

MORE PROBLEMS WITH CRIMINAL TRIALS: THE LIMITED EFFECTIVENESS OF LEGAL MECHANISMS

DAN SIMON*

I INTRODUCTION

Criminal adjudication is a core feature in American public life. Trials are considered “the central institution of law as we know it,”¹ the “crown jewel” of the legal system.² Amidst its multiple purposes, an essential objective of the criminal trial is to determine facts pertaining to the defendant’s guilt. Specifically, the trial is designed to serve the diagnostic function of distinguishing between factually guilty and innocent people. The prevailing sentiment within the American polity and legal profession is that the trial is indeed acutely diagnostic.³ This traditional truism, however, is being challenged by the mounting revelations of false convictions.⁴

Copyright © 2012 by Dan Simon.

This article is also available at <http://lcp.law.duke.edu/>.

* Professor of Law and Psychology at the University of Southern California. For helpful comments on early drafts, I thank the participants at the faculty workshops at Harvard Law School, Northwestern Law School, The University of Chicago Law School, USC Gould School of Law, and Yale Law School. I am particularly grateful for contributions by Bruce Ackerman, Akhil Amar, Jack Balkin, Rebecca Brown, Steve Clark, Mirjan Damaška, Shari Diamond, John Donohue, Brandon Garrett, Heather Gerken, Phil Heymann, Sam Gross, Dan Kahan, Dan Klerman, John Langbein, Tom Lyon, Rob MacCoun, Richard McAdams, Daniel Markovits, Tracey Meares, Janice Nadler, Charles Ogletree, Robert Post, Carol Steiker, Kate Stith, Alan Stone, Lloyd Weinreb, and Jim Whitman. For their wonderful assistance, I thank the staff of the library at USC Law School. My deepest gratitude goes to Anne Simon.

1. ROBERT P. BURNS, *THE DEATH OF THE AMERICAN TRIAL* 3 (2009); JAMES BOYD WHITE, *FROM EXPECTATION TO EXPERIENCE: ESSAYS ON LAW AND LEGAL EDUCATION* 108 (1999).

2. Randall T. Shepard, *The New Role of State Supreme Courts as Engines of Court Reform*, 81 N.Y.U. L. REV. 1535, 1543 (2006).

3. See JEFFREY ABRAMSON, *WE THE JURY: THE JURY SYSTEM AND THE IDEAL OF DEMOCRACY* 5 (1994) (noting that although “juries [do not] always get their verdicts right . . . [t]o get at the good, we must risk the bad”); ROBERT P. BURNS, *A THEORY OF THE TRIAL* 153 (1999) (noting that “[t]he work of juries is generally highly regarded by those in a position to know”); NEIL VIDMAR & VALERIE P. HANS, *AMERICAN JURIES: THE VERDICT* 346 (2007) (“After evaluating all of the evidence, our verdict is strongly in favor of the American jury.”).

For early exceptions to the mainstream view, see EDWIN M. BORCHARD, *CONVICTING THE INNOCENT* xiii (1932) (noting that the errors that lead to false convictions are “typical”); JEROME FRANK & BARBARA FRANK, *NOT GUILTY* 31 (1957) (“The conviction and imprisonment of innocent men too frequently occur to be ignored by any of us. There are too many cases on record to prove the point, and there may be countless others of which we know nothing.”).

4. BRANDON GARRETT, *CONVICTING THE INNOCENT: WHERE CRIMINAL PROSECUTIONS GO*

This article follows in the heels of an article entitled *The Limited Diagnosticity of Criminal Trials (Limited Diagnosticity)*.⁵ Both articles pertain to parts of a larger critique of the criminal justice process that will be published in a book manuscript *In Doubt: The Psychology of the Criminal Justice Process (In Doubt)*.⁶ *Limited Diagnosticity* examined how good fact finders are in determining facts from the types of evidence presented in criminal trials. In other words, how well do juries and judges perform the diagnostic function of distinguishing between factual guilt and innocence? To address this question, that article applied a body of experimental psychological research pertaining to the performance of the tasks involved in the adjudicatory process.⁷

As suggested by its title, *Limited Diagnosticity* concluded that the cognitive processing involved in discovering the truth in difficult cases is more complex and less reliable than generally believed.⁸ One problem with the diagnosticity of criminal adjudication stems from the fact that the evidence that is typically presented in criminal trials is of uneven, and often low, quality. The problems with the integrity of the evidence begin with the fact that criminal investigations are not conducted through best-practice procedures. Courtroom testimony is usually proffered months, sometimes years, following the criminal event, a period during which witnesses repeatedly interact with the criminal process and are subjected to a variety of contaminating sources. As a result, the “raw evidence” perceived by the witness at the criminal event often undergoes editing, embellishment, and alteration. Determining the facts accurately from this “synthesized evidence” is a daunting task, especially given the opacity of the investigative process.

Second, the research examined in *Limited Diagnosticity* indicated that even if the evidence was reliable, drawing the correct inferences from the most common types of evidence presented in criminal trials is a difficult feat. People have a difficult time judging whether a witness identified the perpetrator correctly, whether a witness’s memories of an event are accurate, and whether confessions obtained in police interrogations offer truthful accounts of the suspect’s deeds. The research also indicated that alibi testimony can be hard to produce and is often misleading, and that detecting deceit from people’s

WRONG (2011); Samuel R. Gross et al., *Exonerations in the United States 1989 Through 2003*, 95 J. CRIM. L. & CRIMINOLOGY 523, 523–24 (2005); 250 *Exonerated: Too Many Wrongfully Convicted*, THE INNOCENCE PROJECT 1 (2010), http://www.innocenceproject.org/docs/InnocenceProject_250.pdf.

5. Dan Simon, *The Limited Diagnosticity of Criminal Trials*, 64 VAND. L. REV. 143 (2011).

6. DAN SIMON, *IN DOUBT: THE PSYCHOLOGY OF THE CRIMINAL JUSTICE PROCESS* (forthcoming 2012) (on file with author).

7. The following discussion focuses almost exclusively on determinations of facts that are, at least in principle, discernable—notably, the identity of the perpetrator and the physical acts and circumstances of the criminal event. This article has little to say about value judgments that fact finders are called to make, such as the reasonableness of an act, the morality of a behavior, or the fairness of the law.

8. This project pertains primarily to the class of difficult cases, in which the evidence relies mostly on human testimony and is factually complex. These are the cases that consume most of the adjudicative resources and that put the diagnostic capabilities of the trial to the test.

demeanor is often mistaken.

Third, the fact finding task is further hindered by a variety of factors that are inherent to the environment of criminal trials. These include the excessive persuasive attempts by lawyers, exposure to extra-evidential information, the emotional arousal that accompanies many criminal trials, and racial stereotypes. Another hindrance stems from the fact finder's cognitive process itself, namely, the "coherence effect."⁹ Fact finders tend to overestimate their performance on the fact finding tasks and to underestimate their susceptibility to biasing factors. In sum, *Limited Diagnosticity* concluded that factual findings in criminal trials are bound to contain an appreciable level of inaccuracy and they are also vulnerable to manipulation. While faulty factual determinations can result in dropped charges against truly guilty people and even wrongful acquittals, they lead mostly to the prosecution of innocent people. The prospect of error is generally ignored or denied by those entrusted with governing the criminal justice system, and is not adequately recognized in the scholarly debate. It should be noted that these findings pertain to actors who honestly believe that they are fulfilling their roles properly. Deliberate dishonest conduct raises different sets of issues, which lie beyond the scope of this project.¹⁰

One possible objection to *Limited Diagnosticity* is that it was confined to human performance in decontextualized circumstances that fail to capture the potential for accuracy afforded by the legal procedure. Criminal fact finding might prove to be more diagnostic once embedded in the legal mechanisms that have been designed to correct for errors and to promote accuracy. This article is intended to explore that possibility. It examines the psychological research on the potential of a number of prominent components of the criminal trial to promote the accuracy of the fact finding task. These mechanisms include cross-examination, jury instructions, jurors' assurances of impartiality, the prosecution's heightened burdens, jury deliberation, and judicial review by appellate and post-conviction proceedings. The ubiquitous reverence of the American criminal trial is premised on the assumption that these mechanisms perform their designated functions successfully. The truth evincing potential of these mechanisms is the subject of this article, and they will be examined in turn. In all, the article paints a rather mixed and overall troubling picture. To a

9. See generally Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. CHI. L. REV. 511 (2004); Dan Simon & Kieth J. Holyoak, *Structural Dynamics of Cognition: From Consistency Theories to Constraint Satisfaction*, 6 PERSONALITY & SOC. PSYCHOL. REV. 283 (2002); DAN SIMON, *supra* note 6, at chs. 2, 6.

10. The article will focus mostly on the performance of lay people who serve the fact finding function in the large majority of criminal trials. Its purpose, however, is not to question the suitability of juries for the task, nor to compare them to judges. A mounting body of studies finds that judges do not perform much differently from lay people in many fact finding tasks. See, e.g., Chris Guthrie, Jeffrey J. Rachlinski & Andrew J. Wistrich, *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 826–27 (2001); Jeffrey J. Rachlinski, Chris Guthrie & Andrew J. Wistrich, *Inside the Bankruptcy Judge's Mind*, 86 B.U. L. REV. 1227, 1256–57 (2006). The limitations of human cognition observed in the research appear to exceed any possible differences between the two decisionmaking entities. The suitability of the jury as the preferred fact finding body is left for another day.

limited extent, these mechanisms do indeed enhance diagnosticity, but under a range of realistic circumstances, they turn out to be ineffective, and even detrimental to the process.

II

CROSS-EXAMINATION

The cross-examination of a witness in open court is heralded as a one of the hallmarks of the adversarial system.¹¹ In addition to its important symbolic and political dimensions, cross-examination is deemed a formidable tool for the discovery of truth. The method has long been believed to “beat and bolt out the Truth.”¹² Quoting John Wigmore, the Supreme Court has repeatedly stated that cross-examination is the “greatest legal engine ever invented for the discovery of truth.”¹³ Cross-examination is closely related to the principle of orality, a central feature of the Anglo-American criminal trial by which testimony is given live, in open court.¹⁴ The cross-examination of witnesses could conceivably improve the diagnosticity of the process by either deterring or exposing false testimony.

First, cross-examination could have a prophylactic effect on would-be witnesses. The expectation of having to face tough questioning in open court could indeed have a harnessing effect on witnesses throughout the investigative process and trial in that the fear of exposure could deter witnesses from providing dishonest statements.

Still, the limits of this effect ought not be overlooked. For one, the deterrent effect might restrain would-be liars, but it is doubtful whether it can make much impact on the more challenging cases that involve honest, yet mistaken, witnesses. Memory research indicates that people tend to trust their memories, regardless of the accuracy of those memorial accounts.¹⁵ Given that mistaken witnesses perceive themselves to be accurate, they are unlikely to be deterred from recounting their (actually false) memories any more than accurate witnesses would be deterred from recounting their (truly correct) ones.

The deterrent effect is likely to be diminished also when the witness has little to fear from being exposed on the stand. Witnesses have little to fear when

11. The right to cross-examination is one of the primary justifications for the right to confront one's witness and the right to assistance of counsel, both guaranteed by the Sixth Amendment, which reads in part, “In all criminal prosecutions, the accused shall enjoy the right to . . . be confronted with the witnesses against him . . . and to have the Assistance of Counsel for his defence.” U.S. CONST. amend. VI.

12. JOHN H. LANGBEIN, *THE ORIGINS OF ADVERSARY CRIMINAL TRIAL* 234 (A. W. Brian Simpson ed., 2003) (quoting Matthew Hale).

13. *See, e.g., Lilly v. Virginia*, 527 U.S. 116, 124 (1999); *Watkins v. Sowders*, 449 U.S. 341, 349 (1980).

14. *See Washington v. Crawford*, 541 U.S. 36, 43 (2004); Honoré, T., *The Primacy of Oral Evidence? in CRIME, PROOF AND PUNISHMENT: ESSAYS IN MEMORY OF SIR RUPERT CROSS* 172, 172–92 (Butterworths ed., 1981).

15. *See SIMON, supra* note 6, at ch. 4.

there is no reliable evidence that contradicts their statements. This occurs, for example, when the witness is the sole source of evidence, such as when she was the victim of a crime committed in private, or when she happened to be the only person to witness the criminal event. Due to opacity of the investigations, the limited scope of the prosecution's duty to disclose exculpatory evidence, and the credibility advantage that police officers enjoy over criminal defendants, police investigators will usually have little reason to be concerned about being confronted with contradictory evidence.¹⁶ There will be no reliable evidence to contradict a witness's statement also when that testimony is corroborated by the rest of the available evidence, although the apparent corroboration might be a misleading artifact of the investigative process. Due to the escalating nature of investigations, the accumulation of evidence tends to be driven by the previously available evidence.¹⁷ Paradoxically, it is possible that the corroborating evidence itself was induced by an investigation that was thrown off course by the erroneous testimony now being corroborated.

Witnesses also have little to fear when they know that the jury is strongly disposed to believe them. This is often the case when the witness is a sympathetic victim or, in most jurisdictions, when she is a police officer or public official. Witnesses also have little reason to fear being exposed on the stand when they expect to be discredited no matter what. This is the case for witnesses who perceive that the jury is predisposed to disbelieve them, and it is also generally true for expert witnesses who testify for hire. These experts tend to adopt the position of the party that hires them, and fully expect to be hammered by the opposing side.¹⁸

The second manner in which cross-examination could improve the diagnosticity of the trial is by exposing mistakes or lies in the course of the cross-examination itself. It is beyond dispute that such exposures (call them Perry Mason moments) enhance the accuracy of the evidence presented to the jury. Yet these dramatic instances are few and far between. Witnesses are habitually *prepared* by lawyers, and this usually means that their testimony has been massaged to look as reliable as possible. Witnesses are also advised by the lawyers on what to expect during cross-examination and how best to respond to it.¹⁹

The infrequency of the Perry Mason moments also stems from the strategic risks involved in seeking to expose witnesses on the stand. Attempting, but failing, to expose a witness as deceitful can be very costly. For one, repeated questioning of the witness gives her another opportunity to reiterate her statements. More importantly, weathering a cross-examination can make her

16. On the opacity of investigations, see SIMON, *supra* note 6, at ch. 6. On the limited prosecutorial obligations to disclose exculpatory evidence to the defense, see *id.* at ch. 8.

17. See *id.* at chs. 2, 6.

18. See Ralph Underwager & Hollida Wakefield, *Responding to Improper and Abusive Impeachment Efforts*, 14 AM. J. FORENSIC PSYCHOL. 5 (1996).

19. See SIMON, *supra* note 6, at ch. 6.

testimony seem even more reliable and thus bolster the effectiveness of her testimony. Hence the famous adage, “[Y]ou should never ask a question to which you do not already know the answer.”²⁰ The effectiveness of cross-examination is particularly limited for defense attorneys, who labor under an informational disadvantage. The vast majority of criminal defendants lack the resources, expertise, and legal authority to investigate crimes effectively. While the prosecution has a virtual monopoly on the evidence, its duties to share exculpatory evidence are rather limited.²¹ Vigorous cross-examination can also be perceived by the fact finder as unduly hostile and thus off-putting.²² Again, this concern is particularly constraining for defense attorneys as jurors tend to react negatively to the badgering of victims,²³ bystanders, and police officers.

Moreover, cross-examination can hamper the diagnosticity of the process by undermining the testimony of even honest and accurate witnesses.²⁴ Indeed, some commentators argue that assailing the reliability of all opposing witnesses is a professional duty that a lawyer owes her client in an adversarial system.²⁵ A recent study shows that subjecting honest witnesses to a realistic cross-examination results in substantial changes in their testimony with almost three-quarters of witnesses altering their responses on at least one of the four factual issues that were critical to the outcome of the case.²⁶ One technique for undercutting a witness is to phrase the questions in convoluted language, pejoratively dubbed “lawyerese.” These questions can come in the form of leading questions, questions phrased in the negative or double negative form, and multipart questions. Studies show that these questions increase the rate of erroneous and “don’t know” responses.²⁷ Questions phrased in lawyerese have

20. IRVING YOUNGER, *THE ART OF CROSS-EXAMINATION* 23 (Am. Bar Ass’n Section of Litig., The Section of Litig. Monograph Ser. No. 1, 1976).

21. On the case of *Brady v. Maryland*, 373 U.S. 83, 87 (1963), and its progeny, see Daniel S. Medwed, *Brady’s Bunch of Flaws*, 67 WASH. & LEE L. REV. 1533, 1533–67 (2010); Scott E. Sundby, *Fallen Superheroes and Constitutional Mirages: The Tale of Brady v. Maryland*, 33 MCGEORGE L. REV. 643, 643–63 (2002).

22. For findings of backfiring effects, see Derek D. Rucker & Richard E. Petty, *Effects of Accusations on the Accuser: The Moderating Role of Accuser Culpability*, 29 PERSONALITY & SOC. PSYCHOL. BULL. 1259 (2003).

23. A study found that an innuendo inserted by the defense attorney during cross-examination had a negative impact on jurors’ evaluations of the testimony of an expert witness, but not of the victim. Saul M. Kassir et al., *Dirty Tricks of Cross-Examination: The Influence of Conjectural Evidence on the Jury*, 14 LAW & HUM. BEHAV. 373, 380–81 (1990).

24. This danger was observed by Jerome Frank and Marvin E. Frankel. JEROME FRANK, *COURTS ON TRIAL* 82 (1949) and MARVIN E. FRANKEL, *PARTISAN JUSTICE* 75 (1978).

25. Monroe H. Freedman, *Professional Responsibility of the Criminal Defense Lawyer: The Three Hardest Questions*, 64 MICH. L. REV. 1469, 1475–76 (1966).

26. Tim Valentine & Katie Maras, *The Effect of Cross-Examination on the Accuracy of Adult Eyewitness Testimony*, 25 APPLIED COGNITIVE PSYCHOL. 554, 559 (2011).

