

2005

## Evaluating Juries by Comparison to Judges: A Benchmark for Judging?

Jennifer K. Robbennolot  
jkr@jkr.com

Follow this and additional works at: <http://ir.law.fsu.edu/lr>



Part of the [Law Commons](#)

---

### Recommended Citation

Jennifer K. Robbennolot, *Evaluating Juries by Comparison to Judges: A Benchmark for Judging?*, 32 Fla. St. U. L. Rev. (2005).  
<http://ir.law.fsu.edu/lr/vol32/iss2/8>

This Article is brought to you for free and open access by Scholarship Repository. It has been accepted for inclusion in Florida State University Law Review by an authorized administrator of Scholarship Repository. For more information, please contact [bkaplan@law.fsu.edu](mailto:bkaplan@law.fsu.edu).

# FLORIDA STATE UNIVERSITY LAW REVIEW



## EVALUATING JURIES BY COMPARISON TO JUDGES: A BENCHMARK FOR JUDGING?

*Jennifer K. Robbennolot*

VOLUME 32

WINTER 2005

NUMBER 2

---

Recommended citation: Jennifer K. Robbennolot, *Evaluating Juries by Comparison to Judges: A Benchmark for Judging?*, 32 FLA. ST. U. L. REV. 469 (2005).

# EVALUATING JURIES BY COMPARISON TO JUDGES: A BENCHMARK FOR JUDGING?

JENNIFER K. ROBBENOLT\*

I.	INTRODUCTION.....	469
II.	METHODOLOGICAL APPROACHES FOR COMPARING JURORS AND JUDGES .....	472
	A. <i>Rates of Judge-Jury Agreement</i> .....	472
	B. <i>Archival Analysis</i> .....	474
	C. <i>Experimental Simulations</i> .....	475
	D. <i>The Utility of Multiple Methods</i> .....	476
III.	JUDICIAL PERCEPTIONS OF JURORS .....	476
	A. <i>Judge-Jury Agreement Studies</i> .....	476
	B. <i>Other Estimates of Judge-Jury Agreement</i> .....	480
	C. <i>Judges' Perceptions of Juries</i> .....	481
IV.	ARCHIVAL STUDIES OF OUTCOMES .....	483
V.	EXPERIMENTAL STUDIES .....	486
	A. <i>Consideration of the Evidence</i> .....	487
	1. <i>Scientific, Expert, and Statistical Evidence</i> .....	487
	(a) <i>Scientific Evidence</i> .....	489
	(b) <i>Statistical Evidence</i> .....	491
	2. <i>Inadmissible Evidence</i> .....	492
	B. <i>Damages</i> .....	493
	C. <i>Risk Assessment</i> .....	496
	D. <i>Criminal Cases</i> .....	498
	E. <i>Judges and Juries—The Effects of Group Decisionmaking</i> .....	500
VI.	JUDICIAL DECISIONMAKING AS A BENCHMARK FOR EVALUATING JURY DECISIONMAKING.....	502
	A. <i>Interpreting the Differences</i> .....	503
	B. <i>Research Directions</i> .....	507
VII.	CONCLUSION .....	509

## I. INTRODUCTION

Juries usually do their job very well . . . . But juries also have the ability to disappoint us, sometimes to the point of forcing us to question whether we should have jury trials at all. One of this country's great observers of human nature, Mark Twain, once complained that juries had become "the most ingenious and infallible agency for defeating justice that human wisdom could contrive."<sup>1</sup>

Juries have been widely criticized and widely studied. The popular view is that juries are pro-plaintiff decisionmaking bodies, easily

---

\* Associate Dean for Faculty Research and Development and Associate Professor of Law, University of Missouri-Columbia School of Law. B.S., Willamette University; J.D., Ph.D. (Psychology), University of Nebraska-Lincoln. My thanks go to Phil Peters, Grant Robbennolt, and the symposium participants for their comments on earlier versions of this Article, to Jay Hastings, Cecily Helms, and Elizabeth McNichols for research assistance, and to the University of Missouri Law School Foundation for financial support.

1. Sandra Day O'Connor, *Juries: They May Be Broke, but We Can Fix Them*, 44 FED. LAW. 20, 20 (1997).

swayed by sympathy, highly likely to find liability, and inclined to award high amounts in damages. A wealth of empirical research, however, calls those stereotypes into doubt. Reviews of empirical research examining jury decisionmaking, while finding particular areas in which juries encounter difficulty (for example, understanding legal instructions), attest to the overall competence of juries as decisionmakers.<sup>2</sup>

Nonetheless, criticism of the jury system has led to recent calls to eliminate or to minimize the role of the jury in deciding legal cases. Many of the arguments made against the institution of the jury carry with them—either explicitly or implicitly—the suggestion that jury decisionmaking ought to be replaced with judicial decisionmaking.<sup>3</sup> As Harry Kalven argued:

When one asserts that jury adjudication is of low quality, he must be asserting that jury decisions vary in some significant degree from those a judge would have made in the same cases. If he denies this and wishes to include the judge, he has lost any baseline, and with it any force, for his criticism. . . . Further, trial by judge is the relevant and obvious alternative to trial by jury. To argue against jury trial is, therefore, to argue for bench trial.<sup>4</sup>

In addition to being the most likely alternative mechanism for decisionmaking, judges are also the individuals with whom the jury divides the labor of deciding cases. Cases that go to trial are decided by either a judge or a jury and, within jury trials, judges and juries di-

2. See, e.g., VALERIE P. HANS & NEIL VIDMAR, *JUDGING THE JURY* (1986); Edith Greene et al., *Jurors and Juries: A Review of the Field*, in *TAKING PSYCHOLOGY AND LAW INTO THE TWENTY-FIRST CENTURY* 225 (James R.P. Ogloff ed., 2002); Robert MacCoun, *Inside the Black Box: What Empirical Research Tells Us About Decisionmaking by Civil Juries*, in *VERDICT: ASSESSING THE CIVIL JURY SYSTEM* 137 (Robert E. Litan ed., 1993); Jennifer K. Robbennolt, *Determining Punitive Damages: Empirical Insights and Implications for Reform*, 50 *BUFF. L. REV.* 103 (2002); Neil Vidmar, *The Performance of the American Civil Jury: An Empirical Perspective*, 40 *ARIZ. L. REV.* 849 (1998).

3. See, e.g., Reid Hastie & W. Kip Viscusi, *What Juries Can't Do Well: The Jury's Performance as a Risk Manager*, 40 *ARIZ. L. REV.* 901, 917 (1998); Paul Mogin, *Why Judges, Not Juries, Should Set Punitive Damages*, 65 *U. CHI. L. REV.* 179 (1998); Cass R. Sunstein et al., *Assessing Punitive Damages (with Notes on Cognition and Valuation in Law)*, 107 *YALE L.J.* 2071, 2127 (1998); W. Kip Viscusi, *How Do Judges Think About Risk?*, 1 *AM. L. & ECON. REV.* 26, 60 (1999); see also Mark Curriden, *Putting the Squeeze on Juries*, *A.B.A. J.*, Aug. 2000, at 52, 53; William Glaberson, *Juries, Their Powers Under Siege, Find Their Role Is Being Eroded*, *N.Y. TIMES*, Mar. 2, 2001, at A1; Sir Louis Blom-Cooper, *A Judge Can Do the Work of 12 Amateurs, and Better*, *TIMES* (London), Oct. 21, 2003, at 10. Blom-Cooper argues that:

The jury is the high point of amateurism, potentially a recipe for incompetence and bias. The mood of civilised systems of criminal justice increasingly demands professionalism. I am not contemptuous of the amateur's ability to judge human conduct, only the task of evaluating evidence in the courtroom, which is a job for professionals.

*Id.*

4. Harry Kalven, Jr., *The Dignity of the Civil Jury*, 50 *VA. L. REV.* 1055, 1063 (1964).

vide the work by assuming different roles. Accordingly, any reform of the jury system ought to take into account the degree to which juries and judges would make decisions differently.<sup>5</sup>

Proposals to substitute judicial decisionmaking for that of the jury or to make other jury reforms, however, have been put forth without a clear understanding of the similarities and differences in how juries and judges decide legal cases. Compared to the extensive study of the decisionmaking of jurors and juries, there has been relatively little examination of trial judges' decisionmaking, and even fewer studies have directly compared the decisionmaking of juries and judges.<sup>6</sup> While there are reasonable arguments for why judges and juries might differ in how they make legal decisions, these arguments are, in the absence of more research, speculative at best.<sup>7</sup>

The purpose of this Article is to examine what the limited, but growing, body of empirical research has revealed about the similarities and differences in juror and judicial decisionmaking. Part II describes the range of methods available for comparing jurors and judges. Part III discusses studies examining the degree to which judges agree with decisions made by juries. Part IV examines research that has compared the outcomes of jury trials and bench trials. Part V reviews simulation experiments that have compared the decisionmaking processes of judges and jurors in various legal domains. Part VI describes the normative difficulties in focusing on the comparison of juries to judges, the relative advantages and disadvantages of judicial and jury decisionmaking, and how future research might be structured in order to develop a more useful picture of how these two types of decisionmakers compare. Part VII concludes.

---

5. Phoebe C. Ellsworth, *Jury Reform at the End of the Century: Real Agreement, Real Changes*, 32 U. MICH. J.L. REFORM 213, 219 (1999) ("If juries are susceptible to errors of reasoning on some aspects of their task, and judges are not, the focus of reform efforts should be specifically aimed at the jury; but if judges are susceptible to the same errors as jurors, obviously a different kind of reform is needed.")

6. Michelle Chernikoff Anderson & Robert J. MacCoun, *Goal Conflict in Juror Assessments of Compensatory and Punitive Damages*, 23 LAW & HUM. BEHAV. 313, 328 (1999) ("Considering the amount of attention that has been given to jurors' ability to use legal evidence, it is remarkable how little we know about their professional counterparts [that is, judges]."); Kevin M. Clermont & Theodore Eisenberg, *Trial by Jury or Judge: Transcending Empiricism*, 77 CORNELL L. REV. 1124, 1125 (1992) ("One might suppose that, with so much at stake for so long, we would all know a lot about the ways juries differ from judges in their behavior. In fact, we know remarkably little."); Ellsworth, *supra* note 5, at 218-19 ("[T]here is hardly any research that actually compares judges' and juries' responses to the same trial materials; thus we do not know which of the jury's shortcomings are also characteristic of judges."); Michael J. Saks, *Do We Really Know Anything About the Behavior of the Tort Litigation System—And Why Not?*, 140 U. PA. L. REV. 1147, 1230-31 (1992) ("Judge decision-making has been largely neglected in studies of the tort system, suggesting that juries have been serving well their function as a lightning rod for the judiciary.")

7. See Anderson & MacCoun, *supra* note 6, at 327-28.

## II. METHODOLOGICAL APPROACHES FOR COMPARING JURORS AND JUDGES

It is certainly reasonable to expect differences in how juries and judges decide legal cases. Justice Breyer has noted that “one cannot expect . . . jurors to interpret law like judges, who work within a discipline and hierarchical organization that normally promotes roughly uniform interpretation and application of the law.”<sup>8</sup> By virtue of their differing backgrounds and experiences, one might reasonably predict differences in decisionmaking. In addition, one might expect the different structures of the decisionmaking processes to create some divergence. For example, trial court judges make decisions as individuals, while juries decide cases as groups. However, both juries and judges are human decisionmakers—sharing basic human cognitive processes—who are called upon to make complicated decisions.

At bottom, juries and judges probably do not differ wildly in their decisionmaking, nor is it likely that they would make identical decisions in all circumstances. It is more likely that juries and judges make decisions in legal cases that are mostly similar, but that diverge in some instances. The challenge is to identify the ways in which juries and judges are similar and to distinguish those ways in which their decisionmaking differs. As Lempert argues: “Whatever research may show, the essential point is that it must be comparative. Both jury and judge must be subject to scrutiny.”<sup>9</sup>

Researchers have attempted to do this comparative research in a number of ways, each with its own advantages and disadvantages. These include studies of judicial agreement with jury verdicts, comparisons of outcomes in bench and jury trials, and simulation research that examines the influences on and processes by which judges and jurors make decisions.

### A. Rates of Judge-Jury Agreement

One approach to studying the similarities and differences in the decisionmaking of juries and judges is to ask judges to indicate their agreement or disagreement with the decisions made by juries in cases over which the judges preside. In addition to gathering information about the outcome of the case (that is, the verdict or the damage award or both), it is possible to ask judges to provide other perceptions about the case and the jury. A primary advantage of this approach is that each judge and jury evaluate an identical case. A second advantage is that the focus is on judge and jury reactions to

---

8. *BMW of N. Am., Inc. v. Gore*, 517 U.S. 559, 596 (1996) (Breyer, J., concurring).

9. Richard O. Lempert, *Civil Juries and Complex Cases: Let's Not Rush to Judgment*, 80 MICH. L. REV. 68, 95 (1981).

complete, real-world cases. Since both decisionmakers hear all the evidence and the arguments, their decisions reflect the influences of the entire case.

At the same time, however, there are several disadvantages to this approach. First, it is difficult to make distinctions in patterns of judge-jury agreement among cases using this method.<sup>10</sup> If disagreements are confined to subsets of cases with particular characteristics, rates of overall agreement may fail to reveal these moderators.

Second, there are a variety of perceptual and judgment difficulties with a method that relies on the judge to self-report agreement or disagreement with the jury. Judges, like other people, may have difficulty reporting on their own cognitive processes and identifying the factors that affect their decisions.<sup>11</sup> Relatedly, research has demonstrated that people tend to remember and construe information in ways that are consistent with the decision they have already made.<sup>12</sup>

A third disadvantage relates to the differing roles of the presiding judge and the jury. The trials studied with these methods are, by definition, those in which the juries act as the factfinder. The judges, rather than acting as finders of fact, act in a presiding role. In this role, a judge may behave differently than she does in a bench trial in which she is charged with finding the facts and reaching a final decision in the case. In addition, these roles are interrelated in such a way that judges and juries may influence each other's opinions about the case.<sup>13</sup>

Finally, some studies using this approach have asked judges to indicate their agreement or disagreement with the jury after the jury has rendered a verdict, rather than asking judges to independently indicate what their decision would have been prior to learning the jury verdict. Knowing the jury's verdict may cause a judge to differently evaluate aspects of the case that might affect his or her agreement. Relatedly, it is possible that judges may indicate agreement with a jury verdict that they found reasonable, even if it was not the verdict they would have returned.<sup>14</sup> Moreover, judges' responses may also be influenced by their overall attitude toward the jury system;

10. *Id.* at 103.

11. See Richard E. Nisbett & Timothy DeCamp Wilson, *Telling More than We Can Know: Verbal Reports on Mental Processes*, 84 *PSYCHOL. REV.* 231, 231 (1977).

