JURORS OF THE ABSURD? THE ROLE OF CONSEQUENTIALITY IN JURY SIMULATION RESEARCH

BRIAN H. BORNSTEIN* & SEAN G. MCCABE**

There are a number of situations in which it is impractical, expensive, unethical, or methodologically deficient to study behavior in its natural context. Researchers are generally not allowed in the cockpit of an F-14 to study pilots' decisionmaking, nor are they allowed to traumatize some participants (but not others) to determine whether highly stressful events might be repressed. The legal domain is no exception, as researchers are generally unable to assign decisionmakers (for example, judges and juries) to experimental conditions or observe their decisions as they naturally occur. Archival studies overcome some of these limitations, but they present their own set of problems, such as selective sampling and irremediably confounded variables.[1] The major recourse, as in other arenas of behavioral science, is to conduct simulations.[2] Analyses of the field of jury research

* Brian H. Bornstein is Professor of Psychology and Courtesy Professor of Law at the University of Nebraska-Lincoln. He received his Ph.D. in 1991 from the University of Pennsylvania and Master of Legal Studies in 2001 from the University of Nebraska. His main research interests are jury decisionmaking and eyewitness testimony

** Sean G. McCabe is currently in the J.D.-Ph.D. program at the University of Nebraska-Lincoln. He earned his B.Sc. in Neuroscience and Psychology from Dalhousie University in Halifax, Nova Scotia. His research interests include the social and cognitive factors that affect jury performance and damage awards.

Portions of this Article were presented at the International Interdisciplinary Conference on Psychology & Law, Edinburgh, Scotland (July 2003), and at the Symposium on the Behavioral Analysis of Legal Institutions: Possibilities, Limitations, and New Directions, Tallahassee, Florida (March 2004). We are grateful for the comments of our fellow panelists at both conferences, and especially to Steve Penrod and Greg Mitchell for their organizational efforts.

1. *See, e.g.*, Neil Vidmar, *Making Inferences About Jury Behavior from Jury Verdict Statistics: Cautions About the Lorelei's Lied*, 18 LAW & HUM. BEHAV. 599 (1994). Vidmar also highlights the "apples and oranges" problem involved in comparing case types that differ on multiple dimensions as well as other issues that arise in relying on jury verdict reporters as the source of archival data. *Id.* at 605-06.

2. The pros and cons of simulation methodology generally have been treated else-

suggest that experimental simulations have been increasing in recent years.[3]

Simulated trials differ from real trials in a number of respects, such as the research setting and sample, as well as procedural and substantive verisimilitude. These limitations have been discussed extensively elsewhere.[4] They can be overcome, to some extent, by using a relatively diverse sample instead of only college undergraduates and by using relatively realistic materials and judgment tasks.[5] However, one limitation seems insurmountable, as it is the sine qua non of a simulation; namely, no matter how realistic a simulation is, it is still just a simulation. Simulations can be quite realistic in capturing the verisimilitude of their real-world analog in terms of the participant sample, the procedure, and other methodological charac-

---

where. *See, e.g.*, Elliot Aronson et al., *Experimentation in Social Psychology*, *in* 1 THE HANDBOOK OF SOCIAL PSYCHOLOGY 99 (Daniel T. Gilbert et al. eds., 4th ed. 1998); Ralph Hertwig & Andreas Ortmann, *Experimental Practices in Economics: A Methodological Challenge for Psychologists?*, 24 BEHAV. & BRAIN SCI. 383 (2001); David O. Sears, *College Sophomores in the Laboratory: Influences of a Narrow Data Base on Social Psychology's View of Human Nature*, 51 J. PERSONALITY & SOC. PSYCHOL. 515 (1986).

3. *See* Jeremy A. Blumenthal, *Law and Social Science in the Twenty-First Century*, 12 S. CAL. INTERDISC. L.J. 1 (2002); Brian H. Bornstein, *The Ecological Validity of Jury Simulations: Is the Jury Still Out?*, 23 LAW & HUM. BEHAV. 75, 86-87 (1999) [hereinafter Bornstein, *Ecological Validity*]; Michael T. Nietzel et al., *Juries: The Current State of the Empirical Literature*, *in* PSYCHOLOGY AND LAW: THE STATE OF THE DISCIPLINE 23, 25-26 (Ronald Roesch et al. eds., 1999).

4. *See generally* Bornstein, *Ecological Validity*, *supra* note 3; Robert M. Bray & Norbert L. Kerr, *Methodological Considerations in the Study of the Psychology of the Courtroom*, *in* THE PSYCHOLOGY OF THE COURTROOM 287 (Norbert L. Kerr & Robert M. Bray eds., 1982) [hereinafter Bray & Kerr, *Methodological Considerations*]; 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 (1979) [hereinafter Bray & Kerr, *Use of Simulation*]; James H. Davis et al., *The Empirical Study of Decision Processes in Juries: A Critical Review*, *in* LAW, JUSTICE, AND THE INDIVIDUAL IN SOCIETY: PSYCHOLOGICAL AND LEGAL ISSUES 326 (June Louin Tapp & Felice J. Levine eds., 1977); Shari Seidman Diamond, *Illuminations and Shadows from Jury Simulations*, 21 LAW & HUM. BEHAV. 561 (1997); Ronald C. Dillehay & Michael T. Nietzel, *Constructing a Science of Jury Behavior*, *in* 1 REVIEW OF PERSONALITY AND SOCIAL PSYCHOLOGY 246 (Ladd Wheeler ed., 1980); Kathleen Carrese Gerbasi et al., *Justice Needs a New Blindfold: A Review of Mock Jury Research*, 84 PSYCHOL. BULL. 323 (1977); Nancy J. King, *Postconviction Review of Jury Discrimination: Measuring the Effects of Juror Race on Jury Decisions*, 92 MICH. L. REV. 63 (1993); Vladimir J. Konečni & Ebbe B. Ebbesen, *External Validity of Research in Legal Psychology*, 3 LAW & HUM. BEHAV. 39 (1979); Vladimir J. Konečni & Ebbe B. Ebbesen, *Methodological Issues in Research on Legal Decision-Making, with Special Reference to Experimental Simulations*, *in* PSYCHOLOGY AND LAW: INTERNATIONAL PERSPECTIVES 413 (Friedrich Lösel et al. eds., 1992) [hereinafter Konečni & Ebbesen, *Methodological Issues*]; Robert J. MacCoun, *Experimental Research on Jury Decision-Making*, 244 SCIENCE 1046 (1989); Michael J. Saks, *What Do Jury Experiments Tell Us About How Juries (Should) Make Decisions?*, 6 S. CAL. INTERDISC. L.J. 1 (1997) [hereinafter Saks, *Jury Experiments*]; Wayne Weiten & Shari Seidman Diamond, *A Critical Review of the Jury Simulation Paradigm: The Case of Defendant Characteristics*, 3 LAW & HUM. BEHAV. 71 (1979).