27. Mark R. Keibell & Shane D. Johnson, *Lawyers’ Questioning: The Effect of Confusing Questions on Witness Confidence and Accuracy*, 24 LAW & HUM. BEHAV. 629, 634, 637 (2000); Nancy W. Perry et al., *When Lawyers Question Children: Is Justice Served?*, 19 LAW & HUM. BEHAV. 609, 625–26 (1995). The latter experiment found that convoluted questions had adverse effects of subjects in four age groups ranging from kindergarten to college.

also been found to reduce the confidence of witnesses, regardless of their accuracy,²⁸ thus further weakening the limited diagnostic value of witness's confidence.²⁹

Occasionally, the quest to discredit witnesses injects a degree of acrimony and even viciousness into the proceedings. One manual evokes images of hunting when advising how to conduct cross-examination: "Close all the gaps he might try to slither through," and "pin him down—don't spring the trap too soon."³⁰ The anticipation of being vigorously cross-examined by opposing counsel tends to entrench witnesses in their position and feel more strongly aligned with the party for whom they are testifying. Indeed, lawyers often prime their witnesses for a battle with the opposing attorney. One study showed that preparing a prosecution eyewitness for a hostile cross examination bolstered the incriminating strength of that witness's testimony and resulted in an increased rate of guilty verdicts.³¹

Cross-examination can also compromise the process by influencing the selection of witnesses on grounds other than the accuracy of their testimony. Lawyers might well elect not to bring to the stand an honest and reliable witness because she might "not hold up well" to the vigor of cross-examination, whether due to a personality trait, low intelligence, and the like. By the same token, a lawyer would be tempted to call a witness who is expected to stand up to the task, even if she suspects his integrity or the reliability of his account. Thus, while cross-examination should help deter dishonest witnesses and expose the truth on the stand, its effectiveness and availability are limited, and it can even have an adverse effect on the accuracy of the process.

III

JURY INSTRUCTIONS

A primary justification cited for entrusting jurors with the task of deciding criminal verdicts is that they infuse the decisionmaking process with common sense.³² At the same time, a fundamental tenet in democratic regimes is that the imposition of guilt must be done in accordance with the formal dictates of the law.³³ The primary means by which the legal process seeks to subordinate jury verdicts to the appropriate legal rules is by means of a formal set of instructions

28. Jacqueline M. Wheatcroft et al., *The Influence of Courtroom Questioning Style on Actual and Perceived Eyewitness Confidence and Accuracy*, 9 LEGAL & CRIMINOLOGICAL PSYCHOL. 83, 83 (2004).

29. On the accuracy–confidence relationship, see SIMON, *supra* note 6, at chs. 3, 4.

30. KEITH EVANS, *THE COMMON SENSE RULES OF TRIAL ADVOCACY* 135–36 (1994).

31. Gary L. Wells et al., *The Tractability of Eyewitness Confidence and Its Implications for Triers of Fact*, 66 J. APPLIED PSYCHOL. 688 (1981).

32. As the Supreme Court stated in *Taylor v. Louisiana*, juries embody "the commonsense judgment of the community" as a check against prosecutorial excesses and "in preference to the professional or perhaps over-conditioned or biased response of a judge." 419 U.S. 522, 530 (1975).

33. The principle *nullum crimen sine lege* states that conduct does not constitute a crime unless it has previously been declared to be so by the law.

conveyed to the jury by the trial judge. In addition to the definition of the criminal charge, jury instructions contain numerous rules designed to delineate the framework for deciding the verdict, including the primacy of trial evidence, the presumption of innocence, and the standard of proof. To reach legally warranted verdicts juries need to apply these rules in the prescribed manner. Sometimes, juries deliberately disregard the instructions, thus engaging in jury nullification.³⁴ The following discussion, however, is concerned instead with unwitting failures in applying instructions.³⁵

Driven by a fidelity to the law and by the fear of being overturned on appeal, trial judges typically provide jurors with instructions that adhere to legal rules, even when doing so entails issuing instructions that are beyond the grasp or capabilities of most people. Jury instructions are often complex, couched in alien terminology, and demanding of unfamiliar and even improbable mental exercises. Still, the legal system relies heavily on the assumption that juries conform to the instructions they receive, and the Supreme Court routinely proclaims its faith in their conformity. A crucial assumption underlying that system is that juries will follow the instructions given them by the trial judge.” To support this belief, the Court has offered the circular argument that if jurors did not follow instructions, “it would be pointless for a trial court to instruct a jury, and even more pointless for an appellate court to reverse a criminal conviction because the jury was improperly instructed.”³⁶ Underlying this faith is the belief that although jurors are endowed with a keen eye for discerning human behavior and a developed sense of morality, they are also remarkably malleable to the dictates of the law, even when it conflicts with their personal sense of justice. Just how well jurors comprehend and apply jury instructions is an empirical question. The following discussion tracks the two primary functions served by jury instructions: guiding jurors toward proper decisionmaking, and protecting decisions from potential prejudice.

A. Guidance Instructions

One way to gauge the effectiveness of instructions in guiding jury decisionmaking is to measure how well jurors comprehend them. A sizeable body of research indicates that jurors’ performance falls well short of law’s implicit assumption of complete or near-complete comprehension. The studies find comprehension rates ranging from thirteen to seventy-three percent, levels that are not always better than chance,³⁷ and are far inferior to the law’s implicit

34. See generally Nancy S. Marder, *The Myth of the Nullifying Jury*, 93 NW. U. L. REV. 877 (1999); VIDMAR & HANS, *supra* note 3.

35. It should be acknowledged that a failure to follow jury instructions means that the verdict is not likely to comport with the prescribed law, but that does not necessarily compromise the factual accuracy of the decision. In some instances—such as when the rules of evidence mandate keeping a piece of accurate evidence out of court—the failure to follow the instruction could actually increase the accuracy of the verdict.

36. *Parker v. Randolph*, 442 U.S. 62, 73 (1979).

37. For example, a large post-deliberation study of instruction comprehension found levels of

assumption of complete or near-complete comprehension. The data are mixed as to whether the instructions actually improve jurors' understanding of the law, with some studies finding modest improvement,³⁸ and others finding that instructed jurors are no more knowledgeable than their non-instructed cohorts.³⁹ A study conducted in England found that people tend not to acknowledge their limited levels of knowledge. While only thirty-one percent of respondents understood the instructions, two-thirds of them maintained that they comprehended them correctly. Notably, the level of comprehension increased to almost one-half when the instructions were given in writing.⁴⁰ The comprehension of instructions is particularly low when they pertain to issues about which jurors hold incorrect preexisting beliefs or which entail unfamiliar mental tasks.⁴¹

Similar questions arise when lay people are tested on their ability to apply jury instructions to legal situations. Studies with jurors in Washington state and Florida found that instructed and uninstructed participants performed similarly in applying instructions.⁴² Notably, studies have found that decisions are often unaffected by the instructions given. In deciding homicide cases, simulated jurors rendered the same verdicts under two discrepant definitions of the crime.⁴³ Similar ineffectiveness was found in experiments that compared discrepant instructions for the crime of rape,⁴⁴ and for the insanity defense.⁴⁵

accuracy around 30%. REED HASTIE ET AL., *INSIDE THE JURY* (1983). Comprehension levels on insanity defense instructions were found to range from 15% to 43% correct. James R. P. Ogloff, *A Comparison of Insanity Defense Standards on Juror Decision Making*, 15 *LAW & HUM. BEHAV.* 509, 512 (1991). For a review of the research, see generally Joel D. Lieberman, & Bruce D. Sales, *What Social Science Teaches Us About the Jury Instruction Process*, 3 *PSYCHOL. PUB. POL'Y & L.* 589 (1997).

38. See, e.g., Raymond W. Buchanan et al., *Legal Communication: An Investigation of Juror Comprehension of Pattern Instructions*, 26 *COMM. Q.* 31, 34 (1978); David U. Strawn & Raymond W. Buchanan, *Jury Confusion: A Threat to Justice*, 59 *JUDICATURE* 478, 482 (1976).

39. For example, a study of Michigan jurors found that instruction increased correct responses on procedural rules, but not on definitions of crimes. In this study, the level of comprehension was shy of fifty percent on both types of instructions. Alan Reifman et al., *Real Jurors' Understanding of the Law in Real Cases*, 16 *LAW & HUM. BEHAV.* 539, 547 (1992).

40. Cheryl Thomas, *Are Juries Fair?*, *MINISTRY OF JUSTICE RESEARCH SERIES*, vi (Feb. 2010), <http://www.justice.gov.uk/downloads/publications/research-and-analysis/moj-research/are-juries-fair-research.pdf>.

41. For example, Michigan jurors' responses barely reached chance level on instructions pertaining to the use of evidence for limited purposes and the definition of specific intent. Geoffrey P. Kramer & Doreen M. Koenig, *Do Jurors Understand Criminal Jury Instructions? Analyzing the Results of the Michigan Juror Comprehension Project*, 23 *U. MICH. J. L. REFORM* 401, 417 (1990).

42. In the Washington sample, instructed participants applied the instructions correctly sixty percent of the time. Laurence J. Severance & Elizabeth F. Loftus, *Improving the Ability of Jurors to Comprehend and Apply Criminal Jury Instructions*, 17 *LAW & SOC'Y REV.* 153, 181 (1982); Buchanan et al., *supra* note 38.

43. Nonetheless, respondents relied on the wording of the definitions to justify their decisions. Matthew P. Spackman et al., *An Analysis of the Effects of Subjective and Objective Instruction Forms on Mock-Juries' Murder/Manslaughter Distinctions*, 26 *LAW & HUM. BEHAV.* 605, 612-14 (2002).

44. A study by Dan Kahan found no difference in conviction rates among three different crime definitions (ranging from 53% to 55%). Minor differences were observed with two unconventional instructions (rates of 62% and 65%). Dan M. Kahan, *Culture, Cognition, and Consent: Who Perceives What, and Why, in Acquaintance-Rape Cases*, 158 *U. PA. L. REV.* 729, 773-83 (2010).

Perhaps the strongest impediment to the instruction of juries is the jurors' own preconceptions of law and justice.⁴⁶ People's decisions tend to conform more to their personal preconceptions than to the instructions they receive. Even after receiving instructions on the definitions of crimes such as robbery, burglary, and kidnapping, lay people continued to apply their preexisting, and often mistaken, understandings of these offenses.⁴⁷ When explicitly instructed to use a standard of proof threshold of ninety percent, simulated jurors reported using an average standard of eighty-five percent, which appears to have been a compromise between the instructed threshold and the one they reported absent an instruction (seventy-eight percent).⁴⁸ Notably, judges too seem to subject their compliance with legal rules to their personal sense of justice.⁴⁹

B. Curative Instructions

Another type of jury instructions is designed to "cure" the process of potential biases. Oftentimes, curative instructions are administered to remedy the exposure of jurors to potent information that is not admissible as evidence. Such exposure can originate outside the courtroom, typically in the form of pretrial publicity, or it can derive from statements uttered by the lawyers inside the courtroom, whether unwittingly or deliberately. At times, counsel will request the judge to call a mistrial, but such motions rarely succeed. More frequently, the judge will either deny the motion outright or issue an instruction admonishing the jury to ignore the information.

A number of psychological reasons warrant suspicion about the curative potential of judicial admonitions. Research on "ironic mental processes" has shown that instructing people to suppress a thought is a difficult mental feat, which can even backfire by increasing the salience of the thought.⁵⁰ Indeed, in some jury simulations, issuing instructions to ignore evidence actually increased the impact of the impermissible evidence on the decision.⁵¹ Likewise, research

45. Ogloff, *supra* note 37.

46. See NORMAN J. FINKEL, COMMONSENSE JUSTICE: JURORS' NOTIONS OF THE LAW ch. 17 (1995); PAUL H. ROBINSON & JOHN M. DARLEY, JUSTICE, LIABILITY, & BLAME ch. 7 (1995); Saul M. Kassir & Samuel R. Sommers, *Inadmissible Testimony, Instructions to Disregard, and the Jury: Substantive Versus Procedural Considerations*, 23 PERSONALITY & SOC. PSYCHOL. BULL. 1046, 1051 (1997).

47. Vicki L. Smith, *Prototypes in the Courtroom: Lay Representations of Legal Concepts*, 61 J. PERSONALITY & SOC. PSYCHOL. 857 (1991); Vicki L. Smith, *When Prior Knowledge and Law Collide: Helping Jurors Use the Law*, 17 LAW & HUM. BEHAV. 507, 533 (1993).

48. Mandeep K. Dhami, *On Measuring Quantitative Interpretations of Reasonable Doubt*, 14 J. EXPERIMENTAL PSYCHOL.: APPLIED 353 (2008).

49. A study with federal and state judges observed that the judges were more likely to follow the law and ignore a confession obtained through an impermissibly coercive interrogation when the suspect was charged with a less serious offense than when the offense was more serious. Jeffrey J. Rachlinski et al., *Context Effects in Judicial Decision Making* (forthcoming).

50. See generally Richard M. Wenzlaff & Daniel M. Wegner, *Thought Suppression*, 51 ANN. REV. PSYCHOL. 59 (2000).

51. See, e.g., Kerri L. Pickel, *Inducing Jurors to Disregard Inadmissible Evidence: A Legal Explanation Does Not Help*, 19 LAW & HUM. BEHAV. 407, 422-23 (1995); Sharon Wolf & David A.

on “reactance theory” suggests that people respond negatively to restrictions on their freedom. One reactive response is to increase the value of the deprived freedom, which could amount to placing greater emphasis on the information that is supposed to be ignored.⁵² Research on “hindsight bias” indicates that people have a difficult time imagining their state of knowledge in the absence of information they already have.⁵³ Research on “belief perseverance” indicates that people tend to adhere to beliefs that they have formed, even after learning that the information they used to form those beliefs has been discredited.⁵⁴

The research findings on the effectiveness of curative instructions are somewhat muddled.⁵⁵ Some clarity can be attained, first, by appreciating the role of people’s personal conceptions of justice, which pertain both to their individual policy preferences as well as to the perceived probativeness and reliability of the information. For example, instructions to ignore information were found to be effective when it was unreliable, such as when the wiretapped confession was barely audible,⁵⁶ or when the source of the information originated from a non-credible source.⁵⁷ Conversely, admonitions failed to defuse information that was considered to be probative of the defendant’s guilt.⁵⁸ Admonitions were also ineffective when the grounds for their exclusion could be considered “legal technicalities,” such as when the incriminating conversation was obtained by means of illegal wiretap,⁵⁹ or when the murder weapon was obtained in an illegal search.⁶⁰

In short, people appear to follow instructions that are consistent with their intuitions and inclinations, but disregard instructions that run contrary to

Montgomery, *Effects of Inadmissible Evidence and Level of Judicial Admonishment to Disregard on the Judgments of Mock Jurors*, 7 J. APPLIED SOC. PSYCHOL. 205, 207 (1977).

52. See Jack W. Brehm & John Sensenig, *Social Influence as a Function of Attempted and Implied Usurpation*, 4 J. PERSONALITY & SOC. PSYCHOL. 703, 703–07 (1966).

53. See generally Baruch Fischhoff, *Hindsight—Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty*, 1 J. EXPERIMENTAL PSYCHOL.: HUM. PERCEPTION & PERFORMANCE 288 (1975).

54. See generally Craig A. Anderson et al., *Perseverance of Social Theories: The Role of Explanation in the Persistence of Discredited Information*, 39 J. PERSONALITY & SOC. PSYCHOL. 1037 (1980).

55. See generally Nancy Steblay et al., *The Impact on Juror Verdicts of Judicial Instruction to Disregard Inadmissible Evidence: A Meta-Analysis*, 30 LAW & HUM. BEHAV. 469 (2006).

56. Kassin & Sommers, *supra* note 46.

57. Steven Fein et al., *Can the Jury Disregard that Information? The Use of Suspicion to Reduce the Prejudicial Effects of Pretrial Publicity and Inadmissible Testimony*, 23 PERSONALITY & SOC. PSYCHOL. BULL. 1215, 1223 (1997).

58. Admonitions were found to be ineffective when the evidence to be ignored contained a recording of the defendant’s admission of guilt, provided key facts about an obscure homicide, or connected the defendant’s gun to the crime. See Saul M. Kassin & Holly Sukel, *Coerced Confessions and the Jury: An Experimental Test of the “Harmless Error” Rule*, 21 LAW & HUM. BEHAV. 27, 43 (1997); Christine L. Ruva & Cathy McEvoy, *Negative and Positive Pretrial Publicity Affect Juror Memory and Decision Making*, 14 J. EXPERIMENTAL PSYCHOL.: APPLIED 226, 232 (2008); Stanley Sue et al., *Biasing Effects of Pretrial Publicity on Judicial Decisions*, 2 J. CRIM. JUST. 163, 170 (1974); Fein et al., *supra* note 57.

59. Kassin & Sommers, *supra* note 46.

60. Sue et al., *supra* note 58.

them.⁶¹ It is notable that judges behave in a similar manner. A study of sitting judges found that they cannot refrain from being swayed by probative information that they themselves ruled inadmissible.⁶²

A second way to clarify the research on curative instructions is to distinguish between instructions to disregard evidence completely and instructions to admit evidence for limited purposes. “Limited purpose” instructions are borne by the need to reconcile conflicting policy considerations. When it comes to admitting the defendant’s prior criminal record, for example, the rules of evidence split the difference by positing that evidence of prior convictions cannot be admitted to support inferences about the defendant’s propensity to commit the crime, but it can be admitted to prove other aspects of the charge or to impeach the credibility of the testifying defendant.⁶³ To address this palpable contradiction, the rules advise judges to “restrict the evidence to its proper scope and instruct the jury accordingly.”⁶⁴ Limited purpose instructions are premised on a belief in people’s ability to exert formidable control over their cognitive processing. This assumption runs contrary to the research. Many social judgments occur automatically,⁶⁵ and thus resist conscious control.⁶⁶ Given that turning information on and off at will is an unnatural task that is unparalleled in everyday life, it is not surprising to find that this instruction is basically ineffective in preventing the drawing of impermissible conclusions. A number of jury simulation studies find that despite limiting instructions, exposing jurors to the defendant’s prior criminal record results in higher conviction rates.⁶⁷ A

61. Instructions to ignore hearsay evidence—which can be seen as lying in the gray area of popular conceptions of justice—yield mixed results. Such instructions were unsuccessful in the study by Fein et al., *supra* note 57, at 1223, but generally successful in the study by Pickel, *supra* note 51, at 419.