12. See Mara Mather et al., *Misremembrance of Options Past: Source Monitoring and Choice*, 11 *PSYCHOL. SCI.* 132, 132 (2000).

13. See Peter David Blanck, *What Empirical Research Tells Us: Studying Judges' and Juries' Behavior*, 40 *AM. U. L. REV.* 775, 777 (1991); Peter David Blanck et al., Note, *The Appearance of Justice: Judges' Verbal and Nonverbal Behavior in Criminal Jury Trials*, 38 *STAN. L. REV.* 89, 91 (1985).

14. Lempert, *supra* note 9, at 102 ("[A] judge's recollections may be colored by the verdict returned. In particular, a judge who thought the case close might report agreement with any jury verdict.").

that is, judges with positive views of the jury system may be more likely to report agreement with the jury. A better variant on this approach would be to ask judges to give their verdicts prior to learning the verdict of the jury.<sup>15</sup>

### B. Archival Analysis

A second approach to comparing judges and jurors involves examining cases that have been decided by either judges or juries and comparing the outcomes. One of the main advantages of archival studies is their high verisimilitude, or ecological validity.<sup>16</sup> These studies examine decisions in real cases in which judges or juries were faced with full case facts, subtleties of evidence, and complete legal instructions and were required to make decisions with real consequences. With enough cases and a detailed system for carefully coding case and decisionmaker characteristics, statistical analysis can be used to identify patterns and associations between aspects of the cases and the decisions made. In addition, archival research, unlike the other methods discussed here, does not require judge or jury participation in the research process. Thus, archival research enjoys the practical advantage of not needing to secure the cooperation of judges—a group from which it is notoriously difficult to secure participation.

The difficulty with using the results in decided cases to compare the decisionmaking of judges and juries is that the judges and juries have heard different cases. Each case presents a host of subtle and not so subtle differences—in their facts, injuries, legal claims, legal representation, and parties—which might provide an alternative explanation for any differences obtained. The challenge for archival research, then, is to identify the relevant variables and devise a way to measure them in order to statistically control them. The ability to code and control for variables of interest, however, is frequently limited by the (often insufficient) court records kept. Finally, even if done well, archival research identifies correlations between and among variables, but not causal relationships.<sup>17</sup>

---

15. See, e.g., 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 (2005).

16. Ecological validity refers to the degree to which the conditions of the study reproduce the real world conditions. See *infra* notes 21-23 and accompanying text.

17. For discussions of archival research methods generally, see MacCoun, *supra* note 2, at 141-42, and Neil Vidmar, *Making Inferences About Jury Behavior from Jury Verdict Statistics: Cautions About the Lorelei's Lied*, 18 LAW & HUM. BEHAV. 599 (1994).

### C. Experimental Simulations

A third approach utilizes experimental methodology, which gives the researcher the ability to isolate and manipulate variables of interest and to eliminate alternative explanations. In a typical experiment, a large number of participants evaluate the same simulated case.<sup>18</sup> All characteristics of the case are held constant; the only attribute varied is the characteristic or procedure to be studied. Thus, observed differences in responses, such as verdicts, can be attributed to the variable of interest, unconfounded by other influences. Simulation methodology allows isolation of specific variables, permits observation of deliberation processes, and allows experimental manipulation of legal rules and procedures as well as case and party characteristics.<sup>19</sup> Thus, in order to compare judges and jurors, both groups would be asked to evaluate the same case and their responses would be compared. Differences in their responses can then be attributed to differences between judges and jurors, rather than to differences in the cases faced by each group.

Experimental simulations, however, are limited in the degree to which the research setting can reflect the relevant legal conditions. Departures from the conditions of actual trials may include using written vignettes, participants that do not reflect the jury pool, or choosing not to include group deliberations. Moreover, no matter how sophisticated the simulation, it is still a simulation.<sup>20</sup> Concern about the ability to generalize the experimental simulation research centers on this lack of ecological validity.<sup>21</sup> However, it is not necessary that a study mirror the conditions of an actual trial for the study to have high external validity.<sup>22</sup> As long as the experimental simulation elicits responses similar to those in the real world, it has high gener-

---

18. Because it is not feasible to randomly assign cases to be decided by judge or jury, I focus here primarily on simulation studies rather than field experiments.

19. See generally Brian H. Bornstein, *The Ecological Validity of Jury Simulations: Is the Jury Still Out?*, 23 LAW & HUM. BEHAV. 75 (1999); Robert M. Bray & Norbert L. Kerr, *Use of the Simulation Method in the Study of Jury Behavior: Some Methodological Considerations*, 3 LAW & HUM. BEHAV. 107, 117 (1979); Shari Seidman Diamond, *Illuminations and Shadows from Jury Simulations*, 21 LAW & HUM. BEHAV. 561 (1997); MacCoun, *supra* note 2, at 143-44; Wayne Weiten & Shari Seidman Diamond, *A Critical Review of the Jury Simulation Paradigm: The Case of Defendant Characteristics*, 3 LAW & HUM. BEHAV. 71, 75-83 (1979).

20. See Brian H. Bornstein & Sean G. McCabe, *Jurors of the Absurd? The Role of Consequentiality in Jury Simulation Research*, 32 FLA. ST. U. L. REV. 443 (2005).

21. *Id.*

22. External validity essentially concerns the extent to which a set of research findings can be generalized. The construct applies to generalizations across various persons, times, and settings as well as to generalizations to specific persons, times, and settings. See THOMAS D. COOK & DONALD T. CAMPBELL, *QUASI-EXPERIMENTATION: DESIGN & ANALYSIS ISSUES FOR FIELD SETTINGS* (1979).

alizability even if it does not precisely replicate the particular legal setting.<sup>23</sup>

#### D. *The Utility of Multiple Methods*

Certainly it would be unwise to rely on a single study as dispositive of a research question.<sup>24</sup> Similarly, it would be unwise to rely on a single methodology to explore a research question. A more complete and nuanced understanding of a phenomenon comes not only from multiple studies, but from multiple studies done by different researchers, using a variety of methods, participant populations, approaches, and materials. This is true because each study has unique advantages and disadvantages, and reading multiple studies together makes the most of the advantages while minimizing the disadvantages. To the extent that studies using multiple methods (for example, archival studies and experimental simulations) converge on a common understanding of a phenomenon, greater confidence in the results can be had. To the extent that the results differ, attempts can be made to explain the divergence. Each study, by itself, is interesting, but it is when a number of studies are read together that a more complete picture of the phenomenon of interest is likely to emerge.

### III. JUDICIAL PERCEPTIONS OF JURORS

One approach to comparing juries and judges is to ask judges to indicate their agreement or disagreement with the decisions made by juries in cases over which they preside. Judge-jury agreement has been assessed in large numbers of individual cases; these studies consistently find relatively high rates of judge-jury agreement. More general estimates by judges of the rate at which they agree with juries in all the cases in which they have presided show similarly high rates of agreement. In addition, judges surveyed about their perceptions of jury decisionmaking tend to report relatively high levels of satisfaction with juries.

#### A. *Judge-Jury Agreement Studies*

Kalven and Zeisel conducted the classic large-scale comparison of judges and jurors in both criminal and civil cases.<sup>25</sup> Examining 3576

---

23. See Bornstein, *supra* note 19, at 84. Unfortunately, explicit comparisons of behavior in the real world with behavior under various lab conditions are somewhat rare.

24. See Greene et al., *supra* note 2, at 247 (“[I]t is a rare individual study that is able to address all of these issues. More importantly, it is not clear that individual studies should address every issue.”); Michael J. Saks, *Improving APA Science Translation Amicus Briefs*, 17 LAW & HUM. BEHAV. 235, 240 (1993) (noting the “nearly universal consensus that a single study is not sufficient”).

25. Kalven, *supra* note 4, at 1056 & n.3; HARRY KALVEN, JR. & HANS ZEISEL, *THE AMERICAN JURY* 12 (1966).

criminal trials, they found that judges reported that they would have made the same guilt determinations as did the jury 78% of the time.<sup>26</sup> In 3% of the cases, the jury convicted when the judge would have acquitted; in 19% of the cases, the jury acquitted (or hung) when the judge would have convicted.<sup>27</sup> Of those cases in which there was agreement to convict on the same charge and in which the jury set the sentence, judges indicated agreement with the jury sentence in 57% of the cases.<sup>28</sup>

Notably, Kalven and Zeisel found that judges did not attribute disagreements to jury misunderstandings.<sup>29</sup> In addition, they found that the rates of judge-jury disagreement were comparable not only in cases classified by the judges as easy to understand, but also in difficult cases. Instead, higher rates of judge-jury disagreement were found in cases that the judges classified as “close” than in cases described as “clear.”<sup>30</sup>

Kalven and Zeisel found similar results for civil cases. Across approximately 4000 civil cases, judges reported that their liability decisions would have matched the juries’ decisions 78% of the time.<sup>31</sup> The disagreements were evenly split among cases in which the jury found for the plaintiff when the judge would have found for the defendant (12%) and those in which the jury found for the defendant when the judge would have found for the plaintiff (10%).<sup>32</sup> In cases in which both judge and jury would have awarded damages, judges reported that they would have awarded less than the jury awarded in 52% of

26. KALVEN & ZEISEL, *supra* 25, at 58 tbl.12 (showing that in 14% of cases, both judge and jury would acquit (or hung); in 64% of cases, both judge and jury would convict). Here “convicting” refers to convicting on some charge. *See id.* at 60.

27. *Id.* at 59.

28. *Id.* at 61 tbl.14. Judges would have sentenced more leniently in 16% of these cases and would have sentenced more harshly in 26% of these cases. *Id.* “[T]he jury agrees with the judge often enough to be reassuring, yet disagrees often enough to keep it interesting.” Kalven, *supra* note 4, at 1064.

29. KALVEN & ZEISEL, *supra* note 25, at 152-53 (“[T]he judge almost never advances the inability of the jury to understand as a reason for disagreement.”). Kalven and Zeisel also found that jury “verdicts move basically with the weight and direction of the evidence.” *Id.* at 162.

30. *Id.* at 157. In clear cases, the disagreement rate was quite low whether the case was classified as easy (9%) or difficult (8%). In contrast, higher rates of disagreement were found in both easy (41%) and difficult (39%) cases that were classified as close cases. *Id.* at 157 tbl.50.

The result is a stunning refutation of the hypothesis that the jury does not understand. While, as we can see, jury disagreement is greater in close cases than in clear ones, there is virtually no difference between the frequency of disagreement when the case is easy and when the case is difficult; this holds true for the cases that are clear as well as for the close ones.

*Id.* (footnote omitted).

31. *Id.* at 63 tbl.16 & n.11. In 47% of cases both would have found for the plaintiff; in 31% of cases both would have found for the defendant. *Id.*

32. *Id.*

cases, more than the jury awarded in 39% of cases, and approximately the same amount as the jury awarded in 9% of the cases.<sup>33</sup> On average, Kalven and Zeisel found that juries' damage awards were approximately 20% higher than judges reported they would have awarded.<sup>34</sup>

More recently, in a similar study, Heuer and Penrod examined judge-jury agreement in seventy-seven criminal and sixty-seven civil trials across thirty-three states.<sup>35</sup> For the criminal trials, judges reported that they would have made the same guilt determinations as did the jury in approximately 73% of cases.<sup>36</sup> In approximately 3% of the cases, the jury convicted when the judge would have acquitted; in approximately 25% of the cases, the jury acquitted (or hung) when the judge would have convicted.<sup>37</sup> For the civil trials, judges reported that their liability decisions would have been the same as the juries' in approximately 63% of the cases.<sup>38</sup> Disagreements were evenly split between cases in which judges reported that they would have found for the defendant when the jury found for the plaintiff (18%) and those in which judges reported that they would have found for the plaintiff when the jury found for the defendant (19%).<sup>39</sup> Like Kalven and Zeisel, Heuer and Penrod found no relation between the complexity of the trial or the trial procedures and rates of judge-jury agreement.<sup>40</sup>

Finally, Eisenberg and his colleagues looked at judge-jury agreement rates in a sample of 300 criminal trials in four jurisdictions.<sup>41</sup> Judges filled out two questionnaires, one before the jury reached its verdict and another following the verdict. The agreement rates were strikingly similar to those found by Kalven and Zeisel: judges would have made the same guilt determinations as did the jury 75% of the time.<sup>42</sup> In 6% of the cases, the jury convicted when the judge would have acquitted; in 19% of the cases, the jury acquitted when the

---

33. See Kalven, *supra* note 4, at 1065.

34. *Id.*; see also GORDON BERMANT ET AL., PROTRACTED CIVIL TRIALS: VIEWS FROM THE BENCH AND THE BAR 41-42 (1981) (reporting findings from a study of 18,528 federal trials terminated between July 1, 1976 and December 31, 1978). Bermant and his colleagues found that "rates of appeals after bench [24%] and jury trials [22%] are practically identical" and that rates of successful appeals (resulting in reversal or remand) were comparable as well (19% of bench trials and 17% of jury trials). *Id.* at 42.

35. Larry Heuer & Steven Penrod, *Trial Complexity: A Field Investigation of Its Meaning and Its Effects*, 18 LAW & HUM. BEHAV. 29, 34, 48 (1994) [hereinafter Heuer & Penrod, *Trial Complexity*].

36. *Id.* at 48 tbl.12.

37. *Id.*

38. *Id.* at 48 tbl.13.

39. *Id.*

40. *Id.* at 48-49.

41. Eisenberg et al., *supra* note 15, at 173.

42. *Id.* at 180, 181 tbl.1. In 13% of cases, both judge and jury would acquit; in 62% of cases, both judge and jury would convict. *Id.* at 181 tbl.1.

judge would have convicted.<sup>43</sup> Consistent with other studies of this type, Eisenberg and his colleagues found no relation between the rate of disagreement and the complexity of the trial.<sup>44</sup> Instead, the juries appeared to apply a higher standard for finding guilt beyond a reasonable doubt.<sup>45</sup>

TABLE 1  
COMPARISON OF JUDGE-JURY AGREEMENT STUDIES

	AGREEMENT	JUDGE LENIENT	JURY LENIENT
<b>Criminal Cases</b>			
Kalven & Zeisel (1969)	78%	3%	19%
Heuer & Penrod (1994)	73%	3%	25%
Eisenberg et al. (2004)	75%	6%	19%
<b>Civil Cases</b>			
Kalven & Zeisel (1966)	78%	12%	10%
Heuer & Penrod (1994)	63%	18%	19%

Interestingly, the rates of judge-jury agreement found in these studies compare favorably to rates of agreement between decision-makers considering other types of complex decisions. Diamond reviewed studies reporting rates of agreement for a variety of different decisionmakers who performed a variety of judgment tasks. Judge-jury agreement rates (78%) fare well when stacked up against scientists engaged in peer review (75% agreement rate), employment interviewers (70%), psychiatrists diagnosing psychiatric illness (70%), and physicians diagnosing physical illness (67%-77%).<sup>46</sup>

Importantly, these judge-jury agreement rates are also comparable to the rates of agreement among judges themselves. Diamond and Zeisel examined the sentencing recommendations of judges participating in sentencing councils.<sup>47</sup> Before the meeting of each council, each judge was provided with the presentencing report and then

---

43. *Id.* Note that asymmetry is not universal—the judge convicts more when evidence is medium or strong (from either perspective), but the judge convicts less in cases in which the judge thinks evidence is weak. *Id.* at 186-89.