5. It is worth noting, however, that these particular variables—the composition of the mock juror sample and realism of the mock trial—have little effect on the results of jury simulation research. *See* Bornstein, *Ecological Validity*, *supra* note 3, at 76-84.

teristics. Nonetheless, the essence of a simulation remains, which is that participants' decisions lack real consequences.

Given that these consequences are absent, who would be affected by them if they were present? Most obviously, the litigants would be affected. Mock jurors reach a verdict concerning a paper defendant, who exists solely for purposes of the experiment; but real jurors make decisions concerning a flesh-and-blood defendant, which could entail a prison sentence or hefty damage award.[6] Somewhat less obviously, there are consequences for the jurors themselves. Mock jurors volunteer to participate in a relatively short study, seldom lasting longer than a couple of hours, which they fit into their daily routine, and for which they are paid or receive course credit. Real jurors are summoned to appear at a designated time and place, without regard to the needs and demands of their daily lives, to participate in a trial which might last for weeks, and for which they receive minimal compensation.

The purpose of this Article is to consider the ramifications of this distinction, which we refer to as the "consequentiality issue." Part I elaborates on the principal questions raised by the issue. Part II presents the courts' response to jury simulation research. Part III presents researchers' response to the concerns about consequentiality and describes research findings on the issue, as well as the limitations of that research. Part IV considers relevant research on related topics. The comment concludes in Part V with a discussion of possible alternatives to jury simulation research.

I. QUESTIONS ABOUT SIMULATIONS: IS THE ENTERPRISE ABSURD?

Despite the prevalence of simulation methodology in behavioral science research more broadly, not just regarding juries, its use has not gone unquestioned. As Kühberger et al. observed:

> It is a remarkable fact about decision research that the use of imagined situations is accepted as a legitimate means of studying real decision behavior. In other areas of psychology, such methods would be considered extremely questionable if not absurd. . . . Would any psychophysicist be taken seriously who investigated perceived heaviness not by giving participants actual weights to lift but by asking participants to imagine lifting a two pound weight?[7]

---

6. This is not at all to suggest that mock jurors are cavalier in their approach. Like most jury researchers, we have continually been impressed by how seriously most mock jurors take their task. Deliberations involving hypothetical parties can become quite heated. The question, though, is whether knowing in the back of one's mind that there are ultimately no real consequences for the parties involved can affect either the process or the outcome of mock jurors' reasoning.

7. Anton Kühberger et al., *Framing Decisions: Hypothetical and Real*, 89

Kühberger et al. point out that despite this apparent absurdity, all decisionmaking—even in real situations—is actually hypothetical: it involves weighing evidence to test hypotheses, considering events that may or may not obtain, and anticipating feelings we do not yet have.[8] Even real jurors reason in this fashion; for example, if I accept as true what this witness says, how much does it increase or decrease the probability that the defendant is guilty?[9] How would I feel if I vote to convict the defendant but he really is innocent? So ultimately, in a sense, the disjunction is between thinking hypothetically about real cases and thinking hypothetically about hypothetical cases. The crucial distinction is whether the hypothesis-testing process ends in consequences for the parties involved.

Should psycholegal researchers be concerned? The answer, not surprisingly, is that it depends. In particular, it depends primarily, perhaps exclusively, on whether real and simulated decisionmaking differ in either process or outcome. In terms of process, mock jurors, knowing their decisions lack real consequences, might take their task less seriously and therefore put forth less effort than real jurors. This could lead them to process the evidence less systematically or apply the law inappropriately, thereby denying defendants due process.[10] An abundance of research shows that when decisionmakers are "accountable" for their decisions—that is, when they have to justify or defend them to others—their decisionmaking process is more rational and complex.[11] Real jurors—who face the litigants in open court, answer to the judge, and can be polled afterwards by the attorneys—surely feel accountable; but it is unclear how accountable mock jurors, who are usually nondeliberating students earning extra course credit for completing a questionnaire describing hypothetical parties, would feel.

ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 1162, 1162 (2002).

8. *Id.* at 1163.

9. A number of theorists have explicitly framed jurors' decisionmaking in such hypothesis-testing terms, often invoking Bayes' Theorem as a normative model for doing so. *See, e.g.*, Brian H. Bornstein, *David, Goliath, and Reverend Bayes: Prior Beliefs About Defendants' Status in Personal Injury Cases*, 8 APPLIED COGNITIVE PSYCHOL. 233 (1994); Jonathan J. Koehler & Daniel N. Shaviro, *Veridical Verdicts: Increasing Verdict Accuracy Through the Use of Overtly Probabilistic Evidence and Methods*, 75 CORNELL L. REV. 247 (1990).

10. In support of this possibility, dual-process models of decisionmaking distinguish between "heuristic" and "systematic" processing. The former focuses more on superficial characteristics of the information and less on content, whereas the latter involves more cognitive effort and is more content-driven. *See generally* Shelly Chaiken, *Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion*, 39 J. PERSONALITY & SOC. PSYCHOL. 752 (1980). For a legal application, see Bradley D. McAuliff et al., *Juror Decision-Making in the Twenty-First Century: Confronting Science and Technology in Court*, *in* HANDBOOK OF PSYCHOLOGY IN LEGAL CONTEXTS 303 (David Carson & Ray Bull eds., 2d ed. 2003).

11. *See generally* Jennifer S. Lerner & Philip E. Tetlock, *Accounting for the Effects of Accountability*, 125 PSYCHOL. BULL. 255 (1999).