62. Andrew J. Wistrich et al., *Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding*, 153 U. PA. L. REV. 1251, 1323 (2005).

63. Federal Rule of Evidence 404(b) instructs that “[e]vidence of other crimes, wrongs, or acts is not admissible to prove the character of a person in order to show action in conformity herewith.” Yet, Federal Rule of Evidence 609(a) instructs that under some conditions, evidence of prior convictions can be admitted for the purpose of “attacking a witness’s character for truthfulness.” Rule 404(b) itself also offers exceptions that permits evidence of other crimes, wrongs, or acts is for “other purposes, such as proof of motive, opportunity, intent, preparation, plan, knowledge, identity, or absence of mistake or accident.”

64. FED. R. EVID. 105.

65. See, e.g., Patricia G. Devine, *Stereotypes and Prejudice: Their Automatic and Controlled Components*, 56 J. PERSONALITY & SOC. PSYCHOL. 5 (1989); Daniel T. Gilbert, *Ordinary Personology*, in 2 HANDBOOK OF SOCIAL PSYCHOLOGY 89 (Daniel T. Gilbert et al. eds., 4th ed. 1998); James S. Uleman et al., *Spontaneous Inferences, Implicit Impressions, and Implicit Theories*, 59 ANN. REV. PSYCHOL. 329, 332 (2008).

66. John A. Bargh, *The Four Horsemen of Automaticity: Awareness, Intention, Efficiency, and Control in Social Cognition*, in 1 HANDBOOK OF SOCIAL COGNITION 1 (Robert S. Wyer, Jr. & Thomas K. Srull eds., 2d ed. 1994).

67. Anthony N. Doob & Hershi M. Kirshenbaum, *Some Empirical Evidence of the Effect of Section 12 of the Canada Evidence Act upon the Accused*, 15 CRIM. L.Q. 88, 94–96 (1972); Edith Greene & Mary Dodge, *The Influence of Prior Record Evidence on Juror Decision Making*, 19 LAW & HUM. BEHAV. 67, 72–73 (1995); Valerie P. Hans & Anthony N. Doob, *Section 12 of the Canada Evidence Act and the Deliberations of Simulated Jurors*, 18 CRIM. L.Q. 235, 242 (1976); Roselle L. Wissler & Michael J. Saks, *On the Inefficacy of Limiting Instructions: When Jurors Use Prior Conviction Evidence to*

meta-analysis shows that evidence of prior criminal behavior is generally resistant to curative instructions.⁶⁸ Paradoxically, the instruction limiting the use of prior criminal behavior bears only mixed and weak results with respect to its ostensibly justified objective, namely, impairing the credibility of the defendants' testimony.⁶⁹ Limiting instructions were also found to be ineffective when used to neutralize the effects of evidence of bad character,⁷⁰ and of joining multiple crimes into a single trial.⁷¹

Courts have long been skeptical about of the effectiveness of curative instructions. Justice Robert Jackson described them as an "unmitigated fiction,"⁷² and Judge Learned Hand characterized them as a "mental gymnastic" that is beyond the jury's powers.⁷³ Judge Jerome Frank described them as "a kind of 'judicial lie'" that "damages the decent judicial administration of justice."⁷⁴ Ignoring inadmissible evidence has also been likened to "unringing the bell,"⁷⁵ and "ignoring the smell of a skunk just thrown into the jury box."⁷⁶ For the most part, however, this skepticism has rarely been acknowledged by the Supreme Court in recent years. Under the stewardship of Chief Justices Burger, Rehnquist, and Roberts, the Court has shown no hesitation in asserting that juries can and will disregard inadmissible evidence when instructed to do so.⁷⁷

The practical implications of exposure to inadmissible evidence go beyond the limited effectiveness of curative instructions. The research shows that the judge's decision to admit the contested evidence can actually backfire by boosting its impact on the verdict.⁷⁸ This prospect of an adverse effect poses a serious dilemma for the attorneys. Refraining from objecting to prejudicial

Decide on Guilt, 9 LAW & HUM. BEHAV. 37, 41–42 (1985). Simulated jurors are somewhat more responsive to limiting instructions when the prior and current charges pertain to different crimes. See Wissler & Saks, *supra* note 67. The likely reason is that the intuitive assumptions of recidivism do not hold so strongly between different types of criminal behavior.

68. See generally Steblay et al., *supra* note 55.

69. Compare Wissler & Saks, *supra* note 67, with Sarah Tanford & Michele Cox, *The Effects of Impeachment Evidence and Limiting Instructions on Individual and Group Decision Making*, 12 LAW & HUM. BEHAV. 477, 488 (1988).

70. Jennifer S. Hunt & Thomas Lee Budesheim, *How Jurors Use and Misuse Character Evidence*, 89 J. APPLIED PSYCHOL. 347, 359 (2004).

71. See Edith Greene & Elizabeth F. Loftus, *When Crimes Are Joined at Trial*, 9 LAW & HUM. BEHAV. 193, 203–04 (1985).

72. *Krulewitch v. United States*, 336 U.S. 440, 453 (1949).

73. *Nash v. United States*, 54 F.2d 1006, 1007 (2d Cir. 1932).

74. *United States v. Grunewald*, 233 F.2d 556, 574 (2d Cir. 1956).

75. *Sandez v. United States*, 239 F.2d 239, 248 (9th Cir. 1956).

76. *Dunn v. United States*, 307 F.2d 883, 886 (5th Cir. 1962).

77. Exceptions are made only in face of an "overwhelming probability" that the jury will be unable to do so and there is a "strong likelihood that the effect of the evidence would be 'devastating' to the defendant." *Greer v. Miller*, 483 U.S. 756, 767 (1987) (citations omitted); *Richardson v. Marsh*, 481 U.S. 200 (1987). *But see Cruz v. New York*, 481 U.S. 186 (1987).

78. A meta-analysis showed that decisions to admit the contested evidence result in rates of conviction that are 50% higher than when the evidence is ruled inadmissible and followed by an instruction to ignore (74% v. 46%). Steblay et al., *supra* note 55, at 479.

evidence runs the risk of leaving the jury exposed to that prejudicial information. Yet objecting to the evidence will, at best, result in a relatively ineffective curative instruction and, at worst, will actually bolster the prejudicial effect. This perilous situation is often a double blow for the defendant in that the prejudicial fact—particularly a defendant’s prior record—has possibly already contributed to the police’s decision to investigate him in the first place and the prosecutor’s decision to follow through with the charges.

C. Capital Sentencing

Nowhere do jury instructions play as critical a role in legal decisionmaking as in the context of capital sentencing, and nowhere are the misconceptions of jurors’ performance as consequential.⁷⁹ Since the 1970s, the primary focus of the debate over the death penalty has not been the constitutionality of execution per se, but the fairness of its administration. The current state of the law was forged by two landmark decisions of the Supreme Court. In *Furman v. Georgia*, the majority of the justices decried the extant unbridled jury discretion, determining it to constitute cruel and unusual punishment, and thus violative of the Eighth Amendment.⁸⁰ Four years later, in *Gregg v. Georgia*, the Court reinstated the practice by endorsing Georgia’s regime of “guided discretion,” a sentencing framework that sought to rein in the jury’s discretion while maintaining enough latitude for individualized justice.⁸¹ The Court heralded the proposed regime, characterizing it as a “carefully drafted statute that ensures that the sentencing authority is given adequate information and guidance.”⁸²

While the specific instantiations of the guided discretion regime vary somewhat over time and across jurisdictions, they share three key features. First, the jury must find at least one of the enumerated “aggravating factors” that increase the blame of the defendant (for example, murdering a police officer or killing multiple victims). Typically, the finding of aggravating factors

79. The following discussion does not purport to cover the vast range of psychological issues involved in the topic of capital punishment. It is confined to jurors’ ability to comprehend and follow the sentencing instructions.

80. This ruling struck down all extant capital sentencing statutes and vacated the death sentences of some six hundred people then inhabiting death row. Justice Stewart’s plurality opinion found that, in practice, death sentencing was performed “wantonly and freakishly” befalling only the “capriciously selected random handful.” Receiving the punishment was not unlike “being struck by lightning.” *Furman v. Georgia*, 408 U.S. 238, 309–10 (1972). Justice Brennan likened the arbitrariness to a “lottery system.” *Id.* at 293. While most of the justices shunned the thorny issue of race, racial disparities were clearly on the justices’ mind. Justice Douglas highlighted the death penalty’s discriminative effect on black people, the poor, and the uneducated. For example, Douglas’s opinion included data showing a disparity in the rate of commutations of death sentences (20% for whites and 12% for blacks). *Id.* at 250. The discriminatory effects of the punishment were quite obvious. For example, in the forty years preceding the case, 405 black men had been executed for the crime of rape, whereas only 45 white people were executed for the same crime. CRAIG HANEY, *DEATH BY DESIGN: CAPITAL PUNISHMENT AS A SOCIAL PSYCHOLOGICAL SYSTEM* 14 (Ronald Roesch ed., 2005).

81. *Gregg v. Georgia*, 428 U.S. 153, 195 (1976). On the political environment that spurred the reinstatement of the death penalty statutes, see STUART BANNER, *THE DEATH PENALTY: AN AMERICAN HISTORY* (2002).

82. *Gregg*, 428 U.S. at 195.

must be made unanimously, and be determined beyond a reasonable doubt. Yet, even if aggravating factors are found, the death sentence is not automatic. The jurors must also consider factors that bode for sparing the defendant's life.⁸³ These "mitigating factors" can be established by the relatively low threshold of preponderance of the evidence (or no threshold at all), and do not require the unanimous agreement of all jurors. Mitigating factors need not be enumerated in the statute, but can encompass any fact that tends to lessen the defendant's blame. Finally, jurors are given a "decision rule" that prescribes how to weigh the aggravating factors against the mitigating factors in making the choice between a death sentence and the alternative punishment, which is typically life imprisonment without the possibility of parole.

The legality and legitimacy of the death penalty regime rely heavily on jurors' ability to abide by the procedure of guided discretion. Given that jurors find it difficult to comprehend and apply relatively straightforward instructions in non-capital procedures, it should come as no surprise to find that they fall short in mastering the intricate and unfamiliar mental procedure prescribed by the guided discretion regime. A series of studies by Craig Haney and Mona Lynch tested jury eligible Californians for their comprehension of the state's death sentencing instructions. Even after hearing the instructions three times, jury eligible respondents were largely unsuccessful in defining the terms "aggravating" and "mitigating." Fewer than half of the participants provided accurate definitions for both terms *aggravation* and *mitigation*, and only eight percent provided definitions that were legally correct.⁸⁴ Of particular interest are the misunderstandings of mitigation, which typically offers defendants the only chance of coming out of the process alive. About one-quarter of the respondents mistook two of the mitigating factors to be aggravating factors, and more than one-third misinterpreted the most important mitigating factor—"any other circumstances that extenuate the gravity of the crime"—as grounds to buttress a death verdict.⁸⁵ Respondents also manifested low comprehension of the all-important decision rule. Only one-half stated (correctly) that a death verdict was impermissible when the mitigating factors outweighed the aggravating factors, and only fifteen percent said the same (correctly) when the

83. *Lockett v. Ohio*, 438 U.S. 586, 601 (1978). More recently, the Court has stated that jurors ought to be given the opportunity "to give meaningful consideration and effect to all mitigating evidence that might provide a basis for refusing to impose the death penalty." *Abdul-Kabir v. Quarterman*, 550 U.S. 233, 234 (2007).

84. Correct or partially correct definitions were given for the term "aggravating" by 79% of the respondents, but only by 59% for the term "mitigating." Craig Haney & Mona Lynch, *Comprehending Life and Death Matters: A Preliminary Study of California's Capital Penalty Instructions*, 18 LAW & HUM. BEHAV. 411, 421 (1994) [hereinafter Haney & Lynch, *Comprehending Life and Death Matters*]. See also Craig Haney & Mona Lynch, *Clarifying Life and Death Matters: An Analysis of Instructional Comprehension and Penalty Phase Closing Arguments*, 21 LAW & HUM. BEHAV. 575, 579–80 (1997) [hereinafter Haney & Lynch, *Clarifying Life and Death Matters*].

85. The likely reason for the difficulty in comprehending the catch-all mitigator is the relative unfamiliarity of the word "extenuate." Haney & Lynch, *Comprehending Life and Death Matters*, *supra* note 84.

factors were balanced equally. Forty percent stated (incorrectly) that death was mandatory when the aggravating factors outweighed the mitigating ones.⁸⁶ A follow up study by Haney has found that comprehension improves somewhat thanks to the psycholinguistically improved instructions that were put into effect in California in 2006,⁸⁷ and it can be improved even more by using instructions that provide jurors with thematic and case specific examples presented in a relevant and concrete manner. Still, even with these modified instructions, the level of comprehension remains rather low.⁸⁸

Another series of studies was conducted by Richard Wiener and his colleagues, testing juror comprehension of the Missouri death penalty instructions. These studies found overall comprehension levels of between fifty-five and sixty percent, which is just slightly better than guessing.⁸⁹ Particularly poor levels were observed with respect to some of the procedure's central features, including the need for unanimity in finding aggravating factors, the freedom to consider mitigating factors that are not enumerated in the statute, and the decision rule for weighing the countervailing factors.⁹⁰ Substandard comprehension was also observed in studies testing the instructions used in Illinois,⁹¹ Florida,⁹² Ohio,⁹³ and Tennessee.⁹⁴

Given that the guided discretion regime works primarily to inhibit jurors' discretion to award death sentences, poor comprehension is prone to undermine the protections promised by this procedure. The misconceptions observed in the studies—notably, relaxing the limits on finding aggravating factors, narrowing the scope for establishing mitigating factors, and

86. Haney & Lynch, *Clarifying Life and Death Matters*, *supra* note 84.

87. California's new pattern instructions, CALCRIM, were adopted by Judicial Council in 2005 and came into effect in 2006. The primary innovation in these instructions is the use of psycholinguistic principles to simplify the language of the instructions.

88. Even with the best instructions, lay people reach comprehension scores that are below half of the correct responses. Amy E. Smith & Craig Haney, *Getting to the Point: Attempting to Improve Juror Comprehension of Capital Penalty Phase Instructions*, 35 LAW & HUM. BEHAV. 339, 347–49 (2011).

89. Richard L. Wiener et al., *Comprehensibility of Approved Jury Instructions in Capital Murder Cases*, 80 J. APPLIED PSYCHOL. 455, 459 (1995).

90. Richard L. Wiener et al., *The Role of Declarative and Procedural Knowledge in Capital Murder Sentencing*, 28 J. APPLIED SOC. PSYCHOL. 124, 13 (1998); Richard L. Wiener et al., *Guided Jury Discretion in Capital Murder Cases: The Role of Declarative and Procedural Knowledge*, 10 PSYCHOL. PUB. POL'Y & L. 516, 570 (2004) [hereinafter Wiener et al., *Guided Jury Discretion*].

91. Only 40% of these respondents correctly understood the rule regarding the non-enumerated mitigating factors, and 33% miscomprehended the non-unanimity rule for mitigating factors. Shari S. Diamond & Judith N. Levi, *Improving Decisions on Death by Revising and Testing Jury Instructions*, 79 JUDICATURE 224, 230 (1996).

92. The overall rate of correct responses in this study was 47.5% (for death-eligible respondents). Charles W. Otto et al., *Improving Comprehension of Capital Sentencing Instructions: Debunking Juror Misconceptions*, 53 CRIME & DELINQ. 502, 509 (2007).

93. James Frank & Brandon K. Applegate, *Assessing Juror Understanding of Capital-Sentencing Instructions*, 44 CRIME & DELINQ. 412, 419 (1998).

94. Correct responses ranged from 22% to 83%, with most values around the 50% accuracy. Michael B. Blankenship et al., *Jurors' Comprehension of Sentencing Instructions: A Test of the Death Penalty Process in Tennessee*, 14 JUST. Q. 325, 334–39 (1997).

underappreciating the freedom to vote for imprisonment even in the face of aggravating factors—all result in higher rates of death sentences than intended by the law. Indeed, studies find a consistent relationship between comprehension and death sentencing; the less jurors understand the instructions, the more prone they are to vote for death.⁹⁵ Relatedly, comprehension is inversely correlated with jurors' support for the death penalty, in that those who strongly support the practice tend to display poorer comprehension.⁹⁶ Thus, the procedure of death qualification—which excuses jurors who express a strong aversion to voting for a death sentence—decreases the jury's overall level of comprehension, and thus further increases the likelihood of a death sentence.⁹⁷ Instruction comprehension also interacts with racial bias. Low comprehension has been found to be related to a greater tendency to sentence black defendants to death, particularly when their victims were white.⁹⁸

The findings from the laboratory and field studies are notably consistent with data from the Capital Juror Project, which has conducted in-depth interviews with some 1,200 jurors who served on 350 capital sentencing cases in 14 states. One of these studies found that almost half of the jurors fundamentally misunderstood the concept of mitigation, believing that under certain circumstances death was “the only acceptable punishment.” Not surprisingly, jurors who failed to appreciate their duty to explore mitigating factors were five times more likely to decide prematurely on a death sentence than those who understood the instruction correctly. The influence of jurors' preconceived notions of justice was manifested by a greater endorsement of mandatory death (which is always incorrect) for more reprehensible murder charges.⁹⁹ A significant number of South Carolina jurors mistakenly stated that

95. Diamond & Levi, *supra* note 91, at 231; Wiener et al., *supra* note 90, at 463; Wiener et al., *supra* note 89, at 138.