44. *Id.* at 192. Eisenberg and his colleagues measured complexity in terms of evidentiary complexity and in terms of legal complexity and from the perspective of judges and of jurors. *Id.* at 190-92.

45. *Id.* at 191.

46. Shari Seidman Diamond, *Order in the Court: Consistency in Criminal-Court Decisions*, in 2 THE MASTER LECTURE SERIES: PSYCHOLOGY AND THE LAW 119, 125 (C. James Scheirer & Barbara L. Hammonds eds., 1983); see also STEPHEN L. FIELDING, THE PRACTICE OF UNCERTAINTY 52 (1999).

47. Shari Seidman Diamond & Hans Zeisel, *Sentencing Councils: A Study of Sentence Disparity and Its Reduction*, 43 U. CHI. L. REV. 109, 109 (1975).

recorded a recommendation for sentencing.<sup>48</sup> Diamond and Zeisel found that judges agreed on whether the offender should be sentenced to prison approximately 70% of the time.<sup>49</sup> Thus, the judge-jury agreement rates reported above comport closely with the inter-judge agreement rates. Commenting on such comparisons, Saks noted that “[t]he jury’s concordance rates are particularly impressive given its unique situation of having to deal with a body of cases from which the easiest 80-95% have been removed.”<sup>50</sup>

### B. Other Estimates of Judge-Jury Agreement

In other studies, judges estimated their rates of agreement with juries in cases over which they have presided. For example, a recent survey of 393 Texas state trial judges and 594 federal trial court judges found that large majorities of judges reported agreeing with the jury all or most of the time.<sup>51</sup> These estimates of judge-jury agreement rates across all of the judges’ cases are likely to be less accurate than the estimates obtained by studies asking judges to report their agreement in individual cases. In addition to the difficulties attendant to asking judges to report agreement or disagreement detailed above, judges are likely to experience a variety of difficulties in making estimates across a wide variety of cases, including faulty memory and overreliance on cases that are highly available in memory.<sup>52</sup> Nonetheless, the results of these studies are consistent with the results of the judge-jury agreement studies reported above and contribute to a picture of high levels of agreement between judges and juries.

---

48. *Id.* at 117. The sentencing judge was responsible for making the final sentencing decision, but that decision was informed by the recommendations and discussion of the council. *Id.*

49. *Id.* at 119 tbl.2. For those cases in which the judges agreed that the offender should be incarcerated, the judges’ rate of agreement as to the length of the sentence was quite low—12% in one council and 10% in the other. *Id.* at 120 tbl.3. In comparison, when the judges agreed that the offender should be given a noncustodial sentence, the rate of agreement on the length of that sentence was quite high—98% in one council and 82% in the other. *Id.*

50. Saks, *supra* note 6, at 1236.

51. John B. Attanasio, *Foreword: Juries Rule*, 54 SMU L. REV. 1681, 1684 n.10 (2001) (finding that 92.1% of the federal sample and 90.5% of the state court sample reported agreement in “most” cases; and 4.7% of the federal sample and 5.1% of the state court sample reported agreeing with the jury “all of the time”).

52. Using the availability heuristic, individuals judge the likelihood of an event by the ease with which they can recall examples of similar events. *See generally* Amos Tversky & Daniel Kahneman, *Availability: A Heuristic for Judging Frequency and Probability*, in *JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES* 163 (Daniel Kahneman et al. eds., 1982).

Sentell conducted a series of surveys of 101 state<sup>53</sup> and seventeen federal trial judges in Georgia.<sup>54</sup> He provided participants with findings from Kalven and Zeisel's study and asked them how their experience compared. When presented with the finding that judges agreed with the liability determination of the jury in at least 79% of cases, the vast majority of the judges reported similar or even higher rates of agreement.<sup>55</sup> Similarly, when presented with Kalven and Zeisel's pattern of agreement with damage awards, the majority of judges indicated that this was consistent with their experience.<sup>56</sup> When asked about the reasons for judge-jury disagreement, most Georgia judges agreed with Kalven and Zeisel's assessment that plaintiff leanings by the jury<sup>57</sup> and the jury's inability to understand<sup>58</sup> did not cause judge-jury disagreements. Given potential susceptibility to anchoring and to leading questions, this level of agreement is perhaps not surprising. The study does, however, at least demonstrate that judges find Kalven and Zeisel's findings plausible.

### C. Judges' Perceptions of Juries

Surveys of judges also suggest that the judiciary generally holds a positive view of the jury. Judges report high levels of satisfaction with the jury system overall<sup>59</sup> and view the jury system as an essen-

53. R. Perry Sentell, Jr., *The Georgia Jury and Negligence: The View from the Bench*, 26 GA. L. REV. 85, 94-95 (1991) [hereinafter Sentell, *The View from the Bench*].

54. R. Perry Sentell, Jr., *The Georgia Jury and Negligence: The View from the (Federal) Bench*, 27 GA. L. REV. 59, 66-68 (1992) [hereinafter Sentell, *Federal Bench*].

55. Sentell, *The View from the Bench*, *supra* note 53, at 98-99 (87% of state sample reported that their rate of agreement was "about the same"; 10% as higher; 3% as lower); Sentell, *Federal Bench*, *supra* note 54, at 71 (56% of the federal sample reported that their rate of agreement was "about the same"; 44% reported higher rates of agreement).

56. Among those judges indicating that their experience differed, there appeared to be a split between those who thought that judges would make higher awards in more cases and those who thought that juries would make higher awards in more cases. In addition, Sentell noted that many judges indicated that they agreed with the jury award in a higher percentage of cases than that found by Kalven and Zeisel. Sentell, *The View from the Bench*, *supra* note 53, at 103-04; Sentell, *Federal Bench*, *supra* note 54, at 74-75. When presented with the finding that, on average, juries awarded 20% more than the judges would have in the same cases, 52% of the Georgia state court judges and 44% of the Georgia federal court judges indicated that their experience was comparable. Of those judges who indicated that their experience differed, some agreed that juries awarded more than judges would have but disagreed as to how much, and others indicated that the judge would have awarded more. Sentell, *The View from the Bench*, *supra* note 53, at 104-06; Sentell, *Federal Bench*, *supra* note 54, at 77.

57. Sentell, *The View from the Bench*, *supra* note 53, at 109-11 (86% agree; 14% disagree); Sentell, *Federal Bench*, *supra* note 54, at 80 (87% agree; 13% disagree).

58. Sentell, *The View from the Bench*, *supra* note 53, at 113-15 (94% agree jury generally able to understand; 6% disagree (several limited to "difficult cases" or inability of lawyers)); Sentell, *Federal Bench*, *supra* note 54, at 82 (94% agree; 6% (1 respondent) disagree in complex cases).

59. The judges studied by Kalven and Zeisel reported general overall satisfaction with the jury. For criminal cases, 77% reported that the jury system was "thoroughly satisfactory" and only 3% found the jury system unsatisfactory such that its use "should be

tial safeguard.<sup>60</sup> Judges overwhelmingly believe that juries attempt to follow the judge's legal instructions,<sup>61</sup> demonstrate understanding of the issues in the cases they hear,<sup>62</sup> and are not swayed by emotion.<sup>63</sup>

Judges also recognize challenges for the jury system. One survey found that many judges would support studying alternatives to trial by jury for certain types of cases.<sup>64</sup> Concern about the jury was expressed in particular in the context of complex civil cases, where a majority of the judges believed that "in complex civil cases, the jurors need more guidance than they usually get."<sup>65</sup> Slight majorities of the judges nonetheless rejected the idea of limiting use of juries for com-

---

sharply curtailed." For civil cases, support was slightly lower, but remained high with 64% reporting thorough satisfaction and 9% reporting that the system should be curtailed. Kalven, *supra* note 4, at 1073. Sentell also found relatively high levels of support for the jury among judges—for civil cases, 76.5% of the federal judges surveyed found the jury system "thoroughly satisfactory" and none found it unsatisfactory. Sentell, *Federal Bench*, *supra* note 54, at 87. State court judges also reported a high degree of satisfaction, though somewhat less than the federal judges—57% of the state court judges surveyed found the jury system "thoroughly satisfactory" and only 2% found it unsatisfactory. Sentell, *The View from the Bench*, *supra* note 53, at 120. Most of the judges surveyed reported similar levels of satisfaction with juries in criminal cases. *Id.* at 123; Sentell, *Federal Bench*, *supra* note 54, at 89. The vast majority of the respondents in the SMU/*Dallas Morning News* survey indicated that they thought that juries did "very well" or "moderately well" in reaching "just and fair" verdicts. Attanasio, *supra* note 51, at 1684 n.8 (98.6% of the federal trial judges and 98.3% of the Texas state court judges). In addition, large majorities believed that the jury system is "fine just the way it is" or is in "good" condition needing only minor work. *Id.* at 1684 n.9 (91.1% of the federal judges and 87.8% of the Texas state court judges).

60. A 1987 Louis Harris & Associates survey of 800 state and 200 federal judges who spent at least half of their time on civil cases found that just over three-quarters of the judges agreed that "for routine civil cases, the right to trial by jury is an essential safeguard which must be retained." Louis Harris & Assoc., *Judges' Opinions on Procedural Issues: A Survey of State and Federal Trial Judges Who Spend at Least Half Their Time on General Civil Cases*, 69 B.U. L. REV. 731, 746 (1989) [hereinafter Louis Harris].

61. See, e.g., Sentell, *The View from the Bench*, *supra* note 53, at 126 (93% of the state judges); Sentell, *Federal Bench*, *supra* note 54, at 91 (100% of the federal court judges); Louis Harris, *supra* note 60, at 746 (finding that 99% of the federal judges and 98% of the state judges agreed that "jurors usually make a serious effort to apply the law as they are instructed"); see also BERMANT ET AL., *supra* note 34, at 52 ("Judges and lawyers were uniformly complimentary of the diligence of the juries in these cases; with slightly less unanimity, they also affirmed the validity of the juries' deliberative processes.").

62. Louis Harris, *supra* note 60, at 746 (finding that 73% of the federal judges and 66% of the state judges disagreed with the statement that "too often jurors fail to apply the law because they aren't able to understand it"); Michelle L. Hartmann, *Is It a Short Trip Back to Manor Farm? A Study of Judicial Attitudes and Behaviors Concerning the Civil Jury System*, 54 SMU L. REV. 1827, 1853 (2001) (finding that 93.4% of the federal court judges and 83.9% of the state court judges believed that jurors understand the legal and evidentiary issues "very well" or "moderately well").

63. Louis Harris, *supra* note 60, at 746 (finding that 80% of the federal judges and 69% of the state judges disagreed with the statement that "the feelings jurors have about the parties often cause them to make inappropriate decisions").

64. *Id.* at 747 tbl.7.1 (63% of the federal judges and 68% of the state judges).

65. *Id.* at 746 (66% of the federal judges and 62% of the state judges).

plex civil cases<sup>66</sup> or for “complicated business cases.”<sup>67</sup> Instead, a majority of the judges favored limitations on the use of juries for “minor civil cases involving small sums of money.”<sup>68</sup>

#### IV. ARCHIVAL STUDIES OF OUTCOMES

Another approach to the judge-jury question compares “win rates” between cases heard by juries and those heard by judges. As noted above, the benefit of this approach is that it examines the actual decisions of judges in cases in which they are the decisionmaker, rather than examining their sense of agreement after the fact.

Using the Administrative Office of the United States Courts database, Clermont and Eisenberg compared outcomes in federal civil cases tried before judges and juries between 1979 and 1989.<sup>69</sup> They found that in most types of cases there were no differences in plaintiff win rates before judges and juries.<sup>70</sup> They did find, however, that plaintiffs were more successful in front of judges in several types of cases—most notably products liability, medical malpractice, and motor vehicle cases.<sup>71</sup> In addition, they found a correlation between plaintiff win rates in judge and jury trials across different case types, concluding that “[p]laintiffs tend to do well before judges in the same case categories in which they tend to do well before juries.”<sup>72</sup>

Several studies of this type have used data from the Civil Trial Court Network (CTCN). For example, Eisenberg and his colleagues used this database to compare the decisions of juries and judges across 9000 state court civil cases from 1996.<sup>73</sup> Overall, they found that plaintiffs were more successful in front of judges (62% plaintiff win rate) than in front of juries (47% plaintiff win rate), though rates varied by type of case.<sup>74</sup> Controlling for the size of the compensatory

66. *Id.* at 747 (52% of the federal judges and 59% of the state judges).

67. *Id.* (60% of the federal judges and 58% of the state judges).

68. *Id.* (52% of the federal judges and 71% of the state judges). A small majority of judges agreed that “in some complex civil cases, trial before a panel of experts would be preferable to trial by jury.” *Id.* at 748 tbl.7.2 (53% of the federal judges and 51% of the state judges). Overall, most judges in the SMU/*Dallas Morning News* study thought that the right to a jury trial ought to be “left the same” or even expanded, a substantial minority favored some reduction, and very few favored eliminating the jury. Attanasio, *supra* note 51, at 1685 n.16. Similarly, while some judges favored using juries in fewer types of cases and a handful favored expanding the types of cases that juries hear, most judges favored the status quo. *Id.*

69. Clermont & Eisenberg, *supra* note 6, at 1133.

70. *Id.* at 1137.

71. *Id.* Plaintiffs were more successful in front of juries only in marine personal injury and Federal Employer’s Liability Act cases. *Id.*

72. *Id.* at 1138. In some types of cases, juries awarded more in damages; however, in other types of cases, judges awarded more. *Id.* at 1141.

73. Theodore Eisenberg et al., *Juries, Judges, and Punitive Damages: An Empirical Study*, 87 CORNELL L. REV. 743, 747-49 (2002).