Regarding outcomes, they could differ in terms of a statistical main effect—for example, if there were more convictions in simulated than in real murder trials—or in terms of an interaction, as would be the case if some substantive or procedural variable exerted a greater effect in one situation than the other. For instance, pretrial publicity or expert testimony on the reliability of eyewitness memory might exert a larger effect in simulated than real trials, or vice versa.[12] As several commentators have noted, methodological variables that are involved in interactions are more problematic, though often more interesting theoretically, than variables that exert main effects.[13]

The potential for differences between real and simulated jury decisionmaking has both theoretical/psychological and legal/policy implications. From a psychological-theory perspective, a failure to obtain comparable findings in the two settings would compromise the simulation research's external validity, which "refers to the extent to which a particular causal relationship is robust across populations or settings."[14] The ultimate goal of most psychological research is to discover general principles of behavior; therefore, the discovery of principles that hold true only in limited domains—and artificially contrived ones, at that—has limited utility.[15]

Although it is true that most *psychological* research is guided by a quest for understanding general principles, a great deal of *psycholegal* research is conducted with the additional goal of using the findings to improve functioning of the legal system.[16] Hence the quality of the research has important legal implications as well, in the sense that courts and other policymaking bodies, such as legislatures, can rely on research findings in setting policies regarding jury issues. As discussed below, the courts are not always receptive to experimental research, due largely to these same concerns about external validity—that is, the extent to which the research findings have anything

---

12. We have chosen these particular examples for purely illustrative purposes, though some suggestive evidence on the pretrial publicity question is provided by Geoffrey P. Kramer & Norbert L. Kerr, *Laboratory Simulation and Bias in the Study of Juror Behavior: A Methodological Note*, 13 LAW & HUM. BEHAV. 89 (1989). Although they did not compare jury decisions with and without consequences, they did find that pretrial publicity effects did not appear to differ as a function of the length and complexity of simulated trials. *Id.* at 96-98.

13. *See, e.g.*, Bornstein, *Ecological Validity*, *supra* note 3, at 78; Bray & Kerr, *Methodological Considerations*, *supra* note 4, at 309-13; Saks, *Jury Experiments*, *supra* note 4, at 8.

14. Aronson et al., *supra* note 2, at 130.

15. *Id. See generally* Diamond, *supra* note 4 (discussing examples of limited usefulness of certain simulations).

16. Professor Michael Saks has clearly articulated the relationship between empirical research and legal policy decisions. 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 (1992); Saks, *Jury Experiments*, *supra* note 4; Michael J. Saks, *Legal Policy Analysis and Evaluation*, 44 AM. PSYCHOLOGIST 1110 (1989).

substantial to say about actual legal contexts. It is, therefore, to the field's credit that the experimental methodologies used by researchers are becoming increasingly sophisticated and legally realistic.[17] Nonetheless, an experiment is ultimately still an experiment, raising the issue of whether any simulation can meaningfully speak to real-world legal questions.

II. THE COURTS' RESPONSE TO SIMULATION

Trial and appellate courts have been presented with social scientific research at least as far back as the famous "Brandeis brief" in *Muller v. Oregon*,[18] and considerable evidence suggests that the trend is increasing.[19] Such research has been presented in a number of contexts, ranging from desegregation and punitive damages to affirmative action, eyewitness identification, capital punishment, and countless others.[20] On the whole, courts have been reluctant to base their decisions on social scientific data,[21] though there are, of course, exceptions.[22] The courts' response to psychological research pertaining to capital cases (for example, death qualification and instruction comprehension) is perhaps the most prominent example of judges' reluctance to rely on experimental simulations.[23] In many cases, courts

---

17. *See* Diamond, *supra* note 4; Nietzel et al., *supra* note 3, at 23-24; William C. Thompson, *Research on Jury Decision Making: The State of the Science*, *in* INDIVIDUAL AND GROUP DECISION MAKING: CURRENT ISSUES 203 (N. John Castellan, Jr. ed., 1993).

18. 208 U.S. 412 (1908).

19. For a brief history of social science in the law, see generally MARK COSTANZO, PSYCHOLOGY APPLIED TO LAW 1-30 (2004); JOHN MONAHAN & LAURENS WALKER, SOCIAL SCIENCE IN LAW: CASES AND MATERIALS 1-29 (4th ed. 1998); Blumenthal, *supra* note 3; Phoebe C. Ellsworth & Julius G. Getman, *Social Science in Legal Decision-Making*, *in* LAW AND THE SOCIAL SCIENCES 581 (Leon Lipson & Stanton Wheeler eds., 1986); and Phoebe C. Ellsworth & Robert Mauro, *Psychology and Law*, *in* 2 THE HANDBOOK OF SOCIAL PSYCHOLOGY, *supra* note 2, at 684.

20. A comprehensive listing of cases involving social scientific data is beyond the scope of the present Article. For representative cases, see the following: Gratz v. Bollinger, 539 U.S. 244, 298-300 (2003) (Ginsburg, J., dissenting) (regarding affirmative action); Grutter v. Bollinger, 539 U.S. 306, 330-31 (2003) (regarding benefits of affirmative action); State Farm Mut. Auto. Ins. Co. v. Campbell, 538 U.S. 408 (2003) (regarding punitive damages; although the Court's decision did not cite social scientific research, jury researchers submitted amici curiae on behalf of both plaintiff and defendant); Lockhart v. McCree, 476 U.S. 162, 168-71 (1986) (regarding capital punishment); Brown v. Bd. of Educ., 347 U.S. 483, 494 n.11 (1954) (regarding desegregation); United States v. Hudson, 884 F.2d 1016, 1023-24 (7th Cir. 1989) (regarding eyewitness identification). For general reviews, see Blumenthal, *supra* note 3; MONAHAN & WALKER, *supra* note 19; and Mark I. Satin, *Law and Psychology: A Movement Whose Time Has Come*, 1994 ANN. SURV. AM. L. 581, 600-02.

21. *See, e.g.*, MONAHAN & WALKER, *supra* note 19; J. Alexander Tanford, *The Limits of a Scientific Jurisprudence: The Supreme Court and Psychology*, 66 IND. L.J. 137, 138 (1990) (indicating that the Supreme Court has not welcomed research on jury behavior).

22. *See* Diamond, *supra* note 4, at 569; Tanford, *supra* note 21, at 138.

23. *See, e.g.*, James R. Acker, *A Different Agenda: The Supreme Court, Empirical Research Evidence, and Capital Punishment Decisions, 1986-1989*, 27 LAW & SOC'Y REV. 65 (1993); Diamond, *supra* note 4, at 567-69; Phoebe C. Ellsworth, *Unpleasant Facts: The Supreme Court's Response to Empirical Research on Capital Punishment*, *in*

reviewing death sentences have discounted experimental social scientific findings due to various perceived methodological shortcomings of the research;[24] yet in a number of cases, one of the supposed shortcomings was the simulated nature of the research itself.