96. Mona Lynch & Craig Haney, *Capital Jury Deliberation: Effects on Death Sentencing, Comprehension, and Discrimination*, 33 LAW & HUM. BEHAV. 481, 489 (2009); Mona Lynch & Craig Haney, *Discrimination and Instructional Comprehension: Guided Discretion, Racial Bias, and the Death Penalty*, 24 LAW & HUM. BEHAV. 337, 347 (2000) [hereinafter Lynch & Haney, *Discrimination and Instructional Comprehension*].

97. The more serious problem caused by the policy of death qualification is that it precludes people who tend to be more lenient from sitting on juries. The research shows that the preclusion of these venire members increases the likelihood that the defendants will be convicted in the first place. See Brooke M. Butler & Gary Moran, *The Role of Death Qualification in Venirepersons' Evaluations of Aggravating and Mitigating Circumstances in Capital Trials*, 26 LAW & HUM. BEHAV. 175, 183 (2002); Kevin O'Neil et al., *Exploring the Effects of Attitudes Toward the Death Penalty on Capital Sentencing Verdicts*, 10 PSYCHOL., PUB. POL'Y & L. 443, 463–66 (2004).

98. Of the jurors whose comprehension was in the lower half, 60% voted for the death sentence when the defendant was black, whereas only 41% voted the same when he was white. The disparity was even greater between black defendants charged with killing white victims (68%) and white defendants charged with killing black victims (36%). Lynch & Haney, *Discrimination and Instructional Comprehension*, *supra* note 96, at 350.

99. While some seventy percent of the respondents believed that death was mandatory for repeat murderers, only about one quarter shared the same position for murders that are generally perceived to be less condemnable. William J. Bowers et al., *The Capital Sentencing Decision: Guided Discretion*,

the law prescribed certain aggravating factors that were intuitively plausible, but that were not mentioned in the instructions.¹⁰⁰

The most telling evidence of the shortcomings of the guided discretion regime comes from naturalistic sentencing data. A large body of econometric research shows that death sentencing in the post-*Gregg* regime continues to be performed in an unprincipled manner. The landmark study was conducted by David Baldus and his colleagues on death sentencing in the state of Georgia, where the cases of *Furman* and *Gregg* originated. This study, which examined the 2,484 homicide cases decided between 1973 and 1979, concentrated on racial discrimination, the most salient impermissible influencing factor. The study found that jury-sentencing decisions were affected by the races of the defendants and their victims. White defendants were sentenced to death at a rate of 8% for killing white victims but only 3% for killing black victims. In contrast, 22% of black defendants were sentenced to death for killing white victims, but only 1% received death sentences for killing black victims.¹⁰¹ The findings of this study were presented to the Court in *McCleskey v. Kemp*, but they failed to rattle the justices' trust in the restraining power of the guided discretion doctrine.¹⁰² Racial discrimination in meting out the death penalty has been observed in a number of other jurisdictions, including Northern states,¹⁰³

Reasoned Moral Judgment, or Legal Fiction, in AMERICA'S EXPERIMENT WITH CAPITAL PUNISHMENT: REFLECTIONS ON THE PAST, PRESENT, AND FUTURE OF THE ULTIMATE PENAL SANCTION 413 (James R. Acker et al., eds., 2d ed. 2003).

100. Some thirty percent of jurors who sat on South Carolina death penalty cases stated that the law requires imposition of a death sentence where the jury finds that the defendant will be dangerous in the future. This erroneous belief is patently driven by people's preconceptions, because future dangerousness is not even mentioned in the jury instructions. Theodore Eisenberg & Martin T. Wells, *Deadly Confusion: Juror Instructions in Capital Cases*, 79 CORNELL L. REV. 1, 6 (1993).

101. The study also found strong disparities in the exercise of prosecutorial discretion, as the death penalty was sought in 19% of the cases in which a white man was charged with killing a black man, but in 70% of the cases where the races were reversed. A regression analysis showed that killing a white victim had roughly the same impact on the likelihood of receiving a death sentence as did aggravating factors that were officially prescribed, such as the defendant's serious prior criminal record or the fact that the death was perpetrated for the purpose of an armed robbery. Killing a white defendant had three times the impact as killing a police officer. DAVID C. BALDUS ET AL., *EQUAL JUSTICE AND THE DEATH PENALTY: A LEGAL AND EMPIRICAL ANALYSIS* 327 (1990). See also David C. Baldus et al., *Comparative Review of Death Sentences: An Empirical Study of the Georgia Experience*, 74 J. CRIM. L. & CRIMINOLOGY 661 (1983); David C. Baldus et al., *Race Discrimination and the Death Penalty in the Post-Furman Era: An Empirical and Legal Analysis with Recent Findings from Philadelphia*, 83 CORNELL L. REV. 1638, 1694 (1998); John Blume et al., *Explaining Death Row's Population and Racial Composition*, 1 J. EMPIRICAL LEGAL STUD. 165, 204 (2004).

102. The Court simply asserted that the jury's discretion is "controlled by clear and objective standards so as to produce non-discriminatory application." *McCleskey v. Kemp*, 481 U.S. 279, 302 (1987).

103. A study of capital sentencing in Philadelphia revealed that juries were twice as likely to sentence a black defendant to death as a non-black defendant (24% versus 12%). This study contained 338 sentencing decisions rendered between 1978 and 2000. David C. Baldus & George Woodworth, *Race Discrimination in the Administration of the Death Penalty: An Overview of the Empirical Evidence with Special Emphasis on the Post-1990 Research*, 39 CRIM. L. BULL. 194, 206 (2003). In Maryland, juries sentenced to death almost half of the black defendants convicted of killing white victims but only one-third of white defendants convicted of killing black victims. Overall, taking the entire process into

and has been confirmed by the U.S. General Accounting Office.¹⁰⁴ In sum, it appears that the touted regime of guided discretion provides little guidance, yet allows much discretion. Death penalty sentences continue to be meted in the prejudicial manner found to be unconstitutional in *Furman v. Georgia*.

III

JURORS' ASSURANCE OF IMPARTIALITY

It is well established that people's decisions are affected by their attitudes, preferences, beliefs, moral convictions, and the like. The legal system is cognizant and accepting of these effects on juror decisionmaking, and it would be quite impossible and indeed unwise to try to rid jurors of them. At the same time, there are good reasons to curb the excesses of idiosyncratic influences, since the integrity of the criminal justice process is compromised when the jurors' decision in a given case is overpowered by their prior positions. Hence the Sixth Amendment's guarantee of the right to an impartial jury. Weeding out partial jurors is the primary ostensible objective of the jury selection procedure.¹⁰⁵ A key factor in deciding whether to place a veneer person on the jury is her assurance to the court that she can be unbiased. Assurances of impartiality are used similarly during the trial to allay any suspicion of contamination, such as following exposure to inadmissible information.

The shared wisdom within the legal system is that jurors are quite capable of assessing their objectivity. While the Supreme Court acknowledges the difficulty in knowing "the imponderables which cause one to think what he thinks," it insists that jurors can rise to the occasion: "[S]urely one who is trying as an honest man to live up to the sanctity of his oath is well qualified to say whether he has an unbiased mind in a certain matter."¹⁰⁶ Whether jurors can reliably vouch for their own impartiality is an empirical question.

To meet this expectation, jurors would need, first, to be sufficiently introspective so as to be able to identify which factors actually influence their

consideration, black defendants charged with killing white victims are 4.1 times more likely to receive a death sentence than when the races are reversed. Raymond Paternoster et al., *Justice By Geography and Race: The Administration of the Death Penalty in Maryland, 1978-1999*, 4 MD. L.J. RACE, RELIGION, GENDER 1, 39 (2004). In New Jersey, the rate of death sentences awarded to killers of white victims was 1.7 times higher than for killers of black victims. Among all death eligible cases, 12% of cases resulted in death sentences for killing white victims, but only 7% for killing black victims. DAVID S. BAIME, REPORT TO THE NEW JERSEY SUPREME COURT: SYSTEMIC PROPORTIONALITY REVIEW PROJECT 2004-2005 TERM 6, 17-18 (2005), <http://www.judiciary.state.nj.us/pressrel/Baime2005Report12-16-05.pdf>.

104. This review concluded that even after controlling for all legally relevant variables, differences remain in the likelihood of receiving the death penalty based on race of victim. U.S. GOV'T ACCOUNTABILITY OFFICE, GAO/GGD-90-57, DEATH PENALTY SENTENCING: RESEARCH INDICATES PATTERN OF RACIAL DISPARITIES (1990). For a summary of findings from seventeen states and the federal system, see generally Baldus & Woodworth, *supra* note 103.

105. As some critics charge, the litigants' true objective in jury selection is not to attain a fair representation of the community or to maximize the chances for a fair trial in other ways. Rather, they use the process to stack the jury box with people whom they expect will vote in their favor.

106. *Smith v. Phillips*, 455 U.S. 209, 217 (1982).

decisions. Psychological research has long distinguished between access to “declarative” and “procedural” types of knowledge. People have generally reliable access to the former, which enables fairly accurate reporting of the contents of one’s thoughts, such as one’s beliefs, decisions, and emotional states. At the same time, the cognitive mechanisms and processes used to arrive at those mental states are generally inaccessible to introspection.¹⁰⁷ Notably, people have only limited access to the situational factors that actually influence their decisions¹⁰⁸ and behaviors.¹⁰⁹ Still, people habitually insist on their introspective abilities,¹¹⁰ and when asked about the reasons for their decisions and behaviors, they readily provide spurious explanations: a phenomenon known as “telling-more-than-you-can-know.”¹¹¹

The limitations on introspection are even more pronounced when it comes to the acknowledgment of bias. In a culture that places a premium on rationality, dependability and predictability, the notion of bias carries a pejorative connotation. The ubiquitous motive of maintaining a positive self-conception and public image thus hinders any inclination to concede being biased. Notably, people tend to believe that they are fair and objective.¹¹² While people might be open to admitting in general terms that their judgment can be subject to occasional bias, they rarely concede bias in their current judgments,

107. A sizable body of research shows that much of our reasoning and decisionmaking processes are conducted without awareness. See generally John A. Bargh & Ezequiel Morsella, *The Unconscious Mind*, 3 PERSP. ON PSYCHOL. SCI. 73 (2008). Studies show, for example, that participants are often unaware of changes in their attitudes that were precipitated by the experiment. Daryl J. Bem & H. Keith McConnell, *Testing the Self-Perception Explanation of Dissonance Phenomena: On the Salience of Premanipulation Attitudes*, 14 J. PERSONALITY & SOC. PSYCHOL. 23, 29 (1970); George R. Goethals & Richard F. Reckman, *The Perception of Consistency in Attitudes*, 9 J. EXPERIMENTAL SOC. PSYCHOL. 491, 498 (1973). For findings of unawareness of belief changes, see generally Keith Holyoak & Dan Simon, *Bidirectional Reasoning in Decision Making by Constraint Satisfaction*, 128 J. EXPERIMENTAL PSYCHOL.: GEN. 3 (1999) (studies 2, 3).

108. Richard E. Nisbett & Timothy D. Wilson, *Telling More Than We Can Know: Verbal Reports on Mental Processes*, 84 PSYCHOL. REV. 231, 233 (1977). Studies of decisionmaking find that participants’ reports of the weights they assign to decision factors are often substantially discrepant from the weights objectively derived from their decisions. Paul Slovic & Sarah Lichtenstein, *Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment*, 649 ORGANIZATIONAL BEHAV. & HUM. PERFORMANCE 649, 715–16 (1971).

109. See, e.g., BIBB LATANE & JOHN M. DARLEY, *THE UNRESPONSIVE BYSTANDER: WHY DOESN’T HE HELP?* (1970).

110. The belief in one’s introspective abilities has been labeled the “introspection illusion.” Emily Pronin et al., *Objectivity in the Eye of the Beholder: Divergent Perceptions of Bias in Self Versus Others*, 111 PSYCHOL. REV. 781, 784 (2004).

111. Nisbett & Wilson, *supra* note 108, at 255.

112. For people’s beliefs in their fairness, see generally Wim B. Liebrand et al., *Why We Are Fairer Than Others: A Cross-Cultural Replication and Extension*, 22 J. EXPERIMENTAL SOC. PSYCHOL. 590 (1986). The beliefs in one’s objectivity is captured by a sense of “naïve realism,” which refers to the “unshakable conviction that he or she is somehow privy to an invariant, knowable, objective reality—a reality that others will also perceive faithfully, provided that they are reasonable and rational, a reality that others are apt to misperceive only to the extent that they (in contrast to oneself) view the world through a prism of self-interest, ideological bias, or personal perversity.” Robert J. Robinson et al., *Actual Versus Assumed Differences in Construal: “Naïve realism” in Intergroup Perception and Conflict*, 68 J. PERSONALITY & SOC. PSYCHOL. 404, 405 (1995).

or in any specific judgment for that matter.¹¹³ Limited introspection and the aversion to conceding bias combine to create an “illusion of objectivity.”¹¹⁴

The disinclination to admit to bias has been observed also in studies of jury behavior. Simulated jurors are consistently found to be swayed by extra-evidential factors even as they deny any such influence. Such findings were made in studies that exposed jurors to media publicity about the crime,¹¹⁵ the defendant’s prior record,¹¹⁶ gruesome photographs,¹¹⁷ and to coerced confessions.¹¹⁸ The motivation to appear unbiased is likely to be heightened by the courtroom environment and the avowed commitment to impartiality and fairness. In sum, while there is good reason to exclude from jury service people who say that they cannot be fair and objective,¹¹⁹ the system should be less trusting of self-proclamations of objectivity.

IV

THE PROSECUTION’S HEIGHTENED BURDENS

One of the distinctive features of the criminal process is the imposition of

113. Joyce Ehrlinger et al., *Peering Into the Bias Blind Spot: Peoples Assessments of Bias in Themselves and Others*, 31 PERSONALITY & SOC. PSYCHOL. BULL. 680 (2005). For example, participants denied being biased when they over predicted aggressive behavior of people with typical Afrocentric features. Irene V. Blair et al., *The Use of Afrocentric Features as Cues for Judgment in the Presence of Diagnostic Information*, 35 EUR. J. SOC. PSYCHOL. 59, 64 (2005). In refereeing social science research, scientists denied being influenced by their prior conceptions, even as they were swayed by them. Jonathan J. Koehler, *The Influence of Prior Beliefs on Scientific Judgments of Evidence Quality*, 56 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 28, 39 (1993). In a study of deceit detection that contained biasing information about the witness, only one participant out of a hundred cited that information as a reason for her judgment. Granhag & Strömwall (2000). In a study simulating an investigation of a case of academic misconduct, participants deemed their judgments as highly objective even though they were strongly influenced by their role assignment. Dan Simon et al., Conference on Empirical Legal Studies, On the Objectivity of Investigations: An Experiment (Sept. 9–10, 2008).

114. Ziva Kunda, *The Case for Motivated Reasoning*, 108 PSYCHOL. BULL. 480, 482–83 (1990); Tom Pyszczynski & Jeff Greenberg, *Toward an Integration of Cognitive and Motivational Perspectives on Social Inference: A Biased Hypothesis-Testing Model*, in 20 ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY 297, 330–31 (Leonard Berkowitz ed., 1987); Pronin et al., *supra* note 110. One of the prominent and invariant features of biased processes is that they are deeply hidden beneath the level of conscious awareness. Timothy D. Wilson & Nancy Brekke, *Mental Contamination and Mental Correction: Unwanted Influences on Judgments and Evaluations*, 116 PSYCHOL. BULL. 117, 118 (1994).

115. Gary Moran & Brian L. Cutler, *The Prejudicial Impact of Pretrial Publicity*, 21 J. APPLIED SOC. PSYCHOL. 345, 360–63 (1991); James R. P. Ogloff & Neil Vidmar, *The Impact of Pretrial Publicity on Jurors: A Study to Compare the Relative Effects of Television and Print Media in a Child Sex Abuse Case*, 18 LAW & HUM. BEHAV. 507, 520–21 (1994); Stanley Sue et al., *Authoritarianism, Pretrial Publicity, and Awareness of Bias in Simulated Jurors*, 37 PSYCHOL. REP. 1299, 1302 (1975).

116. Roselle L. Wissler & Michael J. Saks, *On the Inefficacy of Limiting Instructions: When Jurors Use Prior Conviction Evidence to Decide on Guilt*, 9 LAW & HUM. BEHAV. 37, 43–47 (1985); Greene & Dodge, *supra* note 67.

117. Kevin Douglas et al., *The Impact of Graphic Photographic Evidence on Mock Jurors’ Decisions in a Murder Trial: Probative or Prejudicial?*, 21 LAW & HUM. BEHAV. 485, 497–99 (1997).

118. Kassir & Sukel, *supra* note 58.

119. Even if such a response is insincere, these people appear to be driven by ulterior motives, which renders them unsuitable for the task.

heightened demands on the prosecution, namely the “presumption of innocence” and the standard of proof “beyond a reasonable doubt.” The asymmetry underlying these two measures stems to a large extent from the legal system’s avowed discrepant aversion to false convictions and false acquittals. As stated by the Supreme Court, “[I]t is far worse to convict an innocent man than to let a guilty man go free.”¹²⁰ This disparity was famously quantified by Blackstone’s assertion that “the law holds that it is better that ten guilty persons escape than that one innocent suffer.”¹²¹ These measures are generally extolled in the legal discourse, yet their exact meaning and practical effects are far from clear.