74. *Id.* at 749-50.

award, the status of the parties, and the type of case, Eisenberg and his colleagues found no differences between judges and juries in the frequency with which they awarded punitive damages, in the amounts of the awards, or in the amount of punitive damages awarded per unit of compensatory damages.<sup>75</sup> They did find, however, that judges (more so than juries) became more likely to award punitive damages as compensatory damages increased,<sup>76</sup> and they found some evidence suggesting that punitive damage awards made by juries were somewhat more variable than were those of judges.<sup>77</sup>

In a later study, Eisenberg, Rachlinski, and Wells used the subset of cases from the CTCN database in which plaintiffs won and were awarded both compensatory and punitive damages awards to explore whether the ratio of punitive and compensatory damages awarded by judges and juries differed by the type of case tried.<sup>78</sup> They found evidence that judges and juries awarded similar amounts in punitive damages per unit of compensatory damages in non-bodily injury cases, but that judges had higher punitive-compensatory damages ratios in cases involving bodily injury than did juries.<sup>79</sup>

Hersch and Viscusi analyzed the CTCN dataset using different specifications.<sup>80</sup> They found that juries were more likely to award punitive damages than were judges<sup>81</sup> and that juries awarded higher amounts in compensatory<sup>82</sup> and punitive<sup>83</sup> damages than did judges, though this effect appeared to be limited to the high end of the range of awards.<sup>84</sup> They attribute the differences between their findings and those of Eisenberg and his colleagues to differences in specifications in the models used—specifically whether the county of the trial was controlled and the inclusion (in the Eisenberg model) of an interaction term.

Other studies have examined bench and jury trials in particular types of cases. Helland and Tabarrok used data on personal injury cases tried between 1988 and 1996 and reported in *Jury Verdict Re-*

75. *Id.* at 760, 774.

76. *Id.* (finding that juries were more likely to award punitive damages when compensatory damages were low and that judges were more likely to award punitive damages when compensatory damages were high).

77. *Id.* at 775-76 (significant results dependent on inclusion of a particular county). However, the awards made by juries and judges were equally likely to be so extreme as to be outside the range in which 95% of the judges' awards fell. *Id.*

78. Theodore Eisenberg et al., *Reconciling Experimental Incoherence with Real-World Coherence in Punitive Damages*, 54 STAN. L. REV. 1239, 1248-49 (2002).

79. *Id.* at 1253-56.

80. Joni Hersch & W. Kip Viscusi, *Punitive Damages: How Judges and Juries Perform*, 33 J. LEGAL STUD. 1 (2004).

81. *Id.* at 26.

82. *Id.* at 29.

83. *Id.* at 24.

84. *Id.* at 28.

search's *Personal Injury Verdicts and Settlements*.<sup>85</sup> They found that, overall, plaintiffs fared better on liability decisions in front of judges, while juries awarded more in damages in cases in which plaintiffs won.<sup>86</sup> They tested the influence of several factors on the damage awards of juries and judges and found several differences: caps on punitive damages were associated with smaller awards in jury trials, but not in bench trials; the collateral source rule was associated with larger awards in jury trials, but not in bench trials; and trials held in counties with higher poverty rates were associated with higher awards in jury trials, but smaller awards in bench trials.<sup>87</sup>

Using data from the Administrative Office of the United States Courts, Moore analyzed the 1209 federal patent cases that were resolved by either a bench trial or a jury trial between 1983 and 1999.<sup>88</sup> Overall, she found that juries were more likely to find for the patent holder than for the patent infringer, while judges were equally likely to find for either party.<sup>89</sup> Additional differences were found in complex cases. When both the validity of the patent and whether it had been infringed were raised within a case, juries (86%) were more likely than were judges (74%) to find for the same party on both issues.<sup>90</sup> Similarly, when the infringement of multiple patents was raised in a single case, juries (87%) were more likely than were judges (72%) to find for the same party.<sup>91</sup> Overall, juries awarded more in damages than did judges, although both groups awarded more than \$5 million in similar proportions of the cases they heard.<sup>92</sup> Affirmance rates on appeal for bench and jury trials were virtually identical.<sup>93</sup>

These studies show a great deal of similarity in the decisions made by judges and juries, as well as some notable differences. Im-

85. Eric Helland & Alexander Tabarrok, *Runaway Judges? Selection Effects and the Jury*, 16 J.L. ECON. & ORG. 306, 309-10 (2000); see Vidmar, *supra* note 17, at 600 (describing difficulties with using verdict report data).

86. Helland & Tabarrok, *supra* note 85, at 310.

87. *Id.* at 326-28.

88. Kimberly A. Moore, *Judges, Juries, and Patent Cases—An Empirical Peek Inside the Black Box*, 99 MICH. L. REV. 365, 380 (2000).

89. *Id.* at 386 (juries found for the patent holder in 68% of cases; judges found for the patent holder in 51% of cases). Moore also analyzed the data by substantive issue, comparing outcomes on each of four substantive issues: validity; enforceability; infringement; and willfulness. She found that both juries and judges were more likely to find patents valid than invalid, more likely to find patents enforceable than unenforceable, and more likely to find patents infringed than not infringed. Only juries were more likely to find patents willfully infringed than not willfully infringed. *Id.* at 390-91. Overall, juries were more likely to find patents valid, infringed, and willfully infringed than were judges; judges and juries were equally likely to find patents enforceable. *Id.* at 390.

90. *Id.* at 403.

91. *Id.* at 404. Whether the patent holder or the infringer files suit was a significant predictor of who wins in jury trials, but not bench trials. *Id.* at 405.

92. *Id.* at 395 (17% of bench trials and 21% of jury trials).

93. *Id.* at 397 (78% affirmance rate for both).

portantly, there is evidence that judges' and juries' decisions are correlated and that they consider the same or similar factors when making decisions. There is mixed evidence about whether judges or juries give higher damage awards. The reality is likely more nuanced than a generalized finding that one decisionmaker awards more than the other. Research in this area should focus on identifying the conditions under which the damage award decisions of juries and judges depart from each other.

It is difficult, however, to make these nuanced conclusions about the substantive differences in the decisionmaking of judges and jurors from these studies. There is evidence in the studies that juries and judges are faced with differing case mixes. For example, Clermont and Eisenberg found substantial differences in the rates at which different types of cases are heard by judges and juries.<sup>94</sup> It is quite possible that strategic decisions by litigants result in collections of cases with different attributes going before jurors and judges.<sup>95</sup> Clermont and Eisenberg concluded that their data are explained best by "small differences between judges' and juries' treatment of cases and, more importantly, in the parties' varying the selection of cases that reach judge and jury," and they argue that these selection effects are driven by popular misperceptions of the differences between juries and judges.<sup>96</sup> Similarly, Helland and Tabarrok concluded that a large proportion of the differences they found between juries and judges is explained by differences in the cases faced.<sup>97</sup> Thus, care must be used in drawing conclusions about the differences and similarities between judges and juries on the basis of these studies alone—there are likely factors other than the identity of the factfinder that influenced the results. Nonetheless, the findings of these studies provide useful insight into the comparison and can be read in conjunction with the results of other studies using different methodologies.

## V. EXPERIMENTAL STUDIES

Both judge-jury agreement studies and archival studies examine the responses of decisionmakers to real-world cases. The judge-jury agreement studies have the advantage of examining reactions to

---

94. Clermont & Eisenberg, *supra* note 6, at 1141 (finding that juries are more likely to hear personal injury cases and that judges are more likely to hear non-personal injury cases).

95. *See id.* (discussing selection effects); Helland & Tabarrok, *supra* note 85 (same). *See generally* George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1 (1984).

96. Clermont & Eisenberg, *supra* note 6, at 1126.

97. Helland & Tabarrok, *supra* note 85, at 330 (finding two-thirds to three-fourths of the difference in awards were explained by differences in the cases).

identical cases, but the judges and juries are acting in different roles. The archival studies examine the decisions of both groups when they are charged with ultimate decisionmaking responsibility, but the cases decided by the two groups are not identical. Experimental studies that ask jurors and judges to evaluate identical case facts allow for a comparison between the two sets of decisionmakers without the confound of different case mixes. Experimental methods also allow the examination of processes that underlie ultimate verdicts, such as whether the decisionmaker made appropriate use of evidence or legal rules.

Ideally, an experimental comparison between judges and juries would compare the responses of a sample of judges to the responses of a sample of mock juries. Very few, if any, studies fit this ideal model. A more common approach is to compare the responses of judges and individual jurors, but there are also relatively few studies of this type.<sup>98</sup> In addition, there are studies of judicial decisionmaking and studies of juror decisionmaking that have been conducted in similar domains and can be compared. These experimental studies, like the studies reviewed above, show a high degree of similarity in the decisionmaking of judges and jurors.

#### A. *Consideration of the Evidence*

One domain in which the decisionmaking of both jurors and judges has been examined is how each uses the evidence at trial. In general, the research has shown that despite the many extralegal influences on juror decisionmaking, the trial evidence is the best predictor of verdicts.<sup>99</sup> However, some types of evidence have proven to be more difficult to deal with than others.

##### 1. *Scientific, Expert, and Statistical Evidence*

Jurors are often criticized as unable to effectively evaluate expert, scientific, or statistical evidence. A recent review of the empirical research examining jurors' use of expert testimony concludes that "[a]lthough jurors struggle and are occasionally misled, they generally make reasonable use of complex material, utilizing the expert

---

98. For a discussion of potential differences between individual jurors and deliberating juries, see *infra* Part V.E.

99. See HANS & VIDMAR, *supra* note 2, at 154; MICHAEL J. SAKS & REID HASTIE, *SOCIAL PSYCHOLOGY IN COURT* 68 (1978) ("Indeed, the power of evidence is so well recognized by jury researchers that when studying processes other than evidence, they must calibrate the evidence to be moderate so that it leaves some variance to be influenced by the variables under study."); Christy A. Visher, *Juror Decision Making: The Importance of Evidence*, 11 *LAW & HUM. BEHAV.* 1 (1987).

testimony when it is presented in a form that they can use.”<sup>100</sup> There are, nonetheless, particular types of evidence—scientific and statistical evidence—with which jurors appear to have the most difficulty.<sup>101</sup>

However, there is little reason to think that judges are any better equipped to contend with complex scientific or statistical evidence. The educational path of most judges—law school—does not emphasize training in scientific research methods or statistical reasoning.<sup>102</sup> Lehman, Lempert, and Nisbett compared the statistical and methodological reasoning of graduate students in a variety of disciplines.<sup>103</sup> They found that despite few initial differences across disciplines, students in psychology and medicine improved dramatically in their ability to engage in statistical-methodological reasoning in the first three years of graduate school, while students in law did not.<sup>104</sup>

Not surprisingly then, in a national survey of state trial court judges, only 52% of participating judges felt that “their education

100. Neil Vidmar & Shari Seidman Diamond, *Juries and Expert Evidence*, 66 BROOK. L. REV. 1121, 1166-67 (2001).

101. Jurors appear to have difficulty in assessing empirical research. See, e.g., Margaret Bull Kovera et al., *Reasoning About Scientific Evidence: Effects of Juror Gender and Evidence Quality on Juror Decisions in a Hostile Work Environment Case*, 84 J. APPLIED PSYCHOL. 362, 367 (1999) (finding that jurors’ assessments of an empirical study of sexual harassment (presented as part of a sexual harassment trial) were not sensitive to differences in how sexual harassment was measured). Similarly, there is evidence that jurors have difficulty with various types of statistical evidence. See, e.g., David L. Faigman & A.J. Baglioni, Jr., *Bayes’ Theorem in the Trial Process: Instructing Jurors on the Value of Statistical Evidence*, 12 LAW & HUM. BEHAV. 1 (1988); Jonathan J. Koehler et al., *The Random Match Probability in DNA Evidence: Irrelevant and Prejudicial?*, 35 JURIMETRICS J. 201 (1995); Jason Schklar & Shari Seidman Diamond, *Juror Reactions to DNA Evidence: Errors and Expectancies*, 23 LAW & HUM. BEHAV. 159 (1999); William C. Thompson & Edward L. Schumann, *Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor’s Fallacy and the Defense Attorney’s Fallacy*, 11 LAW & HUM. BEHAV. 167 (1987). These results are consistent with a body of evidence that laypersons generally have difficulty with statistical reasoning and with identifying flaws in empirical research. See, e.g., RULES FOR REASONING (Richard E. Nisbett ed., 1993); Christopher Jepson et al., *Inductive Reasoning: Competence or Skill?*, 6 BEHAV. & BRAIN SCI. 494 (1983); Richard E. Nisbett et al., *Teaching Reasoning*, 238 SCIENCE 625 (1987) [hereinafter Nisbett et al., *Teaching Reasoning*]; Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCIENCE 1124 (1974).

102. See Lee Epstein & Gary King, *The Rules of Inference*, 69 U. CHI. L. REV. 1 (2002) (arguing that law schools ought to teach statistics/scientific methodology); Michael Heise, *The Importance of Being Empirical*, 26 PEPP. L. REV. 807 (1999); see also Lempert, *supra* note 9, at 91.

Just as a jury may not include people likely to comprehend the details of complex cases, so a judge, assigned at random, may not be well-equipped to cope with complexity. Many judges are bright and diligent, but others are much less so, and a legal education does not mean that one will have more than a layperson’s understanding of such subjects as economics and statistics.

*Id.*

103. Darrin R. Lehman et al., *The Effects of Graduate Training on Reasoning: Formal Discipline and Thinking About Everyday-Life Events*, 43 AM. PSYCHOLOGIST 431 (1988).

104. *Id.* at 437.

had adequately prepared them to deal with the range of scientific evidence proffered in their courtrooms”; the other 48% felt inadequately prepared to deal with such issues.<sup>105</sup> While 63% reported having had “CLE training about the use of specific types of scientific evidence in the courtroom,” virtually all of the respondents (96%) indicated that they “had *not* received instruction about general scientific methods and principles.”<sup>106</sup> Indeed, recent research provides evidence that judges, like jurors, sometimes have difficulty with assessing the quality of scientific research methods and with statistical reasoning.

(a) *Scientific Evidence*

Gatowski and her colleagues recently conducted a national survey of 400 state trial court judges designed, in part, to assess whether trial court judges are able to operationalize the scientific concepts that they are required to use as gatekeepers under the *Daubert v. Merrell Dow Pharmaceuticals, Inc.*<sup>107</sup> line of cases.<sup>108</sup> Under *Daubert*, trial court judges are required to make “a preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid” using factors such as whether the theory is falsifiable (that is, whether it “can be (and has been) tested”); whether it has been “subjected to peer review and publication”; the “known or potential rate of error”;<sup>109</sup> and the “general acceptance” of the theory in the “relevant scientific community.”<sup>110</sup> The survey asked judges to describe how they would use these concepts in evaluating whether to admit scientific evidence.<sup>111</sup>

Judges were first asked about the concept of falsifiability. While 88% of judges indicated that falsifiability was a useful factor in evaluating scientific evidence, only 6% of judges articulated a description of how they would use the concept to evaluate evidence that indicated a clear understanding of the concept, 59% gave questionable answers, and 35% of judges gave responses that indicated a lack of understanding of the concept.<sup>112</sup> Judges’ responses to a question about error rate evidenced a similar lack of understanding. While 91% of the judges in the sample reported finding a consideration of error rate as a useful factor in evaluating scientific evidence, only 4%

---

105. Sophia I. Gatowski et al., *Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-Daubert World*, 25 LAW & HUM. BEHAV. 433, 442 (2001).