For example, in *Free v. Peters* the Seventh Circuit considered the results of a jury simulation that assessed mock jurors' comprehension of death penalty instructions.[25] Among other shortcomings of the study, the court opined that "[t]he first [fatal flaw] is lack of comparability between the test setting and the setting of the sentencing hearing."[26] In other words, participants in the test setting were deficient because they did not set real sentences. The Missouri Supreme Court reached a similar conclusion in another capital case, *State v. Deck*.[27] The court rejected experimental data showing that penalty phase instructions were poorly understood, concluding:

> [The] study, however, must be discounted because the people interviewed for the study did not act as jurors. They were given hypothetical facts that were different than the facts in this case, and they did not hear the testimony of witnesses, observe physical evidence or deliberate with eleven other jurors.[28]

The capital case that has received the most commentary in this respect is *Lockhart v. McCree*, in which Justice Rehnquist, writing for a majority of the U.S. Supreme Court, harshly criticized several jury simulations that purported to show that death-qualified juries were conviction-prone.[29] Among other failings (for example, lack of deliberation), the Court identified the lack of consequences flowing from the decision as a major factor undermining the studies' implica-

---

CHALLENGING CAPITAL PUNISHMENT: LEGAL AND SOCIAL SCIENCE APPROACHES (Kenneth C. Haas & James A. Inciardi eds., 1988); Tanford, *supra* note 21, at 144-48; *cf.* Konečni & Ebbesen, *Methodological Issues*, *supra* note 4, at 416-18 (criticizing the California Supreme Court's openness toward death-qualification voir dire simulations in true capital-case context and approving of the U.S. Supreme Court's conclusion that the research in that area should not be relied upon).

24. *See generally* sources cited *supra* note 23. This reluctance to adjudicate based on experimental evidence applies not merely to social scientific research, but to experimental data more broadly. Such evidence is often criticized as being too abstract and impersonal to bear on particular case facts, despite its relevance in a technical sense. *E.g.*, *In re* "Agent Orange" Product Liability Litigation, 611 F. Supp. 1223, 1241 (E.D.N.Y. 1985). More generally, it reflects a preference for case-specific information over aggregate data. *See, e.g.*, Brian H. Bornstein, *The Impact of Different Types of Expert Scientific Testimony on Mock Jurors' Liability Verdicts*, 10 PSYCHOL. CRIME & L. 429 (2004). A consideration of other types of experimental data is beyond the scope of the present Article; the focus here is on social/cognitive psychological experimentation.

25. 12 F.3d 700, 705 (7th Cir. 1993).

26. *Id.*

27. 994 S.W.2d 527 (Mo. 1999).

28. *Id.* at 542.

29. 476 U.S. 162, 168-73 (1986). There is sizable literature critiquing the Court's reasoning in *Lockhart*. *See* sources cited *supra* note 20.

tions: "[I]ndividuals . . . were not *actual* jurors sworn under oath to apply the law to the facts of an *actual* case involving the fate of an *actual* capital defendant."[30]

These selected cases should not be taken to imply that the courts are invariably hostile to simulation research. Some courts, in a variety of cases, have been quite receptive to simulation research.[31] Nonetheless, the cases discussed in the preceding paragraphs are a fair representation of appellate courts' general reluctance to base decisions on simulation research, particularly on grounds of external validity.

III. RESEARCHERS' RESPONSE TO THE CONSEQUENTIALITY QUESTION

For the most part, researchers have turned a blind eye to the consequentiality issue, within both psychology in general and psycholegal research in particular.[32] It has been alluded to in critiques of jury research, but overlooked in favor of more researchable problems, such as who the mock jurors are or characteristics of the mock trial.[33] These studies—which tend to find few differences as a function of simulations' methodological characteristics—have focused on the quality of simulations, rather than simulations qua simulations.[34] A large part of the reason for this avoidance of the consequentiality issue is, no doubt, the fact that it is a notoriously intractable research question. A variety of research approaches exist, each with significant limitations.

A. Possible Research Approaches

There are essentially three possible research approaches, each fraught with logistical or ethical complications. First, one could lead mock jurors to believe their decisions have consequences in situations where they really have no consequences. Although such a procedure raises the ethical issue of deception, it does not exceed the degree of deception that is commonplace, and generally ethi-

---

30. *Lockhart*, 476 U.S. at 171 (emphasis added).

31. As Diamond points out, "[I]n a post-*Daubert* world, [courts] are paying more attention to scientific evidence of all kinds." Diamond, *supra* note 4, at 569 (citation omitted).

32. In making this assertion, the authors do not exempt themselves and have even defended simulation methodology elsewhere. *See, e.g.*, Bornstein, *Ecological Validity*, *supra* note 3. There are also several notable exceptions. *See generally* Aronson et al., *supra* note 2; Sears, *supra* note 2. Among jury researchers, Professor Shari Diamond has perhaps paid the greatest attention to the consequentiality issue. *See* Diamond, *supra* note 4; *see also* Weiten & Diamond, *supra* note 4.

33. *See, e.g.*, Bornstein, *Ecological Validity*, *supra* note 3; Nietzel et al., *supra* note 3.

34. *See* sources cited *supra* note 33.

cally acceptable, in social psychological research.[35] However, as a practical matter, it is difficult to convince research participants—who are often fairly savvy psychology undergraduates—that the research means what the experimenter says it means. It would be especially difficult to persuade participants in a psychology research laboratory that they are making decisions about a defendant who is merely shown on videotape or described in a written summary, which are the most commonly used simulation media.[36] For this sort of dissimulation to have even a chance of success, the trial would need to be conducted live in an actual courthouse, or it would need to involve a dispute resolution context where it was plausible that undergraduates in a laboratory setting would be making real decisions. The former approach is difficult, time-consuming, and expensive; the latter approach, though attempted with some success in the studies described below, is also quite complicated and strains participants' credulity.

Second, one could lead real jurors to believe their decisions do not have consequences in situations where they really do. It is hard to imagine a judge who would condone leading some jurors to believe they were making mere hypothetical decisions, for the sake of comparison to other jurors who believed they were making consequential decisions, only to implement those decisions after the fact in both conditions. Such a procedure would adequately address the research question, and it would be fairly easy to implement experimentally; but it is so blatantly unethical and unjust that it hardly warrants serious consideration.[37]

Third, one could randomly, or quasi-randomly, assign participants to conditions that differ solely in their consequences: really real versus really hypothetical. For example, one group of participants would review a set of case facts under simulation assumptions, while another group would review the same case facts but make a consequential decision.[38] This approach is expensive

---

35. *See* Aronson et al., *supra* note 2, at 135-37. For the American Psychological Association's guidelines regarding deception, see AM. PSYCHOLOGICAL ASS'N, ETHICAL PRINCIPLES OF PSYCHOLOGISTS AND CODE OF CONDUCT, Ethical Standard 8.07, *reprinted in* 57 AM. PSYCHOLOGIST 1060, 1070 (2002), *available at* http://www.psycinfo.com/psycarticles/2002-11464-006.htm (last visited Oct. 1, 2004).