A. The Presumption of Innocence

One possible conception of the presumption of innocence is that it is directed at law enforcement agencies and courts as an overarching principle to govern their affairs throughout the criminal justice process. According to this formulation, any ambiguity should be resolved with an eye toward favoring the defendant. The Supreme Court, however, has rejected this interpretation. The Court insists that the presumption is directed only at the fact finder at the trial phase as an aid in the process of deciding verdicts.¹²²

One way in which the presumption of innocence could aid the decisionmaking task is to counter the possible suspicion that arises from the fact that the defendant is the subject of a criminal prosecution.¹²³ Even prior to the presentation of the prosecution’s evidence, it would be rational for a fact finder to regard the accused as the likely perpetrator of the crime: Why else would she be brought to trial? According to one survey, lay people estimate the initial probability of the defendant’s guilt at close to fifty percent.¹²⁴ Although this estimate seems intuitively sensible, it is considerably higher than a presumption of innocence would imply. Whether the judicial instruction is potent enough to counter this suspicion and anchor it at a level that befits the criminal system’s precepts is an empirical question. The scarce available data leave this issue unsolved. For one, it is unclear whether jurors comprehend this instruction correctly. In a field study of Wyoming jurors who had received jury instructions,

120. *In re Winship*, 397 U.S. 358, 372 (1970).

121. 4 WILLIAM BLACKSTONE, COMMENTARIES *358.

122. *Bell v. Wolfish*, 441 U.S. 520 (1979). For an expansive view of the presumption of innocence, see Keith A. Findley, *Defining Innocence*, 74 ALB. L. REV. 1157, 1194 (2011). On the presumption’s possible expressive function as a means of punctuating the respect that society places on the freedoms and reputations of its members, see Laurence H. Tribe, *Trial By Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1370 (1972).

123. For example, the Connecticut pattern jury instructions explain that at the beginning of the trial, the defendant “stood before you free of any bias, prejudice or burden arising from his position as the accused.” 5 DAVID M. BORDEN & LEONARD ORLAND, CONNECTICUT PRACTICE SERIES, CRIMINAL JURY INSTRUCTIONS § 2.8 (4th ed. 2011). This view was espoused by the Supreme Court in *Bell v. Wolfish*, 441 U.S. 520, 533 (1979).

124. Dan Simon et al., Professor of Law and Psychology, Univ. of S. Cal., American Psychology-Law Society Annual Conference: Jurors’ Background Knowledge and Beliefs (Mar. 6–8, 2008).

forty percent indicated that the charges brought against the defendant were evidence of his guilt.¹²⁵ One laboratory simulation demonstrated that the presumption did have a constraining effect,¹²⁶ while another study revealed that simulated jurors assigned low initial probabilities of guilt at first, but abandoned them once the prosecution evidence starts to come in.¹²⁷

The presumption of innocence could also aid the decisionmaking task by instantiating the prosecution's burden of production, that is, the rule that places the onus of producing the evidence on the shoulder of the prosecution, and relieves the defendant from having to prove her innocence. The research shows that jurors do not always comprehend the asymmetric allocation of the burden, a concept that has few parallels in everyday life. After being given Florida's pattern jury instructions, only half of the jurors surveyed in a Florida courthouse maintained that the defendant need not present any evidence to prove his innocence,¹²⁸ and only three in ten of another group of Florida jurors correctly understood the prosecution's burden.¹²⁹ A failure to understand the allocation of the burden was also displayed by one-fifth of Wyoming jurors,¹³⁰ and two-thirds of a sample of Michigan jurors.¹³¹

Finally, the presumption of innocence can be conceived as serving the role of tiebreaker, positing that if the evidence against the defendant does not meet the requisite standard of proof, she should be acquitted. Here, too, there is reason for concern. Almost one-quarter of instructed Florida jurors believed that when faced with two equally reasonable sets of evidence, the defendant should be convicted.¹³²

B. The Standard of Proof: Beyond a Reasonable Doubt

The criminal process's signature standard of proof—beyond a reasonable doubt—has received a fair amount of scholarly attention over the years, yet it remains elusive and deeply disputed.¹³³ For one, there is considerable

125. Half of these respondents interpreted the charges as "strong evidence" of guilt. Bradley Saxton, *How Well do Jurors Understand Jury Instructions? A Field Test Using Real Juries and Real Trials in Wyoming*, 33 LAND & WATER L. REV. 59, 101 (1998).

126. Vicki S. Helgeson & Kelly G. Shaver, *Presumption of Innocence: Congruence Bias Induced and Overcome*, 20 J. APPLIED SOC. PSYCHOL. 276, 289–96 (1990).

127. Thomas M. Ostrom et al., *An Integration Theory Analysis of Jurors' Presumptions of Guilt or Innocence*, 36 J. PERSONALITY & SOC. PSYCHOL. 436, 445–46 (1978).

128. Strawn & Buchanan, *supra* note 38, at 481.

129. The accuracy of non-instructed jurors was even lower, suggesting that the instruction had a moderately positive impact on comprehension. Raymond W. Buchanan et al., *Legal Communication: An Investigation of Juror Comprehension of Pattern Instructions*, 26 COMMUNICATION MONOGRAPHS 31, 34 (1978).

130. Some twenty percent of the respondents stated that once the state has made its case, it becomes the defendant's responsibility to provide proof that he did not commit the crime. Saxton, *supra* note 125, at 97.

131. A majority of these jurors had recently served as jurors in a criminal trial. Reifman et al., *supra* note 39, at 546–47.

132. Strawn & Buchanan, *supra* note 38.

133. Legal historian James Whitman argues that the term was not originally intended as a standard

disagreement as to whether the standard should pertain to the fact finder's subjective state of mind, or to a property of the evidence itself. While most commentators agree with the former formulation,¹³⁴ there is disagreement as to whether the standard posits a numeric criterion for assessing the probability of guilt, or whether it prescribes a threshold for gauging one's belief in the inculpatory power of the evidence.¹³⁵ The latter option, which figures more prominently in courtroom practice and legal literature,¹³⁶ is essentially a meta-cognitive judgment, not unlike a measure of confidence. There is also disagreement about the appropriate meaning of the reasonableness criterion, namely, whether it pertains to the strength of the doubt or to a substantive judgment of its underlying rationality.¹³⁷ Another controversy revolves around the appropriateness of providing jurors with definitions and explanations of the standard. While some commentators and courts propose that the standard should be communicated to the jury without attempting to explain it,¹³⁸ others insist on supplementing the definition with explanatory language. Among the latter, some would focus on the doubt itself, for example, "not a possible doubt, a speculative, imaginary or forced doubt,"¹³⁹ whereas most instructions speak of an absence of such a doubt¹⁴⁰ as in "firmly convinced,"¹⁴¹ "an abiding conviction,"¹⁴² "settled belief," and "near certitude."¹⁴³ In practice, many

of proof, but as a mechanism for providing jurors with moral comfort in the exercise of their solemn task. As such, the standard is ill equipped to aid contemporary jurors in the quest for determining the truth. JAMES Q. WHITMAN, *THE ORIGINS OF REASONABLE DOUBT: THEOLOGICAL ROOTS OF THE CRIMINAL TRIAL* 2–3 (2008).

134. For a minority view on this matter, see Larry Laudan, *Is Reasonable Doubt Reasonable?*, 9 *LEGAL THEORY* 295 (2004).

135. This issue has been a focal point in the debate between the "Bayesian" and "Bayesian-skeptical" camps. Compare RICHARD O. LEMPERT ET AL., *A MODERN APPROACH TO EVIDENCE: TEXT, PROBLEMS, TRANSCRIPTS AND CASES* 228–39 (3d ed. 2000), with Ronald J. Allen, *Rationality, Algorithms, and Juridical Proof: A Preliminary Inquiry*, 1 *INT'L J. EVIDENCE & PROOF* 254, 255–56 (1997).

136. In the seminal *In re Winship*, Justice Brennan explained that the standard of proof instructs the fact finder about "the degree of confidence he is expected to have in the correctness of his factual conclusions." 397 U.S. 358, 370 (1970). This conception is also consistent with the various definitions described below, such as "firmly convinced," "abiding conviction," "near certitude," and the like.

137. For a minority view supporting the latter position, see Laudan, *supra* note 134.

138. For example, in *People v. Malmenato*, the Illinois Supreme Court stated, "Reasonable doubt is a term which needs no elaboration and we have so frequently discussed the futility of attempting to define it that we might expect the practice to be discontinued." 14 Ill. 2d 52, 61 (1958).

139. THE SUPREME COURT COMM. ON STANDARD JURY INSTRUCTIONS IN CRIMINAL CASES, FLORIDA STANDARD JURY INSTRUCTIONS IN CRIMINAL CASES 2.03, available at http://www.floridasupremecourt.org/jury_instructions/instructions.shtml (last visited Jan. 7, 2012).

140. See Lawrence M. Solan, *Refocusing the Burden of Proof in Criminal Cases: Some Doubt About Reasonable Doubt*, 78 *TEX. L. REV.* 105, 111 (1999).

141. See, e.g., 5 DAVID M. BORDEN & LEONARD ORLAND, *supra* note 123, § 2.10 ("If, based on your consideration of the evidence, you are firmly convinced that the defendant is guilty of the crime charged, you must find him guilty."). This version, recommended by the Federal Judicial Center, was also endorsed by Justice Ginsburg in *Victon v Nebraska*. 511 U.S. 1, 25 (1994).

142. For example, the California pattern jury instruction instructs: "Proof beyond a reasonable doubt is proof that leaves you with an abiding conviction that the charge is true." JUDICIAL COUNCIL OF CAL. CRIMINAL JURY INSTRUCTIONS, CALCRIM No. 220 (2011).

jurisdictions mix numerous definitions and explanations, thus making for unwieldy instructions. For example, a Louisiana death sentence was based on a jury instruction that defined the standard in no fewer than eight ways.¹⁴⁴

The available research on the standard of proof leaves much unresolved. A number of surveys have sought to gauge its effect by asking people to express their conception of the standard in terms of a probabilistic assessment of their threshold. The mean responses in most studies hover around 85%,¹⁴⁵ which is a little lower than the value of 90% or higher that is commonly assumed in legal discourse.¹⁴⁶ Although the standard is conceived as invariant to the specifics of the case, varying the case characteristics in the same study has resulted in thresholds ranging from 79% to 94%.¹⁴⁷ Most notably, there is considerable variation in assessments of standard. One large sample of respondents reported an average of 64%,¹⁴⁸ whereas large numbers of respondents insist on complete certainty. Thresholds of 100% were reported by almost 70% of a sample of Michigan jurors,¹⁴⁹ and by more than half of the 500 jury eligible respondents in a Canadian study.¹⁵⁰ A large survey conducted in the United Kingdom obtained thresholds of 100% certainty from no fewer than one-half of the jurors and almost one-third of the legal professionals and magistrates surveyed.¹⁵¹ Naturally, requiring absolute certainty should make convictions virtually impossible; this is clearly contradicted by experimental and naturalistic data. These high standards are also out of whack with lay people's intuitions

143. *Jackson v. Virginia*, 443 U.S. 307, 315 (1979).

144. The definition included "real tangible substantial basis," not "caprice and conjecture," not "give rise to a grave uncertainty," not "a mere possible doubt," "actual and substantial," "reasonable man can seriously entertain," not "an absolute or mathematical certainty," and "moral certainty." *Cage v. Louisiana*, 498 U.S. 39, 41 (1990). The Court struck down the instruction, but not because of its potential to confuse jurors. *Id.*

145. See, e.g., Norbet L. Kerr et al., *Guilt Beyond a Reasonable Doubt: Effects of Concept Definition and Assigned Decision Rule on the Judgments of Mock Jurors*, 34 J. PERSONALITY & SOC. PSYCHOL. 282, 293 (1976) (finding an average assessment of eighty-seven percent). A study of 134 jury-eligible Internet users found an average assessment of 84.5%. Simon et al., *Lay Assessments of Standards of Proof* (unpublished data).

146. The value of 90% (or 91%) is frequently said to be derived from Blackstone's error-ratio of 10:1. That inference is mistaken. To derive the standard of proof from the error ratio, one would need also to know the base rate of guilt among the prosecuted cases and the jury's diagnostic capabilities. See Michael L. DeKay, *The Difference Between Blackstone-Like Error Ratios and Probabilistic Standards of Proof*, 21 L. & SOC. INQUIRY 95, 103–10 (1996).

147. Greene & Dodge, *supra* note 67, at 75.

148. Irwin A. Horowitz & Laird C. Kirkpatrick, *A Concept in Search of a Definition: The Effects of Reasonable Doubt Instructions on Certainty of Guilt Standards and Jury Verdict*, 20 LAW & HUM. BEHAV. 655, 670–71 (1996).

149. Kramer & Koenig, *supra* note 41, at 414.

150. James R. P. Ogloff & V. Gordon Rose, *THE COMPREHENSION OF JUDICIAL INSTRUCTIONS*, in *PSYCHOLOGY AND LAW: AN EMPIRICAL PERSPECTIVE* 418 (Neil Brewer & Kipling D. Williams eds., 2005).

151. Some 40% of a sample of German respondents stated a threshold of 95–99%. Andreas Glöckner & Christoph Engel, *Can We Trust Intuitive Jurors? Standards of Proof and the Probative Value of Evidence in Coherence Based Reasoning*, J. EMPIRICAL LEGAL STUD. (forthcoming 2012) (manuscript at 4).

regarding the appropriate ratio of false convictions to false acquittals.¹⁵²

The limited experimental research testing the effect of the standard of proof on verdict decisions yields mixed results, but it appears that the standard does indeed exert its intended effect, even if only roughly and partly. Some studies have found that instructing jurors with the beyond a reasonable doubt standard reduced conviction rates relative to the “preponderance of the evidence” standard (as it should).¹⁵³ Another experiment found that the standard alone has no appreciable effect, but it reduces conviction rates when combined with the presumption of innocence instruction.¹⁵⁴ However, a field study of more than 300 actual trials conducted by the National Center for State Courts (NCSC) gives reason for pause. This study, observed the relationship between the strength of the evidence and the verdicts in each case. As would be expected, the juries voted to convict at high rates when the strength of the evidence was described as strong. However, more than half of the jurors voted to convict the defendant when they felt that the prosecution’s evidence was of medium strength, and, surprisingly, one in five juries voted to convict even when they felt that the evidence was weak. A similar pattern of judgments was obtained from the judges who sat on these cases.¹⁵⁵

152. While legal observers tend to follow the classic error ratio of 10:1 proposed by Blackstone, lay people appear to be far less sensitive to false convictions. Only 66% of respondents indicated that a false conviction is worse than a false acquittal, with 30% of the respondents maintaining the opposite. Insa Bechert & Markus Quandt, *ISSP Data Report: Attitudes Toward the Role of Government*, SOCIAL SCIENCE OPEN REPOSITORY, 21 (2010), <http://www.ssoar.info/ssoar/files/2011/1535/gs%206%20-%20issp%20data%20report.pdf>. In the *Jurors’ Beliefs Survey* respondents were provided with a continuous scale that listed different ratios of false convictions to false acquittals and asked to indicate their preferred error ratio. The mode response was 1:1, that is, that the errors are equally bad. The median response was: “It is better to acquit one guilty person than convict one innocent person.” Simon et al., *supra* note 145.

The high thresholds provide another reason to doubt that jurors rely on probabilistic assessments in deciding cases. If jurors reach verdicts based on their strength of belief in the evidence, the unrealistic numeric thresholds they report would not necessarily impede on their performance.

153. In one study, the heightened standard reduced the rate of convictions by deliberating participants from 42% to 26%, which was marginally significant. No difference was observed for non-deliberating jurors. Robert J. MacCoun & Norbert L. Kerr, *Asymmetric Influence in Mock Jury Deliberation: Jurors’ Bias for Leniency*, 54 J. PERSONALITY & SOC. PSYCHOL. 21, 29 (1988). Another study found a decrease from 65% to 48% for non-deliberating participants. To prevent a confound between the standard of proof and the presumption of innocence, the 293 jury-eligible participants were informed that this was an arbitration procedure, not a criminal trial. Simon et al., *supra* note 145.

154. Glöckner & Engel, *supra* note 151, at manuscript 6. These data compare the results of the conditions in which the procedure was described as an arbitration and the standard of proof was the preponderance of the evidence (Study 1, Treatments 3 and 4, and Study 2, Treatment 4), with the conditions described as a criminal procedure with the beyond a reasonable doubt standard (Study 1, Treatments 1 and 2, and Study 2, Treatment 1). The latter conditions yielded conviction rates that were less than half of the former conditions (18% v. 44%).

In an early experiment, Lawrence Wrightsman and Saul Kassin found that presenting jurors with the presumption of innocence and beyond a reasonable doubt standard reduced conviction rates from 56% to 35% (data for jurors who were given the instructions prior to the evidence). Saul M. Kassin & Lawrence S. Wrightsman, *On the Requirements of Proof: The Timing of Judicial Instruction and Mock Juror Verdicts*, 37 J. PERSONALITY & SOC. PSYCHOL. 1877, 1880 (1979). This finding should be qualified due to the low number of participants.

155. As one would expect, when the evidence favoring the prosecution was described as strong

Yet, even if the standard of proof did reduce conviction rates, it would not necessarily increase the accuracy of verdicts. The standard is merely a sorting mechanism, and is devoid of any diagnostic properties of its own. It relies on the fact finder's ability to correctly assess the accuracy of evidence in the particular case. If, as the research suggests, the fact finder's perception of the defendant's guilt does not correspond closely to the actual guilt, the standard can do little to differentiate guilty from innocent defendants. High standards are thus bound to result in the acquittal of both innocent and guilty defendants.

V

JURY DELIBERATION

The discussion thus far has focused mostly on the performance of individual jurors. Yet, criminal verdicts are cast collectively by groups, following a process of deliberation. Vesting the decisionmaking power in the hands of juries embodies the ideals of a representative and deliberative democracy, and confers legitimacy on the process.¹⁵⁶ The important question for our current purposes is whether juror deliberation contributes to the accuracy of the decisions made.