106. *Id.*

107. 509 U.S. 579 (1993).

108. Gatowski et al., *supra* note 105, at 433.

109. *Daubert*, 509 U.S. at 592-94.

110. *Id.* at 594 (quoting *United States v. Downing*, 753 F.2d 1224, 1238 (3d Cir. 1985)).

111. Gatowski et al., *supra* note 105.

112. *Id.* at 444-45.

of them explained how they would use the concept in a way that evidenced a clear understanding, 86% gave questionable answers, and 10% gave answers that evidenced a lack of understanding of the concept.<sup>113</sup> Most judges gave responses that indicated a clear understanding of the use of peer review and publication (71%) and of general acceptance (82%), but some judges had difficulty with these concepts as well (for example, 10% of judges evidenced a lack of understanding of the concept of peer review and publication).<sup>114</sup>

Experimental evidence also supports the notion that judges have difficulty assessing scientific evidence. Kovera and McAuliff asked 144 Florida trial court judges to evaluate the admissibility of expert testimony in a simulated sexual harassment case.<sup>115</sup> They varied the description of the study upon which the testimony was based to compare an internally valid version of the study with three methodologically flawed studies.<sup>116</sup> They found that these methodological flaws did not significantly influence trial court judges' decisions about the admissibility of the study as evidence.<sup>117</sup> In addition, when asked to justify their admissibility decisions, few judges mentioned the methodological flaws.<sup>118</sup>

There is also evidence that judges are susceptible to another common difficulty that laypersons have with assessing scientific research—biased assimilation. Biased assimilation is the tendency for evaluations of the methodology and persuasiveness of empirical research to be influenced by the extent to which the results of the research are consistent with the attitudes or expectations of the person doing the evaluation.<sup>119</sup> Thus, “judgments about the validity, reliability, relevance, and sometimes even the meaning of proffered evidence are biased by the apparent consistency of that evidence with the perceiver's theories and expectations.”<sup>120</sup> Redding and Reppucci found a pattern of biased assimilation for state court judges and law students who were asked to make judgments about social science research on

113. *Id.* at 445-47.

114. *Id.* at 447-48.

115. Margaret Bull Kovera & Bradley D. McAuliff, *The Effects of Peer Review and Evidence Quality on Judge Evaluations of Psychological Science: Are Judges Effective Gatekeepers?*, 85 J. APPLIED PSYCHOL. 574, 577 (2000).

116. *Id.* at 578. The methodological flaws included experimenter bias, lack of a control group, or a confound. *Id.*

117. *Id.* at 580; see also Margaret Bull Kovera et al., *Assessment of the Commonsense Psychology Underlying Daubert: Legal Decision Makers' Abilities to Evaluate Expert Evidence in Hostile Work Environment Cases*, 8 PSYCHOL. PUB. POL'Y & L. 180 (2002). Findings for attorneys were similar. *Id.* at 187-89.

118. Kovera & McAuliff, *supra* note 115, at 581.

119. See Charles G. Lord et al., *Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence*, 37 J. PERSONALITY & SOC. PSYCHOL. 2098 (1979).

120. *Id.* at 2099.

capital punishment and deterrence.<sup>121</sup> The results of the empirical studies presented to the U.S. Supreme Court in *Furman v. Georgia*<sup>122</sup> were varied so that they either supported or failed to support the death penalty. Participants judged the research to be more relevant to the legal issues, more likely to be admissible, and to have greater dispositive weight when the results were consistent with their personal attitudes about the death penalty than when the results contradicted those attitudes.<sup>123</sup>

(b) *Statistical Evidence*

Similarly, there is evidence that judges, like jurors, have difficulty with statistical evidence. Recently, Guthrie, Rachlinski, and Wistrich asked 167 federal magistrate judges participating in a workshop by the Federal Judicial Center to respond to a series of vignettes. One of these vignettes was designed to explore “whether judges would commit the inverse fallacy.”<sup>124</sup> Judges were given a *res ipsa loquitur* problem in which the plaintiff was injured and the following facts were determined: (1) when the defendant is negligent there is a 90% chance of injury, (2) when the defendant is not negligent there is a 1% chance of injury, and (3) the defendant is negligent one in 1000 times. Judges were asked to indicate the likelihood that the plaintiff was injured due to the negligence of the defendant (the correct answer is 8.3%) by selecting from the categories 0-25%, 26-50%, 51-75%, or 76-100%.<sup>125</sup> Approximately 60% of the judges answered incorrectly, with most of them selecting the highest category (76-100%).<sup>126</sup> Thus, judges, like jurors, have difficulty using base rate information.

In one of the few studies that has directly compared mock jurors and judges, Wells examined the reactions of laypersons and judges to “naked statistical evidence” in the context of the Blue Bus problem.<sup>127</sup>

---

121. Richard E. Redding & N. Dickon Reppucci, *Effects of Lawyers' Socio-Political Attitudes on Their Judgments of Social Science in Legal Decision Making*, 23 *LAW & HUM. BEHAV.* 31 (1999).

122. 408 U.S. 238 (1972).

123. Redding & Reppucci, *supra* note 121, at 43. The effects were more pervasive for law students (evident in four different ratings) than for judges (evident in ratings of dispositive weight only). *Id.* at 44.

124. Chris Guthrie et al., *Inside the Judicial Mind*, 86 *CORNELL L. REV.* 777, 808 (2001). “The inverse fallacy refers to the tendency to treat the probability of a hypothesis given the evidence . . . as the same as, or close to, the probability of the evidence given the hypothesis . . .” *Id.* at 807.

125. *Id.* at 808.

126. *Id.* at 809.

127. Gary L. Wells, *Naked Statistical Evidence of Liability: Is Subjective Probability Enough?*, 62 *J. PERSONALITY & SOC. PSYCHOL.* 739 (1992). Wells describes the problem as follows:

An accident occurs in which it is known that a bus was at fault, but the specific company is not known. Because the Blue Bus Company accounts for most of

In two different studies, Wells found that the subjective probabilities and verdict patterns of judges, undergraduate students, and MBA students were similar. Across several different versions of the scenario, all three groups made similar subjective assessments of the likelihood that the accident was caused by a Blue Bus. Moreover, all three groups demonstrated a reluctance to find liability on the basis of naked statistical evidence.<sup>128</sup>

## 2. *Inadmissible Evidence*

In addition to concern about their ability to deal effectively with scientific and statistical evidence, jurors are commonly criticized for an inability to ignore irrelevant or inadmissible evidence.<sup>129</sup> However, there is evidence that judges, like ordinary citizens, are unable to ignore inadmissible evidence. Landsman and Rakos compared the responses of eighty-eight judges attending an annual judicial conference and 104 prospective jurors to three versions of a hypothetical vignette.<sup>130</sup> The vignettes differed only in how a particular item of evidence that favored the plaintiff was treated. In one version, the evidence was described and participants were informed of a judicial determination to admit the evidence. In a second version, the evidence was described and participants were told that a judge had decided that the evidence was inadmissible. Control participants were not told about the evidence.<sup>131</sup> Landsman and Rakos found that both judges and jurors who heard the evidence were more likely to find the defendant liable, whether or not they were told that the evidence was inadmissible.<sup>132</sup> Interestingly, Landsman and Rakos found that while the effect of the biasing information was the same for judges

---

the business (say 80%), the balance of probability (in this case 80%) clearly favors the idea that a Blue Bus Company bus caused this accident. . . .

The problem that has plagued the courts and legal scholars is that they at once endorse the balance of probability criterion for such suits and also refuse to rule in favor of the plaintiff when such evidence is presented. Indeed, suits based on naked statistics of this sort are usually thrown out by a summary judgment.

*Id.* at 740 (citations omitted).

128. *Id.* at 744, 747.

129. See Mark Kelman et al., *Context-Dependence in Legal Decision Making*, 25 J. LEGAL STUD. 287 (1996); Stanley Sue et al., *Effects of Inadmissible Evidence on the Decisions of Simulated Jurors: A Moral Dilemma*, 3 J. APPLIED SOC. PSYCHOL. 345 (1973); William C. Thompson et al., *Inadmissible Evidence and Juror Verdicts*, 40 J. PERSONALITY & SOC. PSYCHOL. 453 (1981); Sharon Wolf & David A. Montgomery, *Effects of Inadmissible Evidence and Level of Judicial Admonishment to Disregard on the Judgments of Mock Jurors*, 7 J. APPLIED SOC. PSYCHOL. 205 (1977).

130. Stephan Landsman & Richard F. Rakos, *A Preliminary Inquiry into the Effect of Potentially Biasing Information on Judges and Jurors in Civil Litigation*, 12 BEHAV. SCI. & L. 113 (1994).

131. *Id.* at 120-22.

132. *Id.* at 122-23.

and jurors, the jurors appeared to be more sensitive to their cognitive limitations in disregarding the evidence than did the judges.<sup>133</sup>

More recently, Wistrich, Guthrie, and Rachlinski examined judges' ability to ignore inadmissible evidence in a series of experimental studies.<sup>134</sup> Even when they ruled the evidence inadmissible, judges were unable to ignore the demands made in settlement discussions when setting damage awards, an admission that was protected by attorney-client privilege in determining contract liability, the sexual history of the victim in a sexual assault trial, and the prior criminal conviction of the plaintiff. Judges were, however, able to ignore the results of a search when deciding whether to suppress evidence and an inadmissible confession.<sup>135</sup>

### *B. Damages*

One area, in particular, in which jury decisionmaking has been criticized is the setting of damage awards—particularly the jury's ability to make awards of noneconomic and punitive damages. Consistent with the archival research described above, experimental studies comparing the noneconomic damage awards made by judges and jury-eligible persons primarily show similarities in the damages awarded. Vidmar and Rice asked eighty-nine prospective jurors waiting for jury duty and twenty-one arbitrators to determine noneconomic damages in a medical malpractice case in which the injuries to the female plaintiff resulted in permanent scarring.<sup>136</sup> They found no differences in the amounts potential jurors and arbitrators awarded for pain and suffering.<sup>137</sup> Awards by both jurors and arbitrators for disfigurement were significantly correlated with their beliefs about the degree to which the scar was disfiguring, affected the quality of the plaintiff's life, embarrassed the plaintiff, and was worse than a similar scar on a man.<sup>138</sup> The awards made by both jurors and arbitrators were highly variable, although jurors' awards were more variable than awards made by the arbitrators.<sup>139</sup>

Similarly, Wissler, Hart, and Saks compared how jury-eligible citizens, judges, plaintiffs' attorneys, and defense attorneys make

---

133. *Id.* at 125.

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

135. *Id.*

136. Neil Vidmar & Jeffrey J. Rice, *Assessments of Noneconomic Damage Awards in Medical Negligence: A Comparison of Jurors with Legal Professionals*, 78 IOWA L. REV. 883, 890-91 (1993).

137. *Id.* at 893.

138. *Id.* at 894-95.

139. *Id.* at 893.

pain and suffering awards in personal injury cases.<sup>140</sup> Some overall differences emerged: the awards made by the laypersons were higher and more variable than those made by the judges, and laypersons rated the injuries as more severe than did judges.<sup>141</sup> However, the similarities were striking. They found judges and laypersons were similarly influenced by the characteristics of the injuries studied.<sup>142</sup> The attributes of the injuries influenced both groups in their assessments of the severity of the injuries and in their dollar awards. The dollar awards of both laypersons and judges were somewhat less predictable than their ratings of injury severity, though this was particularly true for laypersons. In addition, the dollar amounts of the two groups were highly correlated, suggesting that they ranked the cases similarly, awarding higher amounts in the same cases.<sup>143</sup>

A number of experimental studies have now examined the punitive damage award decisions of jurors and judges.<sup>144</sup> In one study, Viscusi asked participants at a judicial conference on law and economics as well as laypersons to determine whether a court should award punitive damages in response to several different versions of a hypothetical case.<sup>145</sup> Both judges and laypersons were more likely to think that punitive damages were appropriate as the damages increased; however, judges were less likely than laypersons to think an award of punitive damages was appropriate.<sup>146</sup> However, if Viscusi is correct that in the scenarios described “the firm should not be found negligent, much less be punished with a punitive damages award,”<sup>147</sup> then it is notable that as many as 70% of the judges indicated that they thought punitive damages were appropriate.<sup>148</sup> In a second study using a different scenario, Viscusi found that although all the judges awarded some amount of damages while 5% of jurors awarded nothing, damage awards by jurors were higher and more variable than awards made by judges.<sup>149</sup> Hastie and Viscusi asked citizens and judges to determine whether the defendant in a hypothetical case should be required to pay punitive damages. Citizens appeared

---

140. Roselle L. Wissler et al., *Decisionmaking About General Damages: A Comparison of Jurors, Judges, and Lawyers*, 98 MICH. L. REV. 751 (1999).

141. *Id.* at 798-99.

142. *Id.* at 785-87. Participants rated the degree of disability, mental suffering, disfigurement, and pain caused by an injury. For both groups disability had greatest impact, followed by mental suffering and then disfigurement—pain had no independent significant impact. *Id.*

143. *Id.* at 799.

144. For a more detailed review, see Robbennolt, *supra* note 2, at 146-58.

145. W. Kip Viscusi, *Jurors, Judges, and the Mistreatment of Risk by the Courts*, 30 J. LEGAL STUD. 107 (2001); *see also* Viscusi, *supra* note 3.

146. Viscusi, *supra* note 145, at 112-14.

147. *Id.* at 111.

148. *Id.* at 114 (in the personal injury/plane crash scenario).

149. *Id.* at 129. The type of damages participants were to award was not specified.

to be more likely than judges to believe that punitive damages should be paid.<sup>150</sup>

In another experimental study, Robbennolt found that trial court judges and laypersons considered similar factors when determining punitive damage awards.<sup>151</sup> Both jury-eligible laypersons and trial court judges made higher punitive damage awards when the injury to the plaintiff was more severe, when the defendant's conduct risked more severe harm, and when the defendant was wealthier.<sup>152</sup> In addition, the punitive damage awards of the judges and the laypersons were similar in magnitude and were equally variable.<sup>153</sup>

One particular aspect of damage award decisionmaking that has received a great deal of attention is the frequently replicated finding that jurors are vulnerable to cognitive illusions such as anchoring in determining damage awards. Jurors have been shown to anchor on the amount of the *ad damnum*,<sup>154</sup> on the amount of a damage cap,<sup>155</sup> or on the amount awarded in a similar case as reported in a newspaper story.<sup>156</sup>

More recently, experimental research has demonstrated that judges too are vulnerable to cognitive illusions such as anchoring. In their study of judicial decisionmaking described above, Guthrie, Rachlinski, and Wistrich asked judges to award compensatory damages in a personal injury suit.<sup>157</sup> The suit was in federal court based on diversity jurisdiction and was described such that the damages clearly exceeded the \$75,000 threshold. However, prior to making

150. Hastie & Viscusi, *supra* note 3, at 919; *see also* Viscusi, *supra* note 3, at 42 (discussing the same data).

151. Robbennolt, *supra* note 2.

152. *Id.* at 119-25.

