty of punishment: An exploratory study of an untested proposition in deterrence theory. *Criminology, 17*, 366–379. doi:10.1111/j.1745-9125.1979.tb01302.x

Paternoster, R., Brame, R., Bachman, R., & Sherman, L. W. (1997). Do fair procedures matter? The effect of procedural justice on spouse assault. *Law and Society Review, 31*(1), 163–204. doi:10.2307/3054098

Patton, D. U., Eschmann, R. D., & Butler, D. A. (2013). Internet banging: New trends in social media, gang violence, masculinity and hip hop. *Computers in Human Behavior, 29*(5), A54–A59. doi:10.1016/j.chb.2012.12.035

Pogarsky, G., & Piquero, A. R. (2003). Can punishment encourage offending? Investigating the "resetting" effect. *Journal of Research in Crime and Delinquency, 40*, 95–120. doi:10.1177/0022427802239255

Pratt, T. C., Cullen, F. T., Blevins, K. R., Daigle, L. E., & Madensen, T. D. (2006). The empirical status of deterrence theory: A meta-analysis. In F. Cullen, J. Wright, & K. Blevins (Eds.), *Taking stock: The status of criminological theory* (pp. 367–396). New Brunswick, New Jersey: Transaction Publishers.

Pyrooz, D. C., Decker, S. H., & Webb, V. J. (2014). The ties that bind: Desistance from gangs. *Crime and Delinquency, 60*, 491–516. doi:10.1177/0011128710372191

Ralphs, R., Medina, J., & Aldridge, J. (2009). Who needs enemies with friends like these? The importance of place for young people living in known gang areas. *Journal of Youth Studies, 12*, 483–500. doi:10.1080/13676260903083356

Reid, S. A., Cropley, C. J., & Hogg, M. A. (2005a). *A self-categorization explanation of pluralistic ignorance* (Unpublished manuscript). University of California, Santa Barbara.

Reid, S. A., Giles, H., & Harwood, J. (2005b). A self-categorization perspective on communication and intergroup relations. In J. Harwood & H. Giles (Eds.), *Intergroup communication: Multiple perspectives* (pp. 241–263). New York, NY: Peter Lang.

Rimal, R. N., & Real, K. (2003). Understanding the influence of perceived norms on behaviors. *Communication Theory, 13*, 184–203. doi:10.1111/j.1468-2885.2003.tb00288.x

Sanders, W. B. (1994). *Gangbangs and drive-bys: Grounded culture and juvenile gang violence*. New York, NY: Aldine de Gruyter.

Sharp, C., Aldridge, J., & Medina, J. (2006). *Delinquent youth groups and offending behaviour: findings from the 2004 Offending, Crime and Justice Survey* (Home Office).

Short, J. F, Jr., & Strodtbeck, F. L. (1965). *Group process and gang delinquency*. Chicago, IL: University of Chicago.

Shute, J., & Medina, J. (2014). Hunting gruffalo: 'Gangs', unreason and the big bad coalition. *Criminal Justice Matters, 96*, 26–27. doi:10.1080/09627251.2014.926070

Sidanius, J., & Pratto, F. (1999). *Social dominance*. Cambridge, UK: Cambridge University Press.

Siegel, L. (2003). Gangs and the law. In L. Kontos, D. Brotherton, & L. Barrios (Eds.), *Gangs and society: Alternative perspectives* (pp. 213–227). New York: Columbia University Press.

Skogan, W., & Frydl, K. (2004). *Fairness and effectiveness in policing: The evidence*. Washington, DC: The National Academies Press.

South, R., & Wood, J. L. (2006). Bullying in prisons: The importance of perceived social status, prisonization and moral disengagement. *Aggressive Behavior, 32*, 490–501. doi:10.1002/ab.20149

Stafford, M. C., & Warr, M. (1993). A reconceptualization of general and specific deterrence. *Journal of Research in Crime and Delinquency, 30*, 123–135. doi:10.1177/0022427893030002001

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Chicago, IL: Nelson-Hall.

The Guardian. (2014). *MPs call for urgent review of joint-enterprise law* [News article]. Retrieved from http://www.theguardian.com/law/2014/dec/17/joint-enterprise-law-mps-call-review

Thornberry, T. P., & Krohn, M. D. (1997). Peers, drug use, and delinquency. In D. M. Stoff, J. Breiling, & J. D. Maser (Eds.), *Handbook of antisocial behavior* (pp. 218–233). New York, NY: Wiley.