The field data pertaining to the effects of deliberation are very instructive. Post-trial interviews with real-life jurors reveal a very simple yet noteworthy finding: with very few exceptions, verdicts are determined straightforwardly by the distribution of votes on the first ballot. In total, in 94% of the cases the initial majorities won the day.¹⁵⁷ An early study by Kalven and Zeisel found that juries convicted defendants 94% of the time when more than six jurors first voted for conviction, and acquitted them 97% of the time when fewer than six jurors voted for conviction on the first ballot.¹⁵⁸ Similar findings were made in studies of 179 criminal trials conducted in Indiana,¹⁵⁹ and 43 in Kentucky.¹⁶⁰

(ratings of 6 and 7 on a 1–7 scale), the conviction rate was high: 98% of the juries convicted the defendants and 99% of the judges stated that they would have done the same. Surprisingly, 61% of the juries voted to convict and 78% of the judges would have done the same when the evidence was deemed to be of medium strength (ratings of 3, 4, 5). Even more startling is the fact that 17% of the juries voted to convict when they deemed the evidence to be weak (ratings of 1, 2), and 27% of the judges would have convicted on the basis of that weak evidence (calculated from data in Table 4 on page 186). Theodore Eisenberg et al., *Judge-Jury Agreement in Criminal Cases: A Partial Replication of Kalven and Zeisel's The American Jury*, 2 J. EMPIRICAL LEGAL STUD. 171, 188–89 (2005).

156. ABRAMSON, *supra* note 3, at 3–4; VIDMAR, *supra* note 3, at 347.

157. These data account for the 222 juries for which detailed data are available. These studies are summarized in Robert J. MacCoun & Norbert L. Kerr, *Is the Leniency Bias Really Dead? Misinterpreting Asymmetry Effects in Criminal Jury Deliberation* (under review 2012) (Table 4, coding assumption #2). The following analyses will follow the suggestion of Kerr and MacCoun to split the “undecided” votes on the first ballot evenly between the guilty and not-guilty columns.

158. These data pertain to the non-hanging juries. Sixteen of the trials ended with hung juries. In the ten cases where the first ballots were split evenly, five ended with a conviction and five with an acquittal. This study, comprising of 225 criminal cases conducted in Chicago and Brooklyn, was reported in HARRY KALVEN & HANS ZEISEL, *THE AMERICAN JURY*, CH. 38 (1966).

159. In this study, the rate of conviction was zero when there were four or fewer votes to convict on the first ballot, and was 100% when there were eight or more votes to convict. The conviction rate was around one half (fifty-seven percent) when the first ballot had a majority of no more than two (five to seven votes to convict). These data exclude the eleven hung juries, and were coded per Robert J.

Slightly more nuanced relationships between first ballot votes and final verdicts were found in the NCSC field study.¹⁶¹ The conclusion from these data is that juries behave democratically not in the nuanced deliberative sense of the term, but in the blunt majoritarian sense.¹⁶² Heroic feats by lone jurors, as depicted in the classic film *Twelve Angry Men*, did not occur even once in the thirty-four cases that began with a split of eleven votes to one on the first ballot.¹⁶³

Still, the overwhelming power of numeric majorities should not lead to the conclusion that deliberation has no desirable effect on the process. For one, deliberation serves the important function of identifying the median votes, thus eliminating much of the idiosyncratic bias of individual jury members. The question remains whether deliberation also improves the accuracy of the factual findings that underlie the verdicts. This is particularly important in cases where the initial faction splits are close to equal. The contribution of deliberation is important also to our understanding of what transpires prior to the first ballot vote.

MacCoun & Norbert L. Kerr, *Is the Leniency Bias Really Dead? Misinterpreting Asymmetry Effects in Criminal Jury Deliberation* (Table 4, coding assumption #2). These data underestimate the rate of conviction, as they pertain only to the votes on the most serious charge. Dennis J. Devine, et al., *Deliberation Quality: A Preliminary Examination in Criminal Juries*, 4 J. EMPIRICAL LEGAL STUD. 273, 293 (2007).

160. In this study, the rate of conviction was zero when there were four or fewer votes to convict on the first ballot, and was ninety-four percent when there were eight or more votes to convict. The conviction rate was sixty-six percent when the first ballot had a majority of no more than two (five to seven votes to convict). These data exclude the hung juries, and were coded per Kerr & MacCoun, *supra* note 159; Marla Sandys & Ronald C. Dillehay, *First-Ballot Votes, Predeliberation Dispositions, and Final Verdicts in Jury Trials*, 19 LAW & HUM. BEHAV. 175, 184–85 (1995).

161. In this study, the rate of conviction was 4% when there were four or fewer votes to convict on the first ballot, and was 94% when there were eight or more votes to convict. The conviction rate was 82% when the first ballot had a majority of no more than two (five to seven votes to convict). These data exclude the 25 hung juries, and were coded per Kerr & MacCoun. Paula L. Hannaford-Agor, et al., *Are Hung Juries a Problem?*, THE NATIONAL CENTER FOR STATE COURTS, 67 (2002), <http://contentdm.ncsconline.org/cgi-bin/showfile.exe?CISOROOT=/juries&CISOPTR=27>.

162. There is some debate among scholars about the existence of a skew towards acquittals following deliberation, a phenomenon known as “leniency bias,” see generally MacCoun & Kerr, *supra* note 153, recently labeled the “asymmetry effects in jury deliberation,” *id.* at 22. Compare the work of MacCoun and Kerr with Dennis J. Devine et al., *Jury Decision Making: 45 Years of Empirical Research on Deliberating Groups*, 7 PSYCHOL. PUB. POL’Y & LAW 622, Table 6 (2001). To the extent that the bias exists, it is overwhelmed by the powerful majoritarian dominance. The research is crippled by the lack of data in the important category of juries that were close to an even split on the first ballot (with factions of 5:7, 6:6, or 7:5). Of the 222 studies cases, only one jury had an initial split of five jurors favoring conviction (it voted not guilty), and only four juries had an initial split of 6:6 (three of these juries voted guilty).

It should be noted that Hastie, Penrod & Pennington (1983) found that juries were just as likely to shift towards severity as they are towards leniency (for juries governed by a unanimous decision rule). Mona Lynch & Craig Haney have tested the effects of deliberation in the context of death penalty sentencing, and found evidence of a “severity shift,” with deliberation increasing death sentences from 54% to 66%. Lynch & Haney, *Discrimination and Instructional Comprehension*, *supra* note 96, at 485. It should be kept in mind that leniency and severity biases are likely to have the same effect on guilty and innocent defendants, and thus do not enhance the diagnosticity of the adjudicative process. Kerr & MacCoun, *supra* note 143.

It is widely intuited that groups outperform their individual members. The underlying notion is the belief in collective wisdom, which posits that pooling knowledge and judgment produces the best that the group has to offer and discards the worst. The research, however, indicates that this belief does not always correspond with reality. The preponderance of the research suggests that judgment by groups cannot be said to be generally superior or inferior to the performance of their individual members.¹⁶⁴ On some tasks, groups do indeed outperform their members.¹⁶⁵ On other tasks, however, they perform comparably,¹⁶⁶ or fall short of their members.¹⁶⁷ The respective strengths and weaknesses depend on a host of contextual and group specific factors.¹⁶⁸ Crucially, the effect of deliberation on the group's decision will depend on the accuracy of the faction that wins the day. Groups are bound to reach correct conclusions when the prevailing members hold the correct views, but when they are wrong, deliberation is bound to promote error. Generally speaking, consensus within groups can be achieved by means of two forms of persuasion: informational persuasion and social influence, each of which will be discussed in turn.¹⁶⁹

164. See generally Norbert L. Kerr & R. Scott Tindale, *Group Performance and Decision Making*, 55 ANN. REV. PSYCHOL. 623 (2004).

165. For example, groups are found to outperform individuals in tasks of general knowledge and estimation of quantities. See Janet A. Sniezek & Rebecca A. Henry, *Accuracy and Confidence in Group Judgment*, 43 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 1, 20 (1989); Janet A. Sniezek & Rebecca A. Henry, *Revision, Weighting, and Commitment in Consensus Group Judgment*, 45 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 66, 72 (1990).

166. For illustration, a study that tested pairs of participants in solving math puzzles found that when both members were independently correct or independently wrong, joining forces did not affect their joint accuracy. The performance of mixed pairs was affected by the deliberation, in that the joint solution was determined mostly by the member who was more influential—typically, more confident—irrespective of his or her accuracy. Howard H. Johnson & James M. Torcivia, *Group and Individual Performance on a Single-Stage Task as a Function of Distribution of Individual Performance*, 3 J. EXPERIMENTAL SOC. PSYCHOL. 266, 272 (1967). In performing rule induction problems, groups perform as well as their best members, but only under optimal conditions. If the information or time is limited, groups fail to meet that standard. Patrick R. Laughlin et al., *Collective Versus Individual Induction: Recognition of Truth, Rejection of Error, and Collective Information Processing*, 61 J. PERSONALITY & SOC. PSYCHOL. 50, 62 (1991).

167. For example, groups have been found to be more biased than individuals in the search for information. Stefan Schulz-Hardt et al., *Biased Information Search in Group Decision Making*, 78 J. PERSONALITY & SOC. PSYCHOL. 655, 666–67 (2000). Brainstorming is an acute example of the discrepancy between the widespread belief in the superiority of groups and the empirical findings to the contrary. On the “illusion of group productivity,” see generally Bernard A. Nijstad, et al., *The Illusion of Group Productivity: A Reduction of Failures Explanation*, 36 EUROPEAN J. SOC. PSYCHOL. 31 (2006); Paul B. Pausus et al., *Perception of Performance in Group Brainstorming: The Illusion of Group Productivity*, 19 PERSONALITY & SOC. PSYCHOL. BULL. 78 (1993).

168. See, e.g., Norbet L. Kerr et al., *Bias in Jurors vs Bias in Juries: New Evidence from the SDS Perspective*, 80 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 70 (1999) [hereinafter Kerr et al., *Bias in Jurors*]; Norbet L. Kerr et al., *Bias in Judgment: Comparing Individuals and Groups*, 103 PSYCHOL. REV. 687 (1996).

169. Morton Deutsch & Harold B. Gerard, *A Study of Normative and Informational Social Influences Upon Individual Judgment*, 51 J. ABNORMAL & SOC. PSYCHOL. 629, 634–36 (1955). See generally Wendy Wood, *Motives and Modes of Processing in the Social Influence of Groups*, in DUAL-PROCESS THEORIES IN SOCIAL PSYCHOLOGY 547 (Shelly Chaiken & Yaacov Trope eds., 1999).

A. Informational Persuasion

Informational persuasion is performed by means of conveying information and arguments that lead the listener to alter her position. The research points to three ways in which the exchange of information among jurors could contribute to the accuracy of verdicts. One potential contribution of information persuasion is to improve the jurors' memory of the evidence presented at trial. As it is impossible to remember everything presented over the many hours or days of testimony, jurors enter the deliberation room with memorial accounts that are incomplete and somewhat idiosyncratic.¹⁷⁰ Group deliberation could prove beneficial either by enhancing memory completeness through the pooling of the members' memories or by increasing memory accuracy through the correction of errors by members with better recall. With respect to the latter, the research indicates that deliberation hardly affects the accuracy of the recalled evidence, primarily because most facts recalled by jurors are accurate.¹⁷¹ This is an encouraging finding. The issue of memory completeness, however, is more complicated. The research reveals that memories produced by collaborative recall are somewhat more complete than the recall of the average member, but less complete than the total non-redundant memories recalled by individual members.¹⁷² Studies testing group recall in simulated jury trials indicate that deliberation offers only a modest improvement to the completeness of jurors' memories.¹⁷³

A second way in which deliberation could conceivably improve the accuracy of the process is by enhancing individual jurors' comprehension of the jury instructions. A large jury simulation conducted by Phoebe Ellsworth found that deliberating jurors were no more knowledgeable than non-deliberating jurors, and their comprehension levels were no better than chance. Only half of the references to legal issues that were expressed during the deliberation were

170. Studies of simulated jurors find that people remember about two-thirds of the relevant details presented at trial. Mary E. Pritchard & Janice M. Keenan, *Does Jury Deliberation Really Improve Jurors' Memories?*, 16 *APPLIED COGNITIVE PSYCHOL.* 589, 595 (2002) [hereinafter Pritchard & Keenan, *Jury Deliberation*]; Mary E. Pritchard & Janice M. Keenan, *Memory Monitoring in Mock Jurors*, 5 *J. EXPERIMENTAL PSYCHOL.: APPLIED* 152, 158 (1999) [hereinafter Pritchard & Keenan, *Memory Monitoring*]. In another study, jurors recalled about sixty percent of the evidence presented. See HASTIE ET AL., *supra* note 37.

171. Phoebe C. Ellsworth, *Are Twelve Heads Better Than One?*, 52 *LAW & CONTEMP. PROBS.* 205, 224 (Autumn 1989). For similar findings, see Pritchard & Keenan, *supra* note 170; HASTIE ET AL., *supra* note 37.

172. This phenomenon is labeled "collaborative inhibition." Mary S. Weldon & Krystal D. Bellinger, *Collective Memory: Collaborative and Individual Processes in Remembering*, 23 *J. EXPERIMENTAL PSYCHOL.: LEARNING, MEMORY, AND COGNITION* 1160, 1165 (1997). Collaborative inhibition increases with group size, and is likely to be substantial in groups of twelve members. Barbara H. Basden et al., *Costs and Benefits of Collaborative Remembering*, 14 *APPLIED COGNITIVE PSYCHOL.* 497, 500 (2000). Groups have been found to have superior memory when tested with recognition tasks, which are less relevant to the jury setting. See, e.g., Verlin B. Hinsz, *Cognitive and Consensus Processes in Group Recognition Memory Performance*, 59 *J. PERSONALITY & SOC. PSYCHOL.* 705, 716 (1990).

173. For a review of the research, see Pritchard & Keenan, *Jury Deliberation*, *supra* note 170. Pritchard and Keenan found that group recall was better than recall by individuals by 3.4%. *Id.* at 600.

correct. When jurors were influenced to alter their understanding of the law, they were as likely to substitute accurate understandings with errors as they were to correct mistakes. Opinions expressed most forcefully tended to prevail, regardless of their accuracy.¹⁷⁴ A Canadian study found that in the few instances in which jurors discussed the judicial instructions, only sixty-one percent of the statements were legally correct.¹⁷⁵ Poor comprehension was also found in studies of death sentencing instructions. The Missouri juror studies found that deliberation had a positive, though weak, effect on comprehension of death penalty instructions in one study and no effect at all in another.¹⁷⁶ A study of California jurors found low comprehension of death penalty instructions following deliberation.¹⁷⁷ Another study found that deliberations improved comprehension of a death-sentencing instruction that was correctly understood by two-thirds of the respondents, but not of two other instructions that were correctly understood only by half or fewer.¹⁷⁸ Deliberation could promote accuracy also by improving jurors' ability to comply with curative instructions. Here, too, deliberation yields mixed effects at best. In some instances, deliberation mitigated the effects of extra-legal evidence.¹⁷⁹ Yet in others, it made no difference,¹⁸⁰ and even exacerbated the influence of the impermissible evidence.¹⁸¹

174. See Ellsworth, *supra* note 171.

175. See Ogloff & Rose, *supra* note 150 (describing a study of juries and jury instructions in Canada).

176. Wiener et al., *Guided Jury Discretion*, *supra* note 90, at 571.

177. Mona Lynch & Craig Haney, *Capital Jury Deliberation: Effects on Death Sentencing, Comprehension, and Discrimination*, 33 LAW & HUM. BEHAV. 481, 489 (2009).

178. Diamond & Levi, *supra* note 91.

179. One study found that deliberation reduced guilt ratings following a judge's admonition to disregard inadmissible evidence (5.7 to 4.4 on a scale ranging from 1 to 9). Jeffrey Kerwin & David R. Shaffer, *Mock Jurors Versus Mock Juries: The Role of Deliberation in Reactions to Inadmissible Testimony*, 20 PERSONALITY & SOC. PSYCHOL. BULL. 153, 159 (1994). Another study found that deliberation halved conviction rates given by admonished jurors. Kamala London & Narina Nunez, *The Effect of Jury Deliberations on Jurors' Propensity to Disregard Inadmissible Evidence*, 85 J. APPLIED PSYCHOL. 932, 934 (2000).

180. In one study, deliberation was found to reduce admonished jurors' conviction rates by a mere five percent. Valerie P. Hans & Anthony N. Doob, *Section 12 of the Canada Evidence Act and the Deliberations of Simulated Juries*, 18 CRIM. L.Q. 235, 243 (1976). Deliberation did not make a difference in a study that exposed jurors to pre-trial publicity in a murder case. Christine Ruva et al., *Effects of Pre-Trial Publicity and Jury Deliberation on Juror Bias and Source Memory Errors*, 21 APPLIED COGNITIVE PSYCHOL. 45, 58 (2007). Similarly, deliberation and admonitions made no effect on reducing the hindsight bias in an automobile negligence trial. Alison C. Smith & Edith Greene, *Conduct and Its Consequences: Attempts at Debiasing Jury Judgments*, 29 LAW & HUM. BEHAV. 505, 524 (2005).