153. *Id.* at 149. Consistent with prior studies, the compensatory damages of jurors were marginally higher than those of judges and were more variable. *Id.* at 151.

154. *See* Gretchen B. Chapman & Brian H. Bornstein, *The More You Ask for, the More You Get: Anchoring in Personal Injury Verdicts*, 10 APPLIED COGNITIVE PSYCHOL. 519, 527, 531 (1996); Reid Hastie et al., *Juror Judgments in Civil Cases: Effects of Plaintiff's Requests and Plaintiff's Identity on Punitive Damage Awards*, 23 LAW & HUM. BEHAV. 445 (1999); Verlin B. Hinsz & Kristin E. Indahl, *Assimilation to Anchors for Damage Awards in a Mock Civil Trial*, 25 J. APPLIED SOC. PSYCHOL. 991, 1005, 1010, 1013-15 (1995); John Malouff & Nicola S. Schutte, *Shaping Juror Attitudes: Effects of Requesting Different Damage Amounts in Personal Injury Trials*, 129 J. SOC. PSYCHOL. 491, 494 (1989); Barry Markovsky, *Anchoring Justice*, 51 SOC. PSYCHOL. Q. 213 (1988). For a description of anchoring and adjustment generally, *see* Tversky & Kahneman, *supra* note 101, at 1128-30.

155. *See* Jennifer K. Robbennolt & Christina A. Studebaker, *Anchoring in the Courtroom: The Effects of Caps on Punitive Damages*, 23 LAW & HUM. BEHAV. 353 (1999); Michael J. Saks et al., *Reducing Variability in Civil Jury Awards*, 21 LAW & HUM. BEHAV. 243, 252 (1997).

156. W. Kip Viscusi, *The Challenge of Punitive Damages Mathematics*, 30 J. LEGAL STUD. 313, 331 (2001). For a general discussion of the effects of the media on legal decisionmaking, *see* Jennifer K. Robbennolt & Christina A. Studebaker, *News Media Reporting on Civil Litigation and Its Influence on Civil Justice Decision Making*, 27 LAW & HUM. BEHAV. 5 (2003).

157. Guthrie et al., *supra* note 124, at 791-92.

their damage awards, half of the judges were asked to rule on the defendant's motion to dismiss the case on the ground that the \$75,000 threshold had not been met. While nearly all of the participants who were asked to rule on the motion denied it, the judges who ruled on the motion awarded less in compensatory damages than the judges who were not asked to rule on the motion.<sup>158</sup> The authors concluded that the judges had anchored on the \$75,000 threshold in arriving at their damage awards.<sup>159</sup> Similarly, as noted above, Wistrich, Guthrie, and Rachlinski found that judges' damage awards were influenced by their knowledge of demands made in settlement negotiations.<sup>160</sup>

One other difference between judges and jurors has been found to influence damage awards. Trial court judges are likely to observe a broader range of cases than the typical juror. This may result in trial court judges and laypersons having different perceptions and expectations about the legal system in general and about civil litigation in particular. Wissler and her colleagues have observed that in their studies of noneconomic damages, damage awards are more correlated with case or injury attributes when participants are asked to judge multiple cases or injuries.<sup>161</sup> They suggest that judges may possess "a more complete cognitive reference scale of injuries to provide a context for thinking about the case now under consideration."<sup>162</sup> Several studies have found a connection between perceptions of the civil litigation system and damage awards. Robbennolt found that trial court judges made smaller and more accurate estimates of the percentage of jury awards that are greater than \$1 million and of the percentage of jury awards that include punitive damages than did laypersons.<sup>163</sup> The magnitude of these estimates was related to participants' compensatory damage awards; however, they were not related to participants' punitive damage awards.<sup>164</sup>

### C. Risk Assessment

There is evidence that jurors are vulnerable to cognitive illusions relevant to assessing risk, such as outcome bias (the tendency to judge decision processes and decisionmaking competence more fa-

---

158. *Id.*

159. *Id.* For evidence that lawyers are susceptible to anchoring as well, see Russell Korobkin & Chris Guthrie, *Opening Offers and Out-of-Court Settlement: A Little Moderation May Not Go a Long Way*, 10 OHIO ST. J. ON DISP. RESOL. 1, 18-19 (1994) (discussing the anchoring effect of the first offer in settlement negotiation).

160. Wistrich et al., *supra* note 134.

161. Wissler et al., *supra* note 140, at 808 n.166.

162. *Id.*

163. Robbennolt, *supra* note 2, at 157.

164. *Id.*

vorably when the outcome of the decision turns out to be positive)<sup>165</sup> and hindsight bias (unconsciously overestimating the likelihood one would have assigned to an event once the outcome is known).<sup>166</sup> However, there is evidence that judges, too, are vulnerable to cognitive illusions such as the outcome and hindsight biases.<sup>167</sup>

First, there is evidence that judges are susceptible to outcome bias. Anderson and his colleagues asked sixty-five state and federal judges and fifty-eight auditors to evaluate the correctness of an audit decision.<sup>168</sup> Those participants who were told that the decision resulted in a negative outcome rated the decision more negatively than did those participants who were told that the decision resulted in a positive outcome.<sup>169</sup>

Second, Guthrie, Rachlinski, and Wistrich, in their study of magistrate judges, demonstrated that judges can fall prey to the hindsight bias.<sup>170</sup> Judges read a scenario that detailed the facts of a pro se

165. Jonathan Baron & John C. Hershey, *Outcome Bias in Decision Evaluation*, 54 J. PERSONALITY & SOC. PSYCHOL. 569, 570 (1988).

166. See generally Baruch Fischhoff, *Hindsight ≠ Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty*, 1 J. EXPERIMENTAL PSYCHOL.: HUM. PERCEPTION & PERFORMANCE 288, 288 (1975). For studies demonstrating these phenomena in jurors, see, for example, Reid Hastie et al., *Juror Judgments in Civil Cases: Hindsight Effects on Judgments of Liability for Punitive Damages*, 23 LAW & HUM. BEHAV. 597 (1999); Kim A. Kamin & Jeffrey J. Rachlinski, *Ex Post ≠ Ex Ante: Determining Liability in Hindsight*, 19 LAW & HUM. BEHAV. 89 (1995); Susan J. LaBine & Gary LaBine, *Determinations of Negligence and the Hindsight Bias*, 20 LAW & HUM. BEHAV. 501 (1996); Philip G. Peters, Jr., *Hindsight Bias and Tort Liability: Avoiding Premature Conclusions*, 31 ARIZ. ST. L.J. 1277 (1999); Jennifer K. Robbenolt & Mark S. Sobus, *An Integration of Hindsight Bias and Counterfactual Thinking: Decision-Making and Drug Courier Profiles*, 21 LAW & HUM. BEHAV. 539 (1997); Merrie Jo Stallard & Debra L. Worthington, *Reducing the Hindsight Bias Utilizing Attorney Closing Arguments*, 22 LAW & HUM. BEHAV. 671 (1998). For a review of research on hindsight, see Jeffrey J. Rachlinski, *A Positive Psychological Theory of Judging in Hindsight*, 65 U. CHI. L. REV. 571, 576-90 (1998).

167. In addition to the areas reviewed here and above, there is also evidence that judges are vulnerable to other cognitive illusions—such as egocentric bias and framing—to which laypersons have been shown to be susceptible. Egocentric bias is a cognitive heuristic under which a decisionmaker perceives and interprets information and forms expectations in a manner that is favorable to the decisionmaker. See generally Michael Ross & Fiore Sicoly, *Egocentric Biases in Availability and Attribution*, 37 J. PERSONALITY & SOC. PSYCHOL. 322 (1979). For evidence that judges are susceptible to egocentric bias, see Theodore Eisenberg, *Differing Perceptions of Attorney Fees in Bankruptcy Cases*, 72 WASH. U. L.Q. 979 (1994) (researching bankruptcy judges and lawyers' fees); Guthrie et al., *supra* note 124, at 814 (researching judges' reversal rates). Framing refers to the evaluation of outcomes relative to a neutral reference point. Outcomes are evaluated differently depending on whether they are framed as gains or losses. See generally Daniel Kahneman & Amos Tversky, *Choices, Values, and Frames*, 39 AM. PSYCHOLOGIST 341 (1984). For evidence that judges are susceptible to framing effects, see Guthrie et al., *supra* note 124, at 797. For evidence that lawyers are susceptible to framing effects, see Linda Babcock et al., *Forming Beliefs About Adjudicated Outcomes: Perceptions of Risk and Reservation Values*, 15 INT'L REV. L. ECON. 289 (1995).

168. John C. Anderson et al., *Evaluation of Auditor Decisions: Hindsight Bias Effects and the Expectation Gap*, 14 J. ECON. PSYCHOL. 711 (1993).

169. *Id.* at 727.

170. Guthrie et al., *supra* note 124, at 799-805.

§ 1983 action in which the district court sanctioned the plaintiff and the plaintiff appealed. Judges were told that the appellate court had taken one of three actions: remanded so that a lesser sanction could be imposed, affirmed the sanction, or vacated the sanction. Judges were then given all three possible outcomes and asked to predict which of these decisions the court of appeals was most likely to have made.<sup>171</sup> Judges were more likely to predict the outcome that was consistent with the outcome that they were told occurred; that is, judges who were given a particular outcome were more likely to predict that outcome than judges given the other possible outcomes.<sup>172</sup>

Viscusi explored another possible difference in judges' and jurors' risk beliefs. Consistent with previous psychological research on risk perception,<sup>173</sup> Viscusi has shown that both judges and laypersons tend to overestimate small risks and to underestimate large risks, though these effects were smaller in magnitude for judges than they were for laypersons.<sup>174</sup> It is possible that differences in the ways that judges and laypersons conceptualize risk could influence the ways in which they assess legal cases. There is still much to be learned about whether and how judges and laypersons differ in their conceptions of risk and to what extent these conceptions impact their decisionmaking in the courtroom.

#### D. Criminal Cases

Several direct comparisons have been made of judges' and jurors' decisionmaking in criminal cases. Two studies of blame attributions have found that judges and laypersons are influenced by similar factors in criminal cases. Howe asked sixteen Missouri trial court judges and sixteen students to evaluate a set of scenarios based on two cases in which one person shot another.<sup>175</sup> Several versions of each case were created to reflect different degrees of intent to injure, the justification for the shooting, and the resulting injury. While students attributed more blame to the shooter than did judges across the cases,<sup>176</sup> the blame attributions of both students and judges were

171. *Id.* at 801-02.

172. *Id.* at 802-03.

173. See Paul Slovic et al., *Facts Versus Fears: Understanding Perceived Risk*, in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 463, 467 (Daniel Kahneman et al. eds., 1982); Sarah Lichtenstein et al., *Judged Frequency of Lethal Events*, 4 J. EXPERIMENTAL PSYCHOL.: HUM. LEARNING & MEMORY 551 (1978). But see HOWARD KUNREUTHER ET AL., DISASTER INSURANCE PROTECTION: PUBLIC POLICY LESSONS 168 (1978) (finding that some low risks are ignored).

174. Viscusi, *supra* note 145, at 131-32.

175. Edmund S. Howe, *Integration of Mitigation, Intention, and Outcome Damage Information*, by *Students and Circuit Court Judges*, 21 J. APPLIED SOC. PSYCHOL. 875 (1991).

176. *Id.* at 886. This was true, in particular, when justification for the shooting was higher. "[A]s compared with students, judges assigned markedly less blame for both cases when the level of mitigation was high . . ." *Id.*

similarly influenced by the level of intent, justification, and injury. For both groups, greater blame was attributed when the shooter acted more intentionally, acted with less justification, or caused a more severe injury.<sup>177</sup>

In a second study, Howe and Loftus asked thirteen Missouri circuit court judges and sixteen students to make blame judgments in a case in which a fight led to injury.<sup>178</sup> Consistent with the results reported above, while students assigned more blame than did judges overall, both groups of respondents assigned more blame when the resulting injury was more severe and when the defendant acted more intentionally.<sup>179</sup> Howe and Loftus also found that judges and students displayed “nearly identical pattern[s] of individual differences in integration processes.”<sup>180</sup> While neither of these studies asked participants to render verdicts or to recommend sentences, they do show similarities in the factors that both sets of decisionmakers take into account in assigning blame in criminal cases.

One experimental study has shown differences between judges and jurors in the severity of the sentences that they would assign to criminal defendants. Diamond and Stalans asked 116 Illinois State Court judges who were attending seminars on sentencing, 154 people who had reported for jury duty, and fifty-five students to assign sentences to offenders in four cases.<sup>181</sup> For each of the four cases, participants read a presentencing investigation report, watched a videotape of the sentencing hearing, and were provided with the range of available sentencing options.<sup>182</sup> Although 66% of the lay respondents indicated that they thought that Illinois judges were too lenient in sentencing,<sup>183</sup> lay respondents gave sentences that tended to be more lenient than those given by the judges.<sup>184</sup> The results suggesting that

177. *Id.* at 883, 885 (students and judges respectively).

178. Edmund S. Howe & Thomas C. Loftus, *Integration of Intention and Outcome Information by Students and Circuit Court Judges: Design Economy and Individual Differences*, 22 J. APPLIED SOC. PSYCHOL. 102 (1992).

179. *Id.* at 108.

180. *Id.* at 111.

181. Shari Seidman Diamond & Loretta J. Stalans, *The Myth of Judicial Leniency in Sentencing*, 7 BEHAV. SCI. & L. 73, 74-75 (1989).

182. *Id.* at 75.

183. *Id.* at 86.