36. Bornstein, *Ecological Validity*, *supra* note 3, at 86.

37. It is also quite possibly illegal. With very few exceptions (for example, the occasional investigative TV journalism report), jurors cannot be observed while performing their task. Because of fair-trial concerns, experimental manipulations involving juries are quite rare, and those that have been tried involve much less controversial procedures. *See infra* notes 144-46 and accompanying text. A manipulation that varies the whole essence of the jury's task, though intriguing from a scientific perspective, could never be implemented.

38. A study by Shari Diamond and Hans Zeisel that compared the decisions of actual and shadow juries comes closest to this sort of comparison. *See* Shari Seidman

and time-consuming, as it involves having the hypothetical decisionmakers sit through an entire trial, so that their experience is the same as the real decisionmakers except for the consequences of their decision. And even so, it is difficult to match their experiences completely: real jurors interact with each other, and with courtroom personnel (lawyers and the judge), in ways that shadow jurors would not. Finally, because the deliberations of real jurors are ordinarily off-limits to researchers, as noted above, one could compare the two groups on outcome but not on process.

B. Research Findings on the Consequentiality Question

Perhaps because of these myriad difficulties in conducting the research, we could find only five jury simulation studies that conducted a direct test of the consequentiality question (see Table 1, *infra*).[39] Most of these studies were conducted during the mid to late 1970s, which qualifies as the "heyday" of consequentiality research. It is indeed odd that more recent research has not been conducted, as the question was hardly resolved by this small flurry of activity twenty to thirty years ago. Unfortunately, these studies provide no consensus about the effects of consequentiality on jurors' decisions. One study found that convictions were *less* likely when the decision had real consequences,[40] one study found that convictions were *more* likely when the decision had real consequences,[41] one study found *no difference* between real and hypothetical decisions,[42] and the two remaining studies found no main effect of consequentiality but obtained interacting effects with other variables.[43]

Diamond & Hans Zeisel, *A Courtroom Experiment on Juror Selection and Decision-Making*, 1 PERSONALITY & SOC. PSYCHOL. BULL. 276 (1974). This study is described in more detail in Part III.B.

39. In chronological order, the studies were conducted by Diamond & Zeisel, *supra* note 38; David W. Wilson & Edward Donnerstein, *Guilty or Not Guilty? A Look at the "Simulated" Jury Paradigm*, 7 J. APPLIED SOC. PSYCHOL. 175 (1977); Norbert L. Kerr et al., *Role Playing and the Study of Jury Behavior*, 7 SOC. METHODS & RES. 337 (1979); David Suggs & John J. Berman, *Factors Affecting Testimony About Mitigating Circumstances and the Fixing of Punishment*, 3 LAW & HUM. BEHAV. 251 (1979); and Martin F. Kaplan & Sharon Krupa, *Severe Penalties Under the Control of Others Can Reduce Guilt Verdicts*, 10 LAW & PSYCHOL. REV. 1 (1986).

40. Diamond & Zeisel, *supra* note 38, at 276-77.

41. Wilson & Donnerstein, *supra* note 39, at 185.

42. Kerr et al., *supra* note 39, at 348.

43. Kaplan & Krupa, *supra* note 39, at 8-13; Suggs & Berman, *supra* note 39, at 256.

TABLE 1

STUDIES MEASURING THE EFFECT OF CONSEQUENTIALITY ON JUROR/JURY DECISIONS

STUDY	SETTING	MAIN EFFECT OF CONSEQUENTIALITY	INTERACTING EFFECTS	DELIBERATION
Diamond & Zeisel (1974)	Field	More guilty verdicts for mock juries than real juries	No	Yes
Wilson & Donnerstein (1977)	Laboratory	More guilty verdicts for "real" jurors than hypothetical jurors	No	No
Kerr et al. (1979)	Laboratory	No main effect of consequentiality	No	Yes
Suggs & Berman (1979)	Laboratory	No main effect of consequentiality	Yes	No
Kaplan & Krupa (1986)	Laboratory	No main effect of consequentiality	Yes	No

An early field study by Shari Diamond and Hans Zeisel looked specifically at differences between the verdicts of real and experimental juries.[44] In an experiment designed to examine the effect of peremptory challenges during voir dire on jury composition and verdict, they arranged for three separate juries, one actual and two experimental, to hear one of ten criminal cases in the Northern District of Illinois.[45] The first experimental jury, the "English jury," consisted of a random sample of jurors from the pool who were not selected or questioned by the attorneys.[46] A second experimental jury, the "challenged jury," was made up of those jurors who were removed after peremptory challenges by either the prosecution or defense;[47] these jurors did not know which side had excused them, as all challenges were submitted on special forms.[48] All the jurors were treated simi-

44. Diamond & Zeisel, *supra* note 38, at 276.
45. *Id.*
46. *Id.*
47. *Id.*
48. *Id.*

larly during the trials; they were present in the courtroom during the entire trial, they were paid the standard juror fee, and they all deliberated in separate rooms before reaching a verdict.[49] However, the real jurors knew that they were deciding on an actual verdict, whereas the experimental jurors knew that their decisions would have no effect on the defendant.[50]

Over the series of ten trials, the results revealed a tendency toward more convictions for the experimental juries than for the real juries.[51] This suggests that jurors on the real juries may have used a higher standard for conviction than experimental juries who knew that their decisions would have no consequences for the defendants. While this early study has been critiqued for a number of reasons,[52] it still suggests that important variables to consider in mock-jury research are the role-playing ability of experimental jurors and whether simulations can adequately represent the same decision-making processes of real jurors.

While the results of the Diamond and Zeisel experiment suggested a trend toward more guilty verdicts by jurors who were aware of the experimental nature of their task, a later study produced results indicating the opposite.[53] To test the effect of consequentiality, Wilson and Donnerstein designed a series of studies in which a student was accused of stealing an exam and distributing the questions.[54] The student was tried in the context of a student judicial hearing.[55] The researchers manipulated the character and physical attractiveness of the defendant,[56] but more importantly, participants were informed either that they were taking part in an experimental trial or that the trial was real and would have actual consequences for the student.[57] The results showed that jurors in the real-consequences condition arrived at significantly more guilty verdicts than those in the hypothetical-consequences condition.[58] In addtion, jurors who thought they were participating in an actual trial recalled more of the trial evidence,[59] which suggests that consequentiality can affect the decisionmaking process as well as the outcome.