181. In a study simulating a child molestation charge, jurors presented with inadmissible pre-trial publicity were considerably more likely to convict the defendant following deliberation than they were prior to it (29% v. 11%). This result was obtained where the evidence was ambiguous, which is the most forensically relevant type of situation. Kerr et al., *Bias in Juries*, *supra* note 168, at 79. A large study found that pre-trial publicity affected judgments only following deliberation, increasing convictions from 6% to 20%. Geoffrey P. Kramer et al., *Pretrial Publicity, Judicial Remedies, and Jury Bias*, 14 LAW & HUM. BEHAV. 409, 424 (1990). A limited number of deliberation studies included in the meta-analysis of pre-trial publicity show that deliberation served no appreciable curative effect. Nancy

Third, jury deliberation could promote verdict accuracy by means of improving jurors' ability to draw correct inferences from the evidence. The limited available research reveals only modest benefits in this regard. Social psychological studies find that the effect of group deliberation on social judgments is ameliorative at times, but not always so.¹⁸² Studies on deceit detection found that groups were no more accurate than individuals, though group members were more confident about their (actually poor) performance.¹⁸³

B. Social Influence

Informational persuasion is most likely to succeed when there is a demonstrably correct and evident conclusion,¹⁸⁴ which is rare in difficult criminal cases. In the absence of such favorable conditions, the requisite unanimity will rely more heavily on social influence.¹⁸⁵ Indeed, a distinctive feature of jury deliberation is that unanimity is often achieved through social pressure. The prevalence of social pressures is not altogether surprising given that jury deliberation is a non-structured, non-transparent, one-shot process among strangers, driven by a strong mandate to reach unanimity in cases that are often divisive.¹⁸⁶ The study by the National Center for State Courts (NCSC) included interviews with more than 3,000 jurors. These interviews indicate that the factors that enabled juries to reach a verdict (as opposed to resulting in hung juries) were predominantly social, not evidential. Reaching consensus was found to be explainable by factors such as the perceived open-mindedness and reasonableness of the other jurors, the domination of the deliberation by one or two jurors, and the degree of conflict on the jury.¹⁸⁷ Moreover, this study

Mehrkens Steblay et al., *The Effects of Pretrial Publicity on Juror Verdicts: A Meta-Analytical Review*, 23 LAW & HUM. BEHAV. 219, 230 (1999).

182. Compare Edward F. Wright & Gary L. Wells, *Does Group Discussion Attenuate the Dispositional Bias?*, 15 J. APPLIED SOCIAL PSYCHOL. 531, 542 (1985), with Gwen M. Wittenbaum & Garold Stasser, *The Role of Prior Expectancy and Group Discussion in the Attribution of Attitudes*, 31 J. EXPERIMENTAL SOCIAL PSYCHOL. 82, 99 (1995).

183. In one study, deliberation increased the accuracy rate from 56% to 60%. Groups were found to be more suspicious and were more inclined to judge the statements as deceitful. Mark G. Frank et al., *Individual and Small Group Accuracy in Judging Truthful and Deceptive Communication*, 13 GROUP DECISION & NEGOTIATION 45, 50 (2004). In a second study, deliberation resulted in a non-significant increase in the rate of accuracy from 51.5% to 53%. Ernest S. Park et al., *Group and Individual Accuracy in Deception Detection*, 19 COMMUNICATION RESEARCH REPORTS 99, 103 (2002). The weak contribution of deliberation is likely due to the low level of individual performance in the task of deceit detection.

184. Daniel Gigone & Reid Hastie, *Proper Analysis of the Accuracy of Group Judgments*, 121 PSYCHOL. BULL. 149, 163 (1997).

185. For classic studies on conformity, see Solomon E. Asch, *Studies of Independence and Conformity: I. A Minority of One Against a Unanimous Majority*, PSYCHOLOGICAL MONOGRAPHS 70, 68-70 (1956).

186. Jury deliberation lacks discursive conventions that resemble the idealized forms of deliberative democracy. For example, Jürgen Habermas posits a discourse that aspires to be "public and inclusive, to grant equal communication rights for participants, to require sincerity and to diffuse any kind of force other than the forceless force of the better argument." Jürgen Habermas, *Between Facts and Norms: An Author's Reflections*, 76 DENV. U. L. REV. 937, 940 (1992).

187. These factors were combined to create a Group Dynamics Index, which predicted whether

revealed that the putative unanimity did not reflect genuine consensus. In almost half of the juries that reached a verdict, at least one juror voted with the majority against his or her personal judgment.¹⁸⁸ Acquiescence of minority jurors to the will of the majority has also been observed in experimental studies.¹⁸⁹ Jurors who change their mind tend to be the least confident, not the least accurate.¹⁹⁰ The strength of the social pressure tends to increase as the deliberation progresses,¹⁹¹ and is exacerbated considerably by judicial instructions designed to undo locked juries, also known as *dynamite* charges.¹⁹² It follows that jury deliberation is potentially beneficial when the influential jurors have reached valid conclusions, probably detrimental when they are wrong, and of mixed results when they are evenly split.

Data from laboratory and field studies suggest also that social influence is affected by the race and gender of the jurors, especially of the forepersons. The study of the 179 Indiana juries revealed that juries convicted more frequently when the foreperson voted to convict, especially when the foreperson was white and male.¹⁹³ Likewise, studies conducted by the Capital Jury Project have observed the influence of white male jurors on the likelihood of death sentences.¹⁹⁴ This effect was observed also in a simulated deliberation in a death sentencing study.¹⁹⁵

Two additional aspects of jury deliberation are of note. First, there is preliminary evidence that the quality of the deliberation is affected by the jury's deliberation style, specifically whether it commences with a discussion of the

juries hung or reached a verdict. The index had a scale reliability of Cronbach's $\alpha = .91$. Hannaford-Agor et al., *supra* note 161, at 69.

188. Jurors were asked, "If it were entirely up to you as a one-person jury, what would your verdict have been in this case?" Nicole L. Waters & Valerie P. Hans, *A Jury of One: Opinion Formulation, Conformity, and Dissent on Juries*, 6 J. EMPIRICAL LEGAL STUD. 513, 537 (2009). There is reason to believe that the true number of silent dissenters is even higher given that people are not inclined to admit to having defied their own conscience to appease others.

189. See Sarah Tanford & Steven Penrod, *Jury Deliberations: Discussion Content and Influence Processes in Jury Decision Making*, 16 J. APPLIED SOCIAL PSYCHOL. 322, 330-31 (1986); Park et al., *supra* note 183, at 103.

190. See Ellsworth, *supra* note 171, at 220; Pritchard & Keenan, *Memory Monitoring*, *supra* note 170, at 165. These findings of social influence are consistent with recent research on "social vigilantism," an individual difference construct that corresponds to people's tendency to impress and propagate their beliefs onto others and to resist being persuaded in return. See generally Donald A. Saucier & Russell J. Webster, *Social Vigilantism: Measuring Individual Differences in Belief Superiority and Resistance to Persuasion*, 36 PERSONALITY & SOC. PSYCHOL. BULL. 19 (2009).

191. K. L. Hansen et al., *Temporal Patterns of Normative, Informational, and Procedural-Legal Discussion in Jury Deliberations*, 14 BASIC & APPLIED SOC. PSYCHOL. 33, 40 (1993).

192. See generally Vicki L. Smith & Saul M. Kassin, *Effects of the Dynamite Charge on the Deliberations of Deadlocked Mock Juries*, 17 LAW & HUM. BEHAV. 625 (1993). Saul M. Kassin et al., *The Dynamite Charge: Effects on the Perceptions and Deliberation Behavior of Mock Juries*, 14 LAW & HUM. BEHAV. 537 (1990).

193. Devine et al., *supra* note 159, at 283-87.

194. On the white male dominance phenomenon, see William J. Bowers et al., *Death Sentencing in Black and White: An Empirical Analysis of the Role of Juror Race and Jury Racial Composition*, 3 U. PA. J. CONST. L. 171, 195 (2001).

195. Lynch & Haney, *supra* note 177, at 485.

evidence (evidence-driven deliberation) or with a vote (verdict-driven deliberation). The latter style tends to be more argumentative, as jurors advocate for their espoused verdict. The resulting evaluation of the evidence tends to be more disjointed, the connections between the law and the facts are less developed, and the debate is less rigorous and congenial.¹⁹⁶ A large field study of real-life juries found that the juries that took a vote within the first ten minutes of the deliberation were more likely to hang than juries that deliberated for longer before their first vote.¹⁹⁷ Given the freedom jurors enjoy in structuring their deliberations in the real world, it is likely that different juries follow different procedures.

Second, it is important to acknowledge that jury deliberation has a polarizing effect on the decision. A large body of research shows that discussing and deciding issues in group settings tends to strengthen people's positions in support of the chosen resolution.¹⁹⁸ Typically, this polarization results also in higher levels of confidence.¹⁹⁹ Studies of simulated juries find that deliberation increases jurors' confidence in their memory of the trial evidence (both accurate and inaccurate memories),²⁰⁰ in their monitoring of the sources of their memory,²⁰¹ and in their judgments of the strength of the evidence.²⁰² As mentioned, group deliberation has also been found to inflate members' confidence in judgments of deceit.²⁰³

This polarization is bound to have two practical effects. First, polarizing a fractured jury is likely to deepen the rift between the factions, and thus weaken the prospect of informational persuasion. This effect, in turn, is bound to increase the use of social pressure in order to achieve putative unanimity. Second, jury polarization is likely to drive jurors to feel more strongly about their initial position. Group polarization has the potential to alter the verdict of a juror who is moderately leaning towards conviction. The polarization of that inclination and the concomitant boost in confidence can drive her decision over the threshold of proof and result in a decision to convict.²⁰⁴

196. HASTIE ET AL., *supra* note 37.

197. Hannaford-Agor et al., *supra* note 161.

198. For reviews, see Daniel J. Isenberg, *Group Polarization: A Critical Review and Meta-Analysis*, 50 J. PERSONALITY & SOC. PSYCHOL. 1141, 1142 (1986); CASS R. SUNSTEIN, *WHY GROUPS GO TO EXTREMES* (2008).

199. Sniezek & Henry, *supra* note 165, at 79. Confidence inflation is strongest in tasks that lack demonstrably correct conclusions. Paul Zarnoth & Janet A. Sniezek, *The Social Influence of Confidence in Group Decision Making*, 33 J. EXPERIMENTAL SOC. PSYCHOL. 345, 361 (1997).

200. Pritchard & Keenan, *Jury Deliberation*, *supra* note 170, at 597.

201. Ruva et al., *supra* note 180, at 48.

202. Martin F. Kaplan & Lynn E. Miller, *Reducing the Effects of Juror Bias*, 36 J. PERSONALITY & SOC. PSYCHOL. 1443, 1448-49 (1978).

203. One study found that deliberation increased the confidence of group members from 81% to 89%. Park et al., *supra* note 183, at 53. Another study found that the confidence of members of deliberating groups was higher than that of individuals. Frank et al., *supra* note 183, at 103.

204. The effect of polarization on jurors who are inclined to acquit is unlikely to have an effect on the juror's vote. Due to the asymmetric standard of proof, a juror leaning towards acquittal should vote to acquit regardless of the strength of that leaning. For a similar, intrapersonal polarization borne by

VI

APPELLATE AND POST-CONVICTION REVIEW

Another legal safeguard that is said to promote the integrity of judicial outcomes is the subjection of jury verdicts to post-trial review proceedings. The two common forms of review are the direct appeal process that runs through higher state courts,²⁰⁵ and the collateral post-conviction proceedings, notably habeas corpus, that are conducted in federal and state courts. Both processes are ultimately reviewable by the Supreme Court. Appellate and post-conviction proceedings provide the opportunity to examine trial court verdicts from a more detached and authoritative vantage point, and thus have the potential to increase the diagnosticity of the adjudicative process. Indeed, this review is generally perceived to be an important guarantee of the integrity of criminal verdicts.²⁰⁶

Yet the access to review by appellate and habeas courts is limited by a range of procedural barriers. Access to reviewing courts is limited by intricate statutory and judicially created procedural conditions, which include filing deadlines, contemporaneous objection at trial, narrow categories of cognizable claims, and exhaustion of claims.²⁰⁷ Meeting these conditions is a tall order, especially given that the majority of inmates conduct their post-conviction affairs without the benefit of legal counsel. Indeed, a recent study found that federal habeas courts award relief in only 0.4% of non-capital cases they review,²⁰⁸ and the majority of these reversals were based on legal, not factual, errors. Appellate and post-conviction courts are strongly averse to intervene in factual determinations made at the trial level.²⁰⁹ When they do extend themselves to examine the factual foundation of the verdict, they rely on the record developed by the trial court and examine the evidence in the light most favorable to the prosecution,²¹⁰ which naturally tends to comport with the lower court's findings. These courts also apply high thresholds for intervention,²¹¹ and

the "coherence effect," see SIMON, *supra* note 6, at chs. 2, 6.

205. Convictions for federal crimes are appealed to federal courts of appeal.

206. For example, in the context of capital sentencing, the Supreme Court characterizes the mandatory review by state supreme courts and its own subsequent review as one of the measures that "ensure a degree of care . . . that can be described only as unique." *McCleskey v. Kemp*, 481 U.S. 279, 313 (1987).

207. See, e.g., 28 U.S.C. §§ 2241–2266 (2006); RANDY HERTZ & JAMES S. LIEBMAN, *FEDERAL HABEAS CORPUS PRACTICE AND PROCEDURE*, 1707–1898 (5th ed. 2005).

208. The study examined almost 2,400 habeas corpus proceedings conducted in 2003 and 2004. Among the 267 proceedings in capital cases, some twelve percent of inmates received relief, but only one third of them pertained to the underlying conviction. NANCY J. KING ET AL., U.S. DEP'T OF JUSTICE, NAT'L INST. OF JUSTICE, *FINAL TECHNICAL REPORT: HABEAS LITIGATION IN THE U.S. DISTRICT COURTS* (2007), available at <https://www.ncjrs.gov/pdffiles1/nij/grants/219559.pdf>.

209. In habeas proceedings, this deference is imposed by statute. Title 28, section 2254(d) of the United States Code requires courts to defer to reasonable factual determinations of state courts. As the Court has stated, "Federal courts are not forums in which to relitigate state trials." *Barefoot v. Estelle*, 463 U.S. 880, 887 (1983); see also *Sumner v. Mata*, 449 U.S. 539, 544 (1981).

210. *Jackson v. Virginia*, 443 U.S. 307, 318–19 (1979).

211. *Hernandez v. New York*, 500 U.S. 352, 369 (1991); see also *LAFAVE ET AL.*, *CRIMINAL*

entertain only a “sharply limited” review of claims challenging the sufficiency of evidence.²¹² For example, the Louisiana Supreme Court has ruled that the evidence required for post-conviction relief must be new, material, noncumulative, and conclusive; it must meet “an extraordinarily high standard;” and it must “undermine the prosecution’s entire case.”²¹³ Finally, cases with lingering doubts are often punted onto the lap of the executive branch,²¹⁴ for whom the political costs of freeing convicted inmates—even ones who are most likely innocent—is particularly high.²¹⁵

The meager scope for meaningful review of factual determinations is manifested by courts’ treatment of the DNA exonerees prior to the discovery of the exculpatory DNA evidence. Brandon Garrett has analyzed the appellate and post-conviction review that preceded the exoneration in the first 250 DNA exoneration cases. Of the 165 exoneration cases with written decisions, 69 convicted inmates raised claims challenging the sufficiency of the prosecution’s evidence, yet only one of these challenges succeeded, only to be reversed later on. Of the 70 misidentified people who challenged the integrity of the identification procedures, just five succeeded in undermining the reliability of the procedures,²¹⁶ and just one of the thirteen confessors who challenged the admissibility of their confessions was granted relief.²¹⁷ Of the thirty-eight DNA exonerees who petitioned the Supreme Court, just one was granted review, only to be denied relief. Tellingly, the Court turned down the inmate’s request to obtain the very piece of evidence that ultimately proved his innocence.²¹⁸ Overall, the reviewing courts tended not to reveal any qualms about the underlying convictions. In about one-half of the cases, the courts made matter-of-fact references to the inculpatory evidence, occasionally labeling it

PROCEDURE: WEST’S CRIMINAL PRACTICE SERIES 97–98 (3d ed. 2007). A commonly used standard is whether “any rational trier of fact could have found the essential elements of the crime beyond a reasonable doubt.” *Jackson*, 443 U.S. at 318–19.

212. *Wright v. West*, 505 U.S. 277, 296 (1992). As pointed out by the late William Stuntz, the preeminence of proceduralism results in a diversion of defense claims from factual to procedural issues. See William J. Stuntz, *The Uneasy Relationship Between Criminal Procedure and Criminal Justice*, 107 YALE L.J. 1, 1–76, 37–45 (1997). Ironically, while defense attorneys are all but precluded from challenging the factual foundation of guilty verdicts, they are regularly derided for insisting on procedural exactness, or “technicalities.”

213. See *State v. Conway*, 816 So. 2d. 290, 291 (La. 2002).

214. “Clemency is deeply rooted in our Anglo-American tradition of law, and is the historic remedy for preventing miscarriages of justice where judicial process has been exhausted.” *Herrera v. Collins*, 506 U.S. 390, 411–12 (1993).

215. For illustrations of gubernatorial reticence, see the cases of Earl Washington and “The Norfolk Four.” *Earl Washington*, THE INNOCENCE PROJECT, http://www.innocenceproject.org/Content/Earl_Washington.php (last visited 12/1/2011). See also TOM WELLS & RICHARD A. LEO, *THE WRONG GUYS: MURDER, FALSE CONFESSIONS, AND THE NORFOLK FOUR* (2008).

216. GARRETT, *supra* note 4, at Ch. 7.

217. *Id.* One reversal was awarded for a claim of ineffective assistance of counsel, based on the lawyer’s failure to challenge the confession at trial. Brandon L. Garrett, *The Substance of False Confessions*, 62 STAN. L. REV. 1051, 1107 (2010).

218. *Arizona v. Youngblood*, 488 U.S. 51 (1988); see also *Larry Youngblood*, THE INNOCENCE PROJECT, <http://www.innocenceproject.org/Content/303.php> (last visited 12/1/2011).