184. Lay respondents gave more lenient sentences than judges in two of four cases. *Id.* at 78, 79 (Case 2 and Case 3 respectively). In the other two cases, lay respondents gave sentences that did not differ from those given by Cook County judges and were more lenient than those given by non-Cook County judges. *Id.* at 76-77, 80 (Case 1 and Case 4 respectively). Though not a direct comparison of juror and judges, some related evidence comes from a study by Roberts and Doob. Julian V. Roberts & Anthony N. Doob, *News Media Influences on Public Views of Sentencing*, 14 LAW & HUM. BEHAV. 451, 461-62 (1990) (finding that laypersons who learned about a criminal case by reading a summary of the sentencing hearing were less likely to believe that the judge had been too lenient and more likely to believe that the judge had considered all the appropriate factors than respondents who learned about the case by reading a media account); see also Loretta J. Stalans &

laypersons are more lenient in criminal sentencing are consistent with the results of Kalven and Zeisel and of Heuer and Penrod reported above.<sup>185</sup>

Diamond and Stalans found “few differences in the way judges and laypersons viewed these offenses and offenders” in the cases they studied.<sup>186</sup> In particular, they found that both groups gave similar ratings of offense seriousness, and their ratings of seriousness ordered the cases in the same way.<sup>187</sup> Both groups also made similar ratings of the importance of five different goals of sentencing<sup>188</sup> and made similar predictions of the probability of recidivism.<sup>189</sup>

### *E. Judges and Juries—The Effects of Group Decisionmaking*

One of the disadvantages of most of the simulation research done in this area is that it compares the decisionmaking of judges with individual jurors rather than deliberating juries. In 1964, Kalven argued that:

Often in the debate over the jury the capacity of *one* layman is compared to the capacity of one judge, as though this were the issue. The distinctive strength and safeguard of the jury system is that the jury operates as a group. Whether twelve lay heads are better than one judicial head is still open to argument, but it should be recognized that twelve lay heads are very probably better than one.<sup>190</sup>

His point is no less true forty years later. A trial court judge is typically engaged in making an individual decision, while jurors are allowed to make decisions collectively in groups.

This is not to say that the experimental comparisons between judges and individual jurors are not informative—they clearly are.

---

Shari Seidman Diamond, *Formation and Change in Lay Evaluations of Criminal Sentencing*, 14 LAW & HUM. BEHAV. 199, 211 (1990) (finding that laypersons' sentences were more lenient than required minimum sentence).

185. See KALVEN & ZEISEL, *supra* note 25; Heuer & Penrod, *Trial Complexity*, *supra* note 35; see also Robert J. MacCoun & Norbert L. Kerr, *Asymmetric Influence in Mock Jury Deliberation: Jurors' Bias for Leniency*, 54 J. PERSONALITY & SOC. PSYCHOL. 21 (1988).

186. Diamond & Stalans, *supra* note 181, at 86.

187. *Id.* at 81-82 n.9.

188. *Id.* at 82-85.

189. *Id.* at 82 n.12. There were some differences between the groups, each primarily limited to one of the cases. In one case, judges were less likely to believe that the offender was addicted to heroin. For a case involving the sale of cocaine, judges' sentences were more strongly predicted by the importance they placed on the general deterrent value of sentencing than were laypersons' sentences. For another case, involving a young offender, judges were more likely to believe that a prison sentence would increase the probability of recidivism. *Id.* at 83-85.

190. Kalven, *supra* note 4, at 1067; see also Phoebe C. Ellsworth, *Are Twelve Heads Better than One?*, LAW & CONTEMP. PROBS., Autumn 1989, at 205.

These studies demonstrate many ways in which jurors and judges share similar characteristics in their reasoning and decisionmaking that are likely to impact their decisions. In addition, it is relatively clear that the predeliberation preferences of individual jurors are highly predictive of the final jury determinations.<sup>191</sup> Nonetheless, in this context it is clear that group deliberation has the potential to have an impact on decisionmaking that has important implications for comparisons between juries and judges.

Unfortunately, there is no straightforward conclusion one can draw about how group deliberation will affect the comparison in all cases. There is evidence that deliberation can result in the attenuation of judgmental biases in some instances,<sup>192</sup> improve memory for trial evidence, increase complex reasoning about the evidence and arguments presented,<sup>193</sup> and reduce variability in decisions.<sup>194</sup> At the same time, there is evidence that group deliberation may exacerbate biases under certain conditions<sup>195</sup> and can result in more extreme

191. See, e.g., Reid Hastie et al., *A Study of Juror and Jury Judgments in Civil Cases: Deciding Liability for Punitive Damages*, 22 LAW & HUM. BEHAV. 287, 299 (1998); MacCoun & Kerr, *supra* note 185, at 21-22; Marla Sandys & Ronald C. Dillehay, *First-Ballot Votes, Predeliberation Dispositions, and Final Verdicts in Jury Trials*, 19 LAW & HUM. BEHAV. 175, 175 (1995); David Schkade et al., *Deliberating About Dollars: The Severity Shift*, 100 COLUM. L. REV. 1139, 1152-53 (2000); Sarah Tanford & Steven Penrod, *Jury Deliberations: Discussion Content and Influence Processes in Jury Decision Making*, 16 J. APPLIED SOC. PSYCHOL. 322, 323 (1986). See generally Michael J. Saks, *What Do Jury Experiments Tell Us About How Juries (Should) Make Decisions?*, 6 S. CAL. INTERDISC. L.J. 1, 37 (1997) (discussing the majority effect). Moreover, as noted above in the text accompanying note 23, it is not always necessary for the experimental setting to mirror the legal context for the findings to be generalizable to the legal setting. It is important to carefully consider whether, under what circumstances, and how a departure from the specifics of the legal setting is likely to affect the research findings. See Jennifer K. Robbennolt, *Evaluating Empirical Research Methods: Using Empirical Research in Law and Policy*, 81 NEB. L. REV. 777 (2002).

192. Martin F. Kaplan & Lynn E. Miller, *Reducing the Effects of Juror Bias*, 36 J. PERSONALITY & SOC. PSYCHOL. 1443 (1978). See generally Norbert L. Kerr et al., *Bias in Judgment: Comparing Individuals and Groups*, 103 PSYCHOL. REV. 687 (1996).

193. Monica L. McCoy et al., *The Effect of Jury Deliberations on Jurors' Reasoning Skills*, 23 LAW & HUM. BEHAV. 557, 560-61 (1999) (finding that mock-jurors who engaged in group deliberation considered more counterarguments and made more statements evaluating the evidence).

194. Shari Seidman Diamond et al., *Juror Judgments About Liability and Damages: Sources of Variability and Ways to Increase Consistency*, 48 DEPAUL L. REV. 301, 315-16 (1998) (finding damage awards by mock-juries less variable than those of individual mock-jurors).

195. See generally Kerr et al., *supra* note 192, at 715. In attempting to tease out conditions under which group deliberation is likely to facilitate or hinder unbiased decisionmaking, it is important to pay careful attention to a variety of characteristics of the task and decisionmaking process. Kerr, MacCoun, and Kramer predict that where juries operate under a "majority wins" type of decisionmaking process and where the evidence is close—as opposed to an extreme case, either for or against one party—group deliberation is predicted to exacerbate bias. *Id.* at 697-702. There is empirical evidence to support these predictions. See, e.g., Norbert L. Kerr et al., *Bias in Jurors vs Bias in Juries: New Evidence*

judgments, a phenomenon known as group polarization.<sup>196</sup> Additional research on the effect of jury deliberation generally is clearly warranted as is additional research specifically comparing judges with juries making group decisions.

#### VI. JUDICIAL DECISIONMAKING AS A BENCHMARK FOR EVALUATING JURY DECISIONMAKING

The research reviewed here clearly begins to fill a gap in the literature. While there has been a great deal of research exploring juror and jury decisionmaking and a separate literature examining influences on judicial decisionmaking,<sup>197</sup> there has been surprisingly little research comparing the decisions of juries and trial court judges. The most notable conclusion to be drawn from this emerging literature is that the decisionmaking of judges and jurors is strikingly similar. While there is evidence of some differences, there is a high degree of agreement between the groups, they appear to decide real cases quite similarly, and they show a great deal of similarity in responding to simulated cases designed to examine a variety of legal decisionmaking processes.

One problem with comparing the decisionmaking of juries to that of judges is that to the extent that there are differences in the decisionmaking of judges and juries, there is not always a clear benchmark for determining which group's decision is normatively better. Many of the studies reviewed here—either explicitly or implicitly—assume that judges are “right.” However, differences in the decisionmaking of judges and juries “do not speak for themselves.”<sup>198</sup> For example, even if a large number of studies converged on the finding that judges are more likely to find for plaintiffs than juries or that juries are likely to award higher damages than judges, it would not be clear in the abstract which outcome was more “correct.”<sup>199</sup> The dif-

---

from the SDS Perspective, 80 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 70 (1999).

196. See Martin J. Bourgeois et al., *Nominal and Interactive Groups: Effects of Preinstruction and Deliberations on Decisions and Evidence Recall in Complex Trials*, 80 J. APPLIED PSYCHOL. 58, 63 (1995); Shari Seidman Diamond & Jonathan D. Casper, *Blindfolding the Jury to Verdict Consequences: Damages, Experts, and the Civil Jury*, 26 LAW & SOC'Y REV. 513 (1992); Diamond et al., *supra* note 194, at 315-16; Schkade et al., *supra* note 191, at 1153, 1159. All of these researchers found that deliberation results in higher damage awards than the pre-deliberation awards of the jury members.

197. See generally LEE EPSTEIN & JACK KNIGHT, *THE CHOICES JUSTICES MAKE* (1998); C.K. ROWLAND & ROBERT A. CARP, *POLITICS AND JUDGMENT IN FEDERAL DISTRICT COURTS* (1996).

198. Robert J. MacCoun, *Epistemological Dilemmas in the Assessment of Legal Decision Making*, 23 LAW & HUM. BEHAV. 723, 727 (1999).

199. *Id.* at 726 (“A weakness of this logic is that the observed discrepancy tells us nothing about whether either decisionmaker (or group) is actually accurate; both could be wrong.”). Lempert argues that a problem with comparing judge and jury decisionmaking “is the tendency to treat the judge’s opinion as the ‘correct’ one, and to regard contrary jury

ferences could reflect “jury incompetence, jury lawlessness or nullification, unintended bias, or legally acceptable differences in values.”<sup>200</sup> The degree to which juries ought to make decisions that would mirror those of judges is ill-defined:

We are by no means clear on how much like the judge and how different from him we wish the jury to be. If it is too much like the judge, the jury may lose all claim to a distinctive function. If it is too little like him we are disturbed by how easily jury equity elides into jury anarchy.<sup>201</sup>

#### A. *Interpreting the Differences*

Each type of decisionmaker has different advantages that they bring to their role, and each has characteristics that give them a comparative advantage under some circumstances or for some purposes. Jurors, on the one hand, have the benefit of being able to decide in groups and are expected to bring to the process values that reflect the conscience of the community.<sup>202</sup> As noted above, some dis-

verdicts as deviations from a norm. Yet it is not obvious that the judge’s decision will be nearer the ideal than the jury’s.” Lempert, *supra* note 9, at 103 (footnote omitted). Similarly, Scheiner asserts that:

[J]ury bias can mean that the jury exercises sound discretion, but that jury determinations nevertheless systematically differ from those that judges would make because of moral or experiential differences between judges and juries. Because there is no unbiased baseline, any claim of juror bias of this . . . type can be restated as an equally true claim of judicial bias.

Alan Howard Scheiner, *Judicial Assessment of Punitive Damages, the Seventh Amendment, and the Politics of Jury Power*, 91 COLUM. L. REV. 142, 165 (1991).

200. MacCoun, *supra* note 198, at 727; see also Neal R. Feigenson, *Can Tort Juries Punish Competently?*, 78 CHI.-KENT L. REV. 239 (2003) (reviewing CASS R. SUNSTEIN ET AL., PUNITIVE DAMAGES: HOW JURIES DECIDE (2002)).

[B]ecause juries may legitimately bring to the justice system values that judges may not adequately express, and because those different values may sometimes lead juries to decide cases differently than judges would, it would be odd to rely too heavily on the judicial norm (or the legal norm as construed by judges) to evaluate juries’ decisions.

*Id.* at 244.

201. Harry Kalven, Jr., *The Jury, the Law, and the Personal Injury Damage Award*, 19 OHIO ST. L.J. 158, 178 (1958).

202. See Paula L. Hannaford et al., *How Judges View Civil Juries*, 48 DEPAUL L. REV. 247 (1998) (noting that decisions about damages are particularly for the jury). Kalven noted that:

The judge and jury are two remarkably different institutions for reaching the same objective—fair impersonal adjudication of controversies. The judge represents tradition, discipline, professional competence and repeated experience with the matter. . . . [The jury is] the layman amateur drawn from a wide public, disciplined only by the trial process and by the obligation to reach a group verdict . . . .

Kalven, *supra* note 201, at 178. In addition, MacCoun argues that judges and laypersons are “normatively advantaged in different ways. Judges are clearly more expert in the law . . . . But jurors are better suited for reflecting community standards in an ambiguous domain like tort litigation.” MacCoun, *supra* note 198, at 726-27 (citation omitted).

agreements between judges and juries are desirable to the extent that they “may reflect value judgments that are properly made by the jury in close cases.”<sup>203</sup> Differences that are the result of improved decisionmaking afforded by group processes or that reflect the legitimate value judgments of the jury should be applauded.

Other disagreements, however, may not result from the jury’s comparative advantages. Instead, some differences may be a result of factors for which juries are comparatively disadvantaged in relation to judges. In particular, judges might be expected to have a comparative advantage over jurors due to their training and expertise in the law and in the fact that, as repeat players in the system, they see a large number and a wide variety of cases. Judges are advantaged, it is typically argued, in that they have better memory for the trial evidence (often aided by the trial transcripts), have a better understanding of the law (due to their training and experience), and are better at identifying information presented at trial that is relevant to the legal decisions they are expected to make.<sup>204</sup>

Given the other advantages of juries, however, it may not be appropriate to eliminate the jury in these circumstances, but rather to focus on ways in which the task of the jury can be structured to minimize these disadvantages (that is, to make the jury more like a judge in these particular ways) without diminishing the jury’s unique contribution and role. For example, structurally, juries’ collective memory for trial evidence may exceed that of individual jurors.<sup>205</sup> Moreover, there is evidence that juries’ collective memory for trial evidence can be enhanced by allowing them access to the trial transcripts,<sup>206</sup> to take notes,<sup>207</sup> and to ask questions during trial.<sup>208</sup> There is evidence that juries’ understanding of the legal rules could

203. Lempert, *supra* note 9, at 104.

204. See, e.g., *id.* at 91-92 n.59.

205. See *supra* text accompanying notes 192-96.

206. See, e.g., Martin J. Bourgeois et al., *Effects of Technicality and Access to Trial Transcripts on Verdicts and Information Processing in a Civil Trial*, 19 PERSONALITY & SOC. PSYCHOL. BULL. 220 (1993); Lynne ForsterLee et al., *The Bottom Line: The Effect of Written Expert Witness Statements on Juror Verdicts and Information Processing*, 24 LAW & HUM. BEHAV. 259 (2000).