Kaplan and Krupa's results also suggested higher conviction rates when participants believed they were participating in a real trial, but

---

49. *Id.*
50. *Id.*
51. *Id.*
52. *See* Bray & Kerr, *Use of Simulation*, *supra* note 4; Kerr et al., *supra* note 39.
53. Wilson & Donnerstein, *supra* note 39, at 185.
54. *Id.* at 179.
55. *Id.*
56. *Id.* at 174, 182, 184.
57. *Id.* at 179.
58. *Id.* at 185.
59. *Id.* at 186.

only under some circumstances.[60] They were primarily interested in how the severity of the penalty would affect jurors' decision to convict, as well as how penalty severity might interact with the reality of the consequences. The authors reasoned that because the severity of the criminal convictions in the Diamond and Zeisel study was high, in comparison to the less severe consequences of stealing an exam in the Wilson and Donnerstein study, the severity of the case might determine the effects of consequentiality.[61] To test this hypothesis, Kaplan and Krupa designed an experiment where students were asked to judge the guilt of a fellow student who was accused of cheating.[62] The independent variables were the following: the strength of evidence against the defendant;[63] whether the jurors themselves or an authority would select the punishment; the possible penalty for conviction, which ranged from mild to moderate or from moderate to severe; and whether students were led to believe that their decisions were real and binding or part of an experimental simulation.[64]

There was a main effect for strength of evidence, such that when there was strong evidence of guilt, jurors were more likely to convict, without any interacting effects with the other variables.[65] However, when the evidence of guilt was low and the punishment was controlled by the students, the researchers found that those who were led to believe their decisions had real consequences were more likely to convict than those who made hypothetical decisions.[66] Kaplan and Krupa suggested that these results may have differed from Diamond and Zeisel's results because, in the present study, the jurors were judging other students in a case they had direct personal involvement with: cheating on an exam they had all taken.[67] The students might have adopted a lower conviction criterion and a stronger presumption of guilt when there was a chance someone could have gotten away with not writing the exam they all had to write.

While the research by Kaplan and Krupa suggests a possible explanation for the different results between the Diamond and Zeisel and the Wilson and Donnerstein studies, two other studies that investigated consequentiality found yet other constellations of

---

60. Kaplan & Krupa, *supra* note 39, at 8-13.
61. *Id.* at 2.
62. *Id.* at 3-4.
63. *Id.* at 3. The authors reasoned that when evidence points more clearly to guilt, it should be more difficult to deny the guilt of the defendant merely because a conviction entails severe consequences. *Id.*
64. *Id.*
65. *Id.* at 11.
66. *Id.* This result was obtained regardless of the level of punishment. *Id.* at 12.
67. *Id.* at 13.

results.[68] Kerr et al. also examined the effect of role-playing on juror verdicts by comparing the responses of "actual" and "hypothetical" juries in an experimental student discipline case.[69] The researchers created a hypothetical case involving a student charged with "malicious destruction of University property" at a nearby campus who, because of tremendous publicity, had been granted a change of venue.[70] Participants in the "actual jury" were led to believe that the school was experimenting with a new method of student discipline and that they would receive course credit for their participation in the student's trial.[71] To boost the trial's realism, a confederate, posing as a participant, was excused from the trial after indicating she was at the other campus in the past year and had read about the case in the school newspaper.[72] Another confederate, dressed in coat and tie, had been introduced as the cochairman of the University Committee on Student Discipline and reminded the students of the importance of their task.[73] Participants in the "mock jury" were to place themselves in the role of jurors and decide a previously tried case that would have no actual consequences.[74] All participants were given transcripts of the case and completed questionnaires both prior to and following deliberation.[75]

Results of the study suggested that there were no significant differences between individual and group verdicts of the two juries.[76] There was also a failure to find differences for sentence recommendation, deliberation time, and reasonable doubt criterion.[77] As well, prior to deliberation both groups indicated that they clearly understood their roles and the importance of their decision for the defendant.[78] The authors argued that these results should increase our confidence that mock juries can reliably be utilized in jury behavior research.[79]

Finally, a study by Suggs and Berman provided mixed support for the use of mock juries.[80] In an experiment designed to investigate the effect of mitigating testimony on mock-juror decisions, the researchers included a real-world consequentiality variable to enrich the ex-

68. *See* Kerr et al., *supra* note 39; Suggs & Berman, *supra* note 39.
69. Kerr et al., *supra* note 39, at 341-42, 345.
70. *Id.* at 342-43.
71. *Id.* at 342.
72. *Id.* at 343-44.
73. *Id.* at 343.
74. *Id.* at 345.
75. *Id.* at 344-45.
76. *See id.* at 347-48.
77. *See id.*
78. *Id.* at 347.
79. *Id.* at 351.
80. *See* Suggs & Berman, *supra* note 39.

ternal validity of the results.[81] Participants in the "no-consequence" condition were informed that they were being given a hypothetical plagiarism case in order to investigate student attitudes toward cheating.[82] Those in the "some-consequence" condition were told that the department chair had been confronted with a plagiarism case and wanted student input about appropriate action.[83] Additionally, they were informed that while their opinions would be carefully considered, there was no guarantee that they would be the deciding factor in the case.[84] The researchers also manipulated the presence of mitigating testimony, its source (defendant or third party), and its credibility (high versus low).[85]

Overall, there was no difference between participants in the no- and some-consequence conditions in their assignment of penalties. However, the pattern of results for the other variables was affected by consequentiality. When participants believed their decisions might have some consequence, less severe penalties were given when mitigating testimony for the defendant was presented (regardless of its source) than when it was not.[86] However, for participants whose decisions had no consequence, none of the other variables produced any reliable difference.[87] The researchers concluded that fears about the differences between real-world and role-playing behaviors among jurors may be well founded.[88]

C. Limitations of the Extant Studies

The results of the above studies provide little general consensus about the effect of role-playing and consequences on jury behavior. While a few studies provide support for those who criticize the external validity of experimental jury simulations, the studies themselves fall victim to methodological criticisms. First of all, it is important for any study investigating the effect of role-playing on juror behavior to manipulate the consequentiality variable successfully and be certain participants believe the cover story. If participants who are led to believe they are hearing an actual trial have doubts about its reality or the consequences their decisions will have on the defendant, then it makes little sense to liken their behavior to that of real jurors. In the study by Wilson and Donnerstein, the students participated in the study to receive extra credit in a psychology course.[89] Critics of this

81. *Id.* at 254-55.
82. *Id.* at 255.
83. *Id.*
84. *Id.*
85. *Id.* at 254-55.
86. *Id.* at 258.
87. *Id.* at 258-59.
88. *Id.* at 260.
89. Wilson & Donnerstein, *supra* note 39, at 178.

study note that it seems quite likely that the "real" jurors, expecting to participate in a psychological experiment, would doubt the reality of their task.[90] Unfortunately, Wilson and Donnerstein reported no data measuring the effectiveness of their cover story.[91]

Following the verdict responses of their participants, Kaplan and Krupa asked for their comments, concerns, or feelings about the procedure in order to assess its credibility.[92] Only three participants showed an awareness of the deception and were eliminated from the sample.[93] However, the authors failed to state exactly how these responses were elicited. If participants were indeed just asked to write any general "comments, concerns, or feelings" rather than asked specifically about the importance of the trial's consequences, it might be that not all instances where the cover story failed were identified.