“overwhelming.”²¹⁹ It is also notable that the DNA exonerees were no more likely to receive relief than a matching group of similarly situated inmates whose guilt was not refuted by DNA or by any other exonerating evidence (9% and 10% of non-capital cases, respectively).²²⁰

It is troubling to note that many of the innocent convicts did not even challenge the faulty evidence that led to their convictions. Of those falsely identified by eyewitnesses, only one contested the identification procedures, and fewer than two-thirds of those convicted on the basis of a confession challenged the legality of their interrogations.²²¹ In fact, DNA exonerees did not challenge their convictions any more than did the inmates in the matching group (who are presumably guilty).²²² Whether the impassivity of these innocent people was due to a lack of resources, their mistrust of the system, or any other reason, the reality is that they did not benefit from the legal avenues that were ostensibly open to them. In short, not unlike the other legal safeguards discussed above, the post-trial review proceedings provide sparse assistance in distinguishing between truly guilty and factually innocent convicted inmates.

VII

CONCLUSION AND RECOMMENDATIONS FOR REFORM

The research examined in *Limited Diagnosticity* revealed limitations in the reliability of the evidence typically presented in criminal trials and in the capability of fact finders to fulfill their diagnostic task. This article has examined the extent to which the fact finding task can be improved by the legal mechanisms put in place to promote the diagnosticity of the process. This examination paints a mixed, and overall troubling, picture. To a certain extent, these mechanisms do indeed enhance diagnosticity, but under a wide range of circumstances, they are ineffective, and even detrimental.

Cross-examination can help in deterring perjury and stumping deceitful witnesses, but it is ill equipped to prevent honest mistakes. It is also of limited effectiveness, and can even backfire, especially for litigants who operate with an informational disadvantage, as defense attorneys often do. Cross-examination can also hurt the diagnosticity of the process by inducing biased testimony and flummoxing truthful witnesses. Jury instructions are widely misunderstood and misapplied, and are oftentimes superseded by the jurors' personal beliefs about the law and their conceptions of justice. Jury instructions are particularly ineffective when they call for mental operations that are unnatural or unfamiliar. In particular, the poor performance in following instructions casts serious doubts over the viability of the “guided discretion” regimes by which death penalties are meted out.

219. Garrett, *supra* note 4, at ch. 7

220. Each one of the DNA exonerations was matched with a randomly selected conviction for the same crime, from the same state, decided in the same year. *Id.*

221. *Id.*

222. *Id.*

The research also indicates that people are not well equipped to assess their own objectivity; thus, relying on jurors' assurances of their impartiality amounts to little more than a hollow gesture. The prosecution's heightened burdens, namely the presumption of innocence and the reasonable doubt threshold, appear to reign in verdicts to some extent, but their overall effects are weaker and murkier than generally believed. Jury deliberation influences verdict choices in only a narrow range of cases, and its effects on accuracy are mixed. Finally, the review of factual findings by appellate and post-conviction courts is severely curbed by a slew of procedural restrictions, and by a distinct proclivity to shun factual determinations. As a practical matter, findings made at trial are impervious to reexamination, even in the face of new compelling evidence to the contrary.

Together, these two articles lead to the conclusion that the adjudicatory process has only limited capabilities in distinguishing between accurate and inaccurate evidence. In light of the high expectations from the adjudicatory process,²²³ it is better characterized as "pseudo-diagnostic." To a large degree, criminal verdicts are determined at the investigative phase, with the trial serving primarily as a ritual that delivers more symbolic than real value.²²⁴ This conclusion calls for some specific reforms:

- (1) To make the adversarial trial more diagnostic, the informational asymmetry between the prosecution and the defense should be minimized.
- (2) Jury instructions ought to be drafted in plain language and provided in writing. Juror instructions should include an explicit debunking of common misconceptions and, where possible, concrete examples.
- (3) When trial judges are asked to clarify the instructions, they should provide jurors with helpful explanations rather than merely repeat the (patently unclear) instructions.
- (4) Given the ineffectiveness of curative jury instructions, trial judges ought to be more assertive in limiting the exposure of jurors to extra-evidential information and in preempting any other situations that might require resorting to such instructions.

223. High expectations of accuracy were expressed by respondents in the *Jurors' Beliefs Survey*. In response to a question about the acceptable rate of wrongful convictions, the median response was two out of one thousand convictions, while the mode response was zero. Simon et al., *supra* note 124. A somewhat higher tolerance of error was observed in a smaller survey of 133 college students. These respondents stated that an acceptable rate of wrongful convictions was 5%, and of wrongful acquittals 8%. Hal R. Arkes & Barbara A. Mellers, *Do Juries Meet Our Expectations?*, 26 *LAW & HUM. BEHAV.* 625, 631 (2002).

224. Given the historical development of the common law's criminal justice process, its limited diagnosticity is not altogether surprising. As described by John Langbein, the English criminal process evolved piecemeal, as a series of ad hoc tactical measures intended to balance out the advantages of the opposing adversaries and to circumvent the disbursement of punishments that were discordant with the prevailing public sentiment. These historical developments transpired with little concern for the system's capacity or propensity to ascertain truth. LANGBEIN, *supra* note 12, at 306-36.

- (5) The Supreme Court ought to acknowledge that the “guided discretion” doctrine fails to eliminate the prejudicial administration of the death penalty, and thus violates the Eighth Amendment’s prohibition on cruel and unusual punishment.
- (6) Trial judges should be skeptical of jurors’ assurances of impartiality and should thus apply greater scrutiny when empanelling jurors.
- (7) Jurors should be encouraged to thoroughly deliberate prior to taking the first vote.
- (8) Appellate and *habeas* courts should be more amenable to taking a fresh look at the evidence upon a substantial showing of innocence, and should apply less stringent criteria for assessing convictions.

Yet more sweeping and profound reforms are needed to promote factual accuracy. These reforms require, foremost, a heightened sensitivity towards the integrity of the evidence, on which the entire process relies. As discussed in detail in *In Doubt*, the primary and antecedent causes for faulty evidence lie in the police investigation, and the underlying errors are frequently induced by the investigative procedures used. Thus, the criminal investigation both warrants improvement and is potentially amenable to it. This makes it a promising venue for reform.²²⁵ Hence, the criminal law debate has much to gain by devoting less attention to the courtroom and more to the police station, and by looking beyond constitutional protections and procedural rights towards the actual quality of the evidence that underlies the entire endeavor. Although implementing reform across the thousands of police departments is no easy feat, it is encouraging to note that a number of states and jurisdictions have already made commendable progress. Notable in this regard are North Carolina, Illinois, Wisconsin, and New Jersey, states which seem to have responded to their record of wrongful convictions with critical introspection rather than with denial and self-justification.²²⁶

Many of these reforms can be implemented quite straightforwardly by law enforcement officials and policymakers at almost any level, from chiefs of local police departments, through district attorneys, and state and federal legislatures. Ideally, the effort would also be championed by the judiciary, especially the Supreme Court, which serves as the final arbiter on people’s rights and is the only institution that has jurisdiction over all law enforcement agencies and state courts. Importantly, the Court should abandon its devout and unflinching protection of the procedural rights, and start caring closely about

225. Any effort to reform of criminal investigations must include substantial revision of the collection of forensic evidence. See generally NATIONAL RESEARCH COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD (2009).

226. Reforms of lineup procedures have been put into effect in some ten states, including North Carolina, New Jersey, Rhode Island, Vermont, and Delaware. Mandatory taping of interrogations has been implemented North Carolina, Illinois, Wisconsin, and more than another dozen states. See *Reforms by State*, THE INNOCENCE PROJECT, <http://www.innocenceproject.org/news/LawView5.php> (last visited Dec. 1, 2012).

the inadequate investigative procedures that it condones and the reliability of the evidence they produce. Thus, the Court should endorse a robust version of the “freestanding” conception of due process,²²⁷ which should be guided by the objective of maximizing accuracy.²²⁸ In plain language, the Bill of Right’s Due Process clauses should be understood as providing criminal suspects and defendants the most basic and genuine form of due process, designed to ensure that their liberty and life will be deprived only following a process that minimizes the risk of mistakes to an unavoidable minimum. It must be acknowledged that the Court does not appear to be heading in this direction.²²⁹

One obvious way to enhance the accuracy of the evidence is to ensure that investigations are conducted with scrupulous adherence to “best-practice procedures.” Naturally, investigations that follow best practices are bound to produce more accurate evidence. Recommendations for reform include better management of the investigative process, and reforms of lineup procedures, interviews with cooperative witnesses, and interrogations of suspects.²³⁰ A second promising way to enhance the integrity of the evidence is to make criminal investigations transparent. This can be achieved by creating an electronic record of all investigative procedures and make the record available to all parties. The recording should capture all encounters with witnesses and suspects, including all lineups, interviews, and interrogations.

In *In Doubt*, I offer a detailed discussion of the advantages of the proposed reforms and of the hurdles they face. In brief, the investigative record will effectively freeze the witness’ statements at their *raw* state, and thereby minimize the effects of memory decay, contamination, and any biases or distortions borne by the investigative and pre-trial processes. This should bind witnesses to their original statements, and also reduce the pressure applied on them to alter their testimony. The recording of investigations is expected also to have an ameliorative effect on the investigative process itself, primarily by increasing investigators’ sense of accountability for the way they conduct their investigations. Transparency would help ensure that investigators adhere to best practices by providing law enforcement agencies with a tool for training,

227. See generally Jerold H. Israel, *Free-standing Due Process and criminal procedure: The Supreme Court’s Search for Interpretive Guidelines*, 45 ST. LOUIS U. L.J. 303 (2001); Andrew E. Tazlitz, *What Remains of Reliability: Hearsay and Freestanding Due Process After Crawford v. Washington*, CRIM. JUSTICE, Summer 2005, at 39.

228. DONLAD A. DRIPPS, ABOUT GUILT AND INNOCENCE: THE ORIGINS, DEVELOPMENT, AND FUTURE OF CONSTITUTIONAL CRIMINAL PROCEDURE (2003)

229. At the oral arguments in *Perry v. New Hampshire* (No. 10-8974, November 2, 2011), the justices seemed resolutely averse to the concept of freestanding due process rights. See Dahlia Lithwick, *See No Evil: Eyewitness Testimony May be Unreliable, but the Supreme Court Doesn’t Want to be the One to Say So*, SLATE (Nov. 02, 2011), http://www.slate.com/articles/news_and_politics/supreme_court_dispatches/2011/11/perry_v_new_hampshire_the_supreme_court_looks_at_eyewitness_evid.single.html. The case dealt with the admissibility of an eyewitness identification that was obtained from a suggestive lineup, which is the most notable application of a freestanding due process right. See *Manson v. Brathwaite*, 432 U.S. 98 (1977).

230. See Simon, *supra* note 6, at chs. 2–5

oversight, and quality assurance, and it should also help deter police misconduct. The record is also bound to serve as an informational tool by capturing forensic details that would otherwise not be collected.

Making the evidence more accurate and transparent is bound to make an appreciable impact throughout the legal proceedings, starting with the widespread practice of plea bargaining. Prosecutors should offer plea deals that are more fair and justified, and defendants should be better able to assess their situation before signing off on long terms of imprisonment. More accurate and transparent testimony can be expected to reduce the inclination to sort out murky facts through the costly, cumbersome, and imprecise process of litigation. When cases go to trial, the heightened quality of evidence is expected to reduce the distrust between the adversarial parties, and soften the contentiousness of the process. The range of plausible claims will be curbed, with the effect of narrowing the opportunities for unjust prosecutions and frivolous defenses.²³¹ Prosecutors would be in a position to pursue strong cases more forcefully, and defense attorneys would be better equipped to defend innocent defendants and to pursue valid claims of their clients. Freeing police detectives from excessive court proceedings, hostile cross examinations, and swearing contests should reduce the adversarial pressures they encounter and enable them to devote more effort to solving crimes.

Naturally, fact finders will be in a better position to determine the facts when presented with more accurate accounts of the criminal event. Fact finders should also benefit from gaining better access to the investigative procedures used to elicit the testimony. In particular, greater transparency should help fact finders determine whether the testimony might have been swayed or otherwise biased by the investigation itself.²³² Making plea negotiations and trials more accurate and less contentious should also result in fewer appeals, habeas proceedings, civil suits, and damage payouts.

The advantages conferred by recording investigations punctuate the realization that the methods that were available for the presentation and preservation of evidence in the eighteenth century in England are not necessarily the best ways of doing so today.²³³ The deterioration and

231. Indeed, law enforcement agencies that record interrogations report that the number of motions to suppress confessions has been reduced dramatically, and in some jurisdictions, has been eliminated altogether. Thomas P. Sullivan et al., *The Case for Recording Police Interrogations*, 34 LITIG. 30 (2008).

232. It must, however, be acknowledged that the creation of the electronic record itself can introduce bias. Factors such as camera perspective, framing, lighting, and editing can shape the narrative impact of the film. For example, research on taping interrogations has found that focusing the video camera exclusively on the suspect inflates the perceived voluntariness of the suspect's statements, which results in unwarranted trust in coerced confessions. See G. Daniel Lassiter et al., *Videotaping Custodial Interrogations: Toward a Scientifically Based Policy*, in POLICE INTERROGATIONS AND FALSE CONFESSIONS: CURRENT RESEARCH, PRACTICE, AND POLICY RECOMMENDATIONS 143–60 (G. Daniel Lassiter et al. eds., 2010); 14 LEGAL & CRIM. PSYCHOL. 157, 167 (2008). For a review and recommendations to minimize this phenomenon, see *id.* at 143–60.

233. For an illuminating view of the development of the English criminal process, see LANGBEIN,

contamination of testimony en route to the trial, coupled with the technological capability to memorialize witnesses' statements, makes the orthodox adherence to the principle of orality seem arcane and deeply misguided. Historically, the primacy of oral testimony was based on its perceived superiority over documentary testimony.²³⁴ But that advantage is no longer relevant, as the *synthesized* oral testimony given in court fares poorly in comparison with a digital recording of that witness's raw statements. The oral presentation of evidence will better serve the fact-finding goal if witnesses were to testify under the shadow of their own raw statements, and when need be, the courtroom testimony were supplemented by those recorded statements.

To be sure, the recommendation to create investigative records is bound to be met with some resistance from law enforcement agencies.²³⁵ These concerns, however, have been largely dispelled by the experience of law enforcement agencies in the handful of states and scores of counties that mandate the practice. The consistent reaction of police personnel and prosecutors in these jurisdictions is nothing short of enthusiastic support. Transparency has turned out to be a beneficial tool in the hands of law enforcement, even as it continues to receive backing from defense attorneys.²³⁶

Critics are also likely to resist the sharing of investigative recordings with the defense, due primarily to fears of exposing witnesses to briber and even intimidation. While this argument has a kernel of truth, it is grossly overblown. A number of states, including Arizona, Colorado, New Jersey, and North Carolina, have implemented regimes that afford substantial discovery, with no apparent regrets.²³⁷ Some states, notably Florida and Vermont, even give

supra note 12.

234. As Jeremy Bentham stated, "oral testimony has a great superiority over written testimony." M. DUMONT, A TREATISE ON JUDICIAL EVIDENCE EXTRACTED FROM THE MANUSCRIPTS OF JEREMY BENTHAM, ESQ. 198 (1825).

235. A fuller discussion of these objections is provided in *In Doubt*. See SIMON, *supra* note 6, at ch. 8.

236. As stated by a Minnesota official, the order by the state supreme court to tape interrogations was "the best thing we've ever had rammed down our throats." Statement of Alan K. Harris, Deputy Prosecutor Hennepin County, quoted in Thomas P. Sullivan, *Electronic Recordings of Custodial Interrogations: Everybody Wins*, 95 J. CRIM. L. & CRIMINOLOGY 1127, 1127 (2005). Likewise, a North Carolina police chief explains that having "greater documentation and better evidence . . . will be a greater thing when it comes to successful prosecution of the case." A survey of 630 active police investigators found that eighty-one percent of the respondents believed that interrogations ought to be recorded in full. Saul M. Kassin et al., *Police Interviewing and Interrogation: A Self-Portrait Survey of Police Practices and Beliefs*, 31 LAW & HUM. BEHAV. 381, 393 (2007).

For a critique of the FBI's arguments, see Thomas P. Sullivan, *Recording Federal Custodial Interviews*, 45 AM. CRIM. L. REV. 1297, 1315-36 (2008). Undoubtedly, some investigations will go unrecorded. This could occur due to equipment failure, witness non-cooperation, or police misconduct. Based on the experience with taping of custodial interrogations, there is reason to believe that the majority of investigations will be taped, primarily because the police will be motivated to do so. Thomas P. Sullivan & Andrew W. Vail, *The Consequences of Law Enforcement Officials' Failure to Record Custodial Interviews as Required by Law*, 99 J. CRIM. L. & CRIMINOLOGY 215, 221 (2009).

237. For a discussion, see *Expanded Discovery in Criminal Cases; A Policy Review*, THE JUSTICE PROJECT (2007), http://www.pewtrusts.org/uploadedFiles/wwwpewtrustsorg/Reports/Death_penalty_reform/Expanded%20discovery%20policy%20brief.pdf.

criminal defendants the right to depose the prosecution's witnesses ahead of the trial.²³⁸ These states do not appear to show any of the calamities predicted by the critics of such policies.²³⁹

At the end of the day, the combination of best-practice investigative procedures and transparent investigations has the potential to bring about a transformative change throughout the criminal justice process. More accurate evidence will enable the process to meet its epistemic demands without having to rely on the legal mechanisms embedded in the criminal trial.

238. The deposition of witnesses in criminal proceedings is permitted in some form also in Indiana, Iowa, Missouri, and North Dakota. See discussion in Thomas, *supra* note 40, at 596–98.

239. Critics might oppose this proposed measure also on the grounds that it alters the principle of orality, a central feature of the Anglo-American criminal trial by which testimony is heard live, in open court. *Washington v. Crawford*, 541 U.S. 36, 43 (2004). The proposition, however, should not be seen as a challenge to the principle of orality, as the recorded testimony should supplement, rather than replace, oral testimony.