Thornberry, T. P., Krohn, M. D., Lizotte, A. J., Smith, C., & Tobin, K. (2003). *Gangs and delinquency in developmental perspective*. Cambridge: Cambridge University Press.

Turner, J. C., Hogg, M., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford, UK: Blackwell.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131. doi:10.1126/science.185.4157.1124

Vargas, R. (2014). Criminal group embeddedness and the adverse effects of arresting a gang's leader: a comparative case study. *Criminology, 52*(2), 143–168. doi:10.1111/1745-9125.12033

Vigil, J. D. (1988). *Barrio gangs: street life and identity in southern California*. Austin, TX: University of Texas Press.

Viki, G. T., & Abrams, D. (2012). The social influence of groups on individuals. In J. L. Wood & T. A. Gannon (Eds.), *Crime and crime reduction: The importance of group processes* (pp. 3–33). East Sussex, UK: Routledge.

Waldeck, S. (2000). Cops, community policing, and the social norms approach to crime control: Should one make us more comfortable with the others? *Georgia Law Review*, *34*, 1253–1297. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1024750

Watkins, A. M., Huebner, B. M., & Decker, S. H. (2008). Patterns of gun acquisition, carrying, and use among juvenile and adult arrestees: Evidence from a high-crime city. *Justice Quarterly, 25*, 674–700. doi:10.1080/07418820801930118

Weerman, F. M., Maxson, C. L., Esbensen, F. A., Aldridge, J., Medina, J., & van Gemert, F. (2009). *Eurogang program manual: Background, development, and use of the Eurogang instruments in multi-site, multi-method comparative research*. St Louis, MO: University of Missouri at St Louis.

Weisburd, D., Einat, T., & Kowalski, M. (2008). The miracle of the cells: An experimental study of interventions to increase payment of court-ordered financial obligations. *Criminology and Public Policy, 7*, 9–36. doi:10.1111/j.1745-9133.2008.00487.x

Wellford, C. F., Pepper, J. V., & Petrie, C. V. (2005). *Firearms and violence: A critical review*. Washington, DC: National Academies Press.

Wilson, J. M., & Chermak, S. (2011). Community-driven violence reduction programs. *Criminology and Public Policy, 10*, 993–1027. doi:10.1111/j.1745-9133.2011.00763.x

Wilson, J. Q., & Kelling, G. L. (1982). Broken windows. *Atlantic Monthly*, *249*, 29–38. Retrieved from http://www.lantm.lth.se/fileadmin/fastighetsvetenskap/utbildning/Fastighetsvaerderingssystem/BrokenWindowTheory.pdf

Wood, J. L. (2014). Understanding gang membership: The significance of group processes. *Group Processes and Intergroup Relations, 17*, 710–729. doi:10.1177/1368430214550344

Wood, J. L., & Alleyne, E. (2010). Street gang theory and research: Where are we now and where do we go from here? *Aggression and violent behavior*, *15*, 100–111. doi:10.1016/j.avb.2009.08.005

Wood, J. L., Alleyne, E., Mozova, K., & James, M. (2014). Predicting involvement in prison gang activity: Street gang membership, social and psychological factors. *Law and Human Behavior, 38*(3), 203–211. doi:10.1037/lhb0000053

Wood, J. L., Moir, A., & James, M. (2009). Prisoners' gang-related activity: The importance of bullying and moral disengagement. *Psychology, Crime and Law, 15*, 569–581. doi:10.1080/10683160802427786

Wood, J. L., Williams, G. R., & James, M. J. (2010). Incapacitation and imprisonment: Prisoners' involvement in community-based crime. *Psychology, Crime, and Law, 16*, 601–615. doi:10.1080/10683160902971071

Zimring, F. E., & Hawkins, G. (1973). *Deterrence: The legal threat in crime control*. Chicago: University of Chicago Press.

Index

B.H. Bornstein and M.K. Miller (eds.), *Advances in Psychology and Law*,
Advances in Psychology and Law 2, DOI 10.1007/978-3-319-43083-6