Suggs and Berman did explicitly measure the effectiveness of their consequentiality manipulation for participants in the some-consequence condition, but they did not find particularly satisfying results. These participants were specifically asked if they doubted that the case was real; 30% answered that they had no doubt, 53% reported some doubt, 15% reported strong doubt, and 1% were convinced it was not a real case.[94] The participants were also asked whether they had doubts that their opinions would have any impact on the final decision, and 64% said yes.[95] Again, students volunteered for the study in order to fulfill course requirements, which may have tipped them off as to the simulated nature of the case.

Noting the importance of the consequentiality manipulation, Kerr et al. sought to manipulate it effectively by creating a mock trial that was as realistic as possible.[96] All participants were recruited to participate in a vaguely worded "Jury project."[97] Because the student participants expected to receive course credit for their participation, the researchers assumed that most would expect the jury project to be some sort of experiment.[98] For this reason, they used confederates and other props to enhance the credibility of the cover story.[99]

Kerr et al. also checked for participant suspicion of the experimental manipulation. Of 108 "actual" jurors, seven indicated some misunderstanding or suspiciousness of the cover story in their writ-

---

90. Bray & Kerr, *Use of Simulation*, *supra* note 4, at 113; Kerr et al., *supra* note 39, at 339.
91. *See* Wilson & Donnerstein, *supra* note 39.
92. Kaplan & Krupa, *supra* note 39, at 7.
93. *Id.*
94. Suggs & Berman, *supra* note 39, at 256.
95. *Id.*
96. *See* Kerr et al., *supra* note 39.
97. *Id.* at 341.
98. *Id.*
99. *Id.*

ten remarks, but they also indicated that they treated the case as genuine in case their suspicions were without merit.[100] Also, predeliberation questionnaires indicated that actual jurors clearly believed their verdicts would have consequences for the defendant and felt that their decisions were significantly more important for the defendant than did the mock jurors.[101]

Another limitation of these studies is that only two of the five included jury deliberations, an extremely important part of the adversarial process. A lack of jury deliberation in experimental studies has been identified as one of the major threats to external validity and policy relevance.[102] Despite the frequent suggestion that jury verdicts are essentially determined by the distribution of verdict preferences prior to deliberation,[103] there is reason to believe that deliberation can influence jury outcomes in certain situations. A recent meta-analysis of research that investigated deliberating juries between 1955 and 1999 argues that in one out of ten trials, deliberation results in a reversal of the verdict preference initially favored by the majority;[104] given the large number of jury trials each year, a substantial number of trial outcomes could therefore hinge on the deliberation process. Other than the field experiment by Diamond and Zeisel, the only laboratory study of consequentiality that included deliberation was the study by Kerr et al.; but even then, the deliberations were limited to forty-five minutes.[105] With unlimited time, dissenting members of hung mock juries might be more willing to acquiesce than those in the "actual" juries, due to the differential importance of reaching a final decision in the two conditions.[106]

Because of the difficulties associated with trying to manipulate experimentally mock jurors' perceptions of the consequences of their decisions, a more effective method might be to eliminate the consequentiality manipulation and use shadow juries during actual trials, as was done by Diamond and Zeisel.[107] However, this field method is not without its own limitations. The small number of cases and lack of experimental control over potentially confounding variables, such

---

100. *Id.* at 346.

101. *Id.* at 347.

102. *E.g.*, Diamond, *supra* note 4, at 564-65; Weiten & Diamond, *supra* note 4, at 78-79.

103. *E.g.*, REID HASTIE ET AL., INSIDE THE JURY 169-71 (1983); HARRY KALVEN, JR. & HANS ZEISEL, THE AMERICAN JURY (1966).

104. Dennis J. Devine et al., *Jury Decision Making: 45 Years of Empirical Research on Deliberating Groups*, 7 PSYCHOL. PUB. POL'Y & L. 622 (2001). Factors suggested by the meta-analysis for how deliberation may affect the verdict outcome include the deliberation style (evidence or verdict-driven), polling methods, and the collectively accepted interpretation of instructions.

105. Kerr et al., *supra* note 39, at 344.

106. Bray & Kerr, *Use of Simulation*, *supra* note 4, at 114.

107. Diamond & Zeisel, *supra* note 38, at 276.

as the juror selection procedure, make the results of the Diamond and Zeisel study difficult to assess as well.[108]

Finally, all of these studies involved criminal or quasi-criminal offenses. Consequentiality might have different effects in civil trials, where the outcomes usually involve monetary damages instead of criminal sanctions, such as imprisonment or other penalties (for example, failing a class or expulsion in the student honor-code-violation cases), and where a unanimous decision rule is less common.

IV. OTHER RELEVANT RESEARCH

An inspection of these studies leaves unresolved the question whether mock jurors behave similarly to actual jurors when they are aware that their decisions have no real consequences. In addition, the amount of deception necessary for effective consequentiality manipulations makes it very difficult to perform research on the consequentiality of jury decisions. Because of these inconsistencies and limitations, it is helpful to review other bodies of research to see if results from analogous experiments can be used either to support or undermine the use of mock trials in studying juror behavior. Relevant research exists within the jury research domain as well as from research on decisionmaking in nonlegal contexts.

A. Within the Legal/Jury Domain

As mentioned in Part III, Kaplan and Krupa suggested that it is important to include considerations of penalty severity when investigating consequentiality.[109] Intuitively, this makes sense: if we are considering the consequences that a jury's decisions (that is, some versus none) will have on a defendant, then the magnitude of those consequences (small versus large) might be expected to operate in an analogous fashion.