207. See, e.g., Lynne ForsterLee et al., *Effects of Notetaking on Verdicts and Evidence Processing in a Civil Trial*, 18 LAW & HUM. BEHAV. 567 (1994); Irwin A. Horowitz & Lynne ForsterLee, *The Effects of Note-Taking and Trial Transcript Access on Mock Jury Decisions in a Complex Civil Trial*, 25 LAW & HUM. BEHAV. 373 (2001); David L. Rosenhan et al., *Notetaking Can Aid Juror Recall*, 18 LAW & HUM. BEHAV. 53 (1994).

208. See, e.g., Larry Heuer & Steven Penrod, *Increasing Jurors’ Participation in Trials: A Field Experiment with Jury Notetaking and Question Asking*, 12 LAW & HUM. BEHAV. 231 (1988); Larry Heuer & Steven D. Penrod, *Instructing Jurors: A Field Experiment with Written and Preliminary Instructions*, 13 LAW & HUM. BEHAV. 409 (1989) [hereinafter Heuer & Penrod, *Instructing Jurors*]; Heuer & Penrod, *Trial Complexity*, *supra* note 35; Larry Heuer & Steven Penrod, *Juror Notetaking and Question Asking During Trials: A National Field Experiment*, 18 LAW & HUM. BEHAV. 121 (1994) [hereinafter Heuer & Penrod, *Juror Notetaking*].

be aided by more clearly and carefully written jury instructions,<sup>209</sup> by providing written instructions,<sup>210</sup> and by providing reasons for the legal rules they are asked to apply.<sup>211</sup> Moreover, there is evidence that juries' ability to identify and use relevant evidence could be enhanced by giving them instruction on the law before they hear evidence in the case so that they have a template, or schema, with which to organize the trial information.<sup>212</sup> Additional research is needed to explore how these structural changes might moderate any differences in the decisionmaking of juries as compared to judges.

It is also clear that there are certain types of errors that are made by both jurors and judges. Thus, even high rates of agreement are of some concern to the extent that such agreement is due to both groups making similar errors. For example, both judges and jurors have been shown to have difficulty assessing scientific and statistical evidence and in ignoring inadmissible evidence. It is possible to deal with these shared shortcomings in several different ways—either by devising ways to help both groups improve their decisionmaking or by structuring trials to minimize such problems.

First, at least for some difficulties in decisionmaking, both groups might be assisted in ways that can improve their decisionmaking. Take the example of the difficulties that both jurors and judges appear to have with scientific and statistical evidence. Both groups might be helped by instruction in these types of reasoning. There is evidence that relatively short, targeted training sessions can improve

---

209. There is some evidence that group deliberation improves juror understanding if instructions are rewritten according to psycholinguistic principles. See Joel D. Lieberman & Bruce D. Sales, *What Social Science Teaches Us About the Jury Instruction Process*, 3 PSYCHOL. PUB. POLY & L. 589, 634 (1997).

210. See, e.g., Heuer & Penrod, *Instructing Jurors*, *supra* note 208, at 419-21 (finding that while jurors who were provided written instructions found them helpful, no difference in understanding was detected); 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, 429-30 (1990) (finding better comprehension with some written instructions); Irene Glassman Prager et al., *Improving Juror Understanding for Intervening Causation Instructions*, 3 FORENSIC REP. 187, 191-92 (1989) (same).

211. See, e.g., Diamond & Casper, *supra* note 196, at 534; Duane T. Wegener et al., *Flexible Corrections of Juror Judgments: Implications for Jury Instructions*, 6 PSYCHOL. PUB. POLY & L. 629, 646 (2000); Roselle L. Wissler et al., *The Impact of Jury Instructions on the Fusion of Liability and Compensatory Damages*, 25 LAW & HUM. BEHAV. 125, 134-35 (2001).

212. See, e.g., Lynne ForsterLee & Irwin A. Horowitz, *Enhancing Juror Competence in a Complex Trial*, 11 APPLIED COGNITIVE PSYCHOL. 305, 306 (1997); Lynne ForsterLee et al., *Juror Competence in Civil Trials: Effects of Preinstruction and Evidence Technicality*, 78 J. APPLIED PSYCHOL. 14, 19 (1993); Heuer & Penrod, *Instructing Jurors*, *supra* note 208, at 425-26; 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, 1881 (1979); Vicki L. Smith, *The Feasibility and Utility of Pretrial Instruction in the Substantive Law: A Survey of Judges*, 14 LAW & HUM. BEHAV. 235, 244 (1990).

individuals' ability to engage in statistical and methodological reasoning.<sup>213</sup> Open empirical questions remain regarding whether and to what extent such training could be effectively incorporated in the trial process, the most effective way to structure it, and whether it would be equally effective for jurors and judges. Similarly, to address concerns about hindsight bias, Hastie and Viscusi suggest that "[i]t is even possible that the performance of judges can be improved by adding information from professional risk analysis experts at trial, perhaps in the form of expert testimony."<sup>214</sup> If such testimony proved to be helpful, it is likely that it would be helpful to juries as well.

Second, there are ways in which the division of labor within jury trials provides a structural mechanism that attempts to minimize certain shortcomings in decisionmaking. For example, as described above, there is evidence that both jurors and judges have difficulty ignoring inadmissible evidence. Jury trials are already structured so as to minimize this potential problem. First, there is some evidence that group deliberation can moderate the effect of inadmissible evidence on decisionmaking.<sup>215</sup> Second, because a judge will typically make a ruling in a pretrial hearing about whether particular evidence is admissible, the jury, in most cases, will be shielded from the inadmissible evidence. In a bench trial, in contrast, the judge as fact-finder cannot avoid being exposed to the evidence.

---

213. Geoffrey T. Fong et al., *The Effects of Statistical Training on Thinking About Everyday Problems*, 18 COGNITIVE PSYCHOL. 253 (1986); see also Richard P. Larrick et al., *Teaching the Use of Cost-Benefit Reasoning in Everyday Life*, 1 PSYCHOL. SCI. 362 (1990); Richard P. Larrick et al., *Who Uses the Cost-Benefit Rules of Choice? Implications for the Normative Status of Microeconomic Theory*, 56 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 331, 338-40 (1993); Lehman et al., *supra* note 103, at 440-41; Darrin R. Lehman & Richard E. Nisbett, *A Longitudinal Study of the Effects of Undergraduate Training on Reasoning*, 26 DEVELOPMENTAL PSYCHOL. 952, 959 (1990); Nisbett et al., *Teaching Reasoning*, *supra* note 101, at 627-28. It is reasonable to believe that juries may have a comparative advantage in this context:

Judges are lawyers; lawyers are people who, disproportionately more than most educated Americans, are uncomfortable with quantitative, scientific, and technological information; avoided it as students, and are incompetent with it as adults. By contrast, a well assembled jury containing a high school science teacher, an accountant, or an engineer, should have greater potential than the average judge to understand complex technical or quantitative evidence.

Saks, *supra* note 191, at 43.

214. Hastie & Viscusi, *supra* note 3, at 917.

215. See, e.g., Thomas R. Carretta & Richard L. Moreland, *The Direct and Indirect Effects of Inadmissible Evidence*, 13 J. APPLIED SOC. PSYCHOL. 291, 306-07 (1983); Kamala London & Narina Nunez, *The Effect of Jury Deliberations on Jurors' Propensity to Disregard Inadmissible Evidence*, 85 J. APPLIED PSYCHOL. 932, 937-38 (2000). But see William C. Thompson et al., *Inadmissible Evidence and Juror Verdicts*, 40 J. PERSONALITY & SOC. PSYCHOL. 453, 461 (1981) ("[T]here are strong indications that the biasing effects of inadmissible evidence persisted following deliberations.").

### B. Research Directions

In order to be able to devise mechanisms by which decisionmaking can be improved—either by aiding decisionmakers (judges or juries) or by allocating responsibility for decisionmaking—we must have a better understanding of judge and jury decisionmaking that goes beyond simple comparisons. Hastie and Rasinski have classified four different mechanisms, or “logics,” for measuring “accuracy” in judgments that may prove to be helpful in guiding future research in this area.<sup>216</sup> One of these approaches is to make a direct comparison between different decisionmakers.<sup>217</sup> The research reviewed here comparing the decisions of judges and juries exemplifies this approach. As Hastie and Rasinski explain, this method “involves comparing the judgment of one subject with the judgment of a second subject to determine whether the two judges agree. If the two judges disagree, then someone must be in error.”<sup>218</sup> Despite the problems with this approach, research making a direct comparison between judges and jurors or juries has been useful in demonstrating that there are fewer differences in the decisions made by judges and juries than might have been imagined. Moreover, these comparisons are useful in identifying areas of agreement and disagreement which can then be further studied in order to more carefully determine the nature of these agreements or disagreements and the processes by which they occur.

However, Hastie and Rasinski also distinguish three other logics. First, they argue that the “most basic” comparison is one between a judgment and a demonstrably correct answer.<sup>219</sup> Ideally, one could compare the decisions of judges and juries with the correct outcome to determine which decisionmaker comes closest and under what circumstances. Feigenson has distinguished a set of criteria which could be used to define the correct answers necessitated by this logic: a “legal norm” by which the correct answer is determined by “what the relevant legal rules appear to require”; an “economic norm” by which the correct answer is determined by “the outcome dictated by the consequentialist theory of optimal deterrence”; and a “rational judgment norm” by which the correct answer is determined by “the standards of rational decision-making commonly employed in cogni-

---

216. Reid Hastie & Kenneth A. Rasinski, *The Concept of Accuracy in Social Judgment*, in *THE SOCIAL PSYCHOLOGY OF KNOWLEDGE* 193 (Daniel Bar-Tal & Arie W. Kruglanski eds., 1988). Robert MacCoun has previously identified the usefulness of this taxonomy for thinking about these comparisons. MacCoun, *supra* note 198, at 725-26.

217. Hastie & Rasinski, *supra* note 216, at 197.

218. *Id.*

219. *Id.* at 196.

tive and social psychological research.”<sup>220</sup> However, as MacCoun notes, cases that come to trial are frequently those cases that lack such clear answers.<sup>221</sup> To limit the judge-jury comparison to cases in which the answer is clear misses far too much.

Two other logics identified by Hastie and Rasinski involve examining the information that decisionmakers use in reaching a decision. Decisionmakers may err because they “use a bad cue” such that their decision was influenced by a factor that should not have influenced the decision<sup>222</sup> or they may “miss a good cue” such that their decision was not influenced by a factor which should have provided an influence.<sup>223</sup> For example, decisionmakers use good cues when they base their punitive damage awards on the nature of the harm and the reprehensibility of the defendant’s conduct. In contrast, decisionmakers miss a good cue when they do not appropriately use base-rate information, and they use bad cues when they base decisions on flawed empirical research or inadmissible evidence. Following these logics, one could compare judges and juries in their abilities to reach decisions using appropriate cues.

Research that combines a direct comparison of decisionmakers (that is, judges and juries) with another of these logics or norms (for example, using bad cues or missing good cues) would prove useful in determining not just whether the decisions of judges and jurors differ, but how the two groups compare against some other benchmark. For example, the Landsman and Rakos study of inadmissible evidence is a useful model of this approach. It directly compared judges and laypersons in their use of inadmissible evidence—a bad cue—as defined by the legal rules of evidence—a legal norm.<sup>224</sup> Similarly, Robbennolt directly compared judges and laypersons in their use of both good and bad cues for determining compensatory and punitive damages, as defined by case law—legal norms.<sup>225</sup>

In addition, future research might also carefully explore and compare the norms, attitudes, and understandings that jurors and judges bring with them into the courtroom. Such factors can influence how judges and jurors decide cases, particularly in close

220. Feigenson, *supra* note 200, at 242. Feigenson also described a “judicial norm” (that is, “what a judge would have decided”) which corresponds to the “logic” of direct comparison between juries and judges. *Id.*

221. See MacCoun, *supra* note 198, at 726 (“This logic has been quite fruitful in psychophysics, but perhaps less so in social psychology, where we often lack objective measures of the ‘true’ state of the sociopolitical environment.”).

222. Hastie & Rasinski, *supra* note 216, at 198; see also Kerr et al., *supra* note 192, at 689 (describing “sins of commission”).

223. Hastie & Rasinski, *supra* note 216, at 198-99; see also Kerr et al., *supra* note 192, at 689 (describing “sins of omission”).

224. Landsman & Rakos, *supra* note 130, at 120-22.

225. Robbennolt, *supra* note 2, at 146-58.

cases,<sup>226</sup> and may play an important role in those cases in which juries and judges would reach different results. Existing research, for example, has explored juror attitudes toward tort plaintiffs, plaintiffs' lawyers, corporate defendants, and physicians.<sup>227</sup> In addition, psychologists have explored the notions that laypersons hold about justice and fairness.<sup>228</sup> Political science research has explored the influences of a variety of individual difference factors, including political affiliation, on judicial decisionmaking.<sup>229</sup> Very little research has explored the similarities and differences of jurors and judges on these types of normative, attitudinal, and conceptual factors. Whether these predispositions are "good cues," "bad cues," or simply legitimate value differences, a more nuanced understanding of these notions would help to inform this debate.

## VII. CONCLUSION

Research has begun to explore the similarities and differences in the decisionmaking of juries and judges—although direct comparisons are still relatively uncommon. Research that has been done to date demonstrates some differences, but shows striking similarities in the decisionmaking of judges and jurors. While there appear to be some differences in outcomes under some circumstances, judges and jurors generally appear to be influenced by similar factors and suffer from many of the same difficulties in making their decisions. More research is needed to define the contours of these differences and similarities and to inform mechanisms by which the decisionmaking process can be improved. In particular, future research ought to combine a focus on making direct comparisons between judges and juries with evaluation against external benchmarks.

---

226. Consistent with the findings of jury research, *see* sources cited *supra* note 99, studies of district court decisionmaking find that legal factors have the greatest impact. Extralegal factors, however, are found to play a greater role in close cases. *See* Tracey E. George, *Judicial Independence and the Ambiguity of Article III Protections*, 64 OHIO ST. L.J. 221, 242-43 (2003) (citing empirical studies).

227. *See, e.g.*, VALERIE P. HANS, *BUSINESS ON TRIAL: THE CIVIL JURY AND CORPORATE RESPONSIBILITY* (2000); NEIL VIDMAR, *MEDICAL MALPRACTICE AND THE AMERICAN JURY* (1995).

228. *See, e.g.*, NORMAN J. FINKEL, *COMMONSENSE JUSTICE: JURORS' NOTIONS OF THE LAW* (1995); NORMAN J. FINKEL, *NOT FAIR! THE TYPOLOGY OF COMMONSENSE UNFAIRNESS* (2001).

229. For a summary of the research exploring the effects of individual differences on judicial decisionmaking, *see* Tracey E. George, *Court Fixing*, 43 ARIZ. L. REV. 9 (2001); and George, *supra* note 226, at 242-44 (discussing empirical studies showing a relationship between the appointing President and decisionmaking by district court judges).