The Supreme Court addressed the issue of willingness to convict in death penalty cases in *Witherspoon v. Illinois*[110] and later in *Lockhart v. McCree*,[111] indicating the Court's concern with the impact of penalty severity on jurors' decisions. If jurors are assumed to focus on avoiding false convictions of innocent people, then the greater the penalty severity, the greater the perceived cost of error. As this perceived cost of error increases, jurors should require more evidence of guilt before voting to convict. With this shift in criterion, jurors

108. Bray & Kerr, *Use of Simulation*, *supra* note 4, at 113; Kerr et al., *supra* note 39, at 338.

109. Kaplan & Krupa, *supra* note 39.

110. 391 U.S. 510 (1968).

111. 476 U.S. 162 (1986).

should therefore be less likely to vote guilty when the penalty is more severe.[112]

The idea that more severe penalties lead to less willingness to convict is not new to the experimental arena. For example, an early study by Vidmar found that when mock jurors were presented with several different guilt decision alternatives (each carrying a mandatory sentence), the highest number of not-guilty verdicts came from the condition with only two extreme alternatives (not guilty or first degree murder) as opposed to conditions with several or less extreme levels of guilt.[113] Similar studies varying the latitude and severity of possible sentences offered to mock jurors have also found evidence suggesting that larger potential penalties lead to fewer convictions.[114] However, Freedman and his colleagues argue that because the design of these studies did not vary the available evidence for different charges of guilt, even though the law often requires additional evidence to prove more serious charges than less serious ones, jurors were essentially forced to vote not guilty for the more serious charges.[115] In their own research, Freedman et al. equated the evidence required for guilt for all charges and found no indication that mock jurors were less likely to vote guilty when penalties were relatively severe.[116] Thus, the extent to which penalty severity can serve as a useful analog for consequentiality is unclear.[117]

The underlying premise of this line of research is that the relationship between no consequence and some consequence is the same as the relationship between a small consequence and a large consequence. Is this a reasonable analogy? If so, then the difference between no jail time and six months imprisonment should be the same

---

112. Norbert L. Kerr, *Severity of Prescribed Penalty and Mock Jurors' Verdicts*, 36 J. PERSONALITY & SOC. PSYCHOL. 1431, 1431-32 (1978).

113. Neil Vidmar, *Effects of Decision Alternatives on the Verdicts and Social Perceptions of Simulated Jurors*, 22 J. PERSONALITY & SOC. PSYCHOL. 211, 211 (1972).

114. *See, e.g.*, Kalman J. Kaplan & Roger I. Simon, *Latitude and Severity of Sentencing Options, Race of the Victim and Decisions of Simulated Jurors: Some Issues Arising from the "Algiers Motel" Trial*, 7 LAW & SOC'Y REV. 87, 90, 96 (1972).

115. Jonathan L. Freedman et al., *Severity of Penalty, Seriousness of the Charge, and Mock Jurors' Verdicts*, 18 LAW & HUM. BEHAV. 189, 191 (1994).

116. *Id.* at 189.

117. In response to the Freedman et al. study, Kaplan criticized the research for, among other reasons, failing to address the consequences of conviction for the defendant or society. Martin F. Kaplan, *Setting the Record Straight (Again) on Severity of Penalty: A Comment on Freedman et al.*, 18 LAW & HUM. BEHAV. 697, 697 (1994). Further, Kaplan argued that rather than negate the effect of penalties, the research findings as a whole set limits on the penalty-severity effect and support the possibility of a criterion shift due to penalty. *See id.* at 698. The crucial variable, according to Kaplan, is whether the jurors believe the decision involves real consequences. *See id.* For a continuation of this debate, see Jonathan L. Freedman, *Penalties and Verdicts: Keeping the Record Straight*, 18 LAW & HUM. BEHAV. 699 (1994), and Martin F. Kaplan, *Keeping the Record Complete*, 18 LAW & HUM. BEHAV. 702 (1994).

as the difference between six and twelve months imprisonment.[118] Although that pattern might correspond well to people's perceptions, it is also possible that there is a discontinuity between no penalty and any penalty, in much the same way that there is a discontinuity between a lengthy prison term and the death penalty.[119]

B. Other Domains

The debate over the use of hypothetical situations to study human decisionmaking is not limited to the legal domain. It is a widely accepted practice to use hypothetical situations when investigating judgment and decisionmaking in a variety of contexts.[120] As with jury decisionmaking, researchers in these fields have occasionally (though also surprisingly seldom) considered the difference between decisions made with hypothetical versus real consequences.[121]

One research area in particular that has found a need for the use of hypothetical situations is the domain of risk-taking behavior. A common approach in this area is to study how people behave when faced with a risky situation, such as a hypothetical gamble, that can be framed in a number of different ways. Because of the dependency on hypothetical situations in this research, its validity has often been questioned, which has led to research comparing real and hypothetical decisionmakers.[122] Mirroring the inconsistent findings on consequentiality in the jury simulation literature, this research has found that people make riskier decisions when real consequences are used,[123]

---

118. The analogy holds even if length of imprisonment is perceived as a nonlinear scale.

119. Another way of expressing it is that the extreme alternatives—no consequences at all at one end of the continuum and capital punishment at the other end—are qualitatively different from the range of alternatives, such as prison sentences of varying duration, that comprise the large middle of the continuum. If so, then comparing some-penalty versus no-penalty involves comparing apples and oranges, whereas comparing more versus less severe penalties involves comparing apples of different sizes.

120. For a general overview of research on judgment and decisionmaking, see JONATHAN BARON, THINKING AND DECIDING (3d ed. 2000). A sizable literature has developed on the applications of behavioral-decision theory to law. *See, e.g.*, Gregory Mitchell, *Why Law and Economics' Perfect Rationality Should Not Be Traded for Behavioral Law and Economics' Equal Incompetence*, 91 GEO. L.J. 67 (2002); Jeffrey J. Rachlinski, *The "New" Law and Psychology: A Reply to Critics, Skeptics, and Cautious Supporters*, 85 CORNELL L. REV. 739 (2000).

121. *See, e.g.*, Hertwig & Ortmann, *supra* note 2, at 419-20. The examples in the following discussion are not intended to provide an exhaustive review of fields that have raised the consequentiality issue. For other domains, see, for example, Eva E.A. Skoe et al., *The Role of Reported Emotion in Real-Life and Hypothetical Moral Dilemmas*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 962, 964 (2002).

122. For examples of the literature, see Hertwig & Ortmann, *supra* note 2; and Kühberger et al., *supra* note 7.

123. *See* David B. Wiseman & Irwin P. Levin, *Comparing Risky Decision Making Under Conditions of Real and Hypothetical Consequences*, 66 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 241 (1996).

when hypothetical consequences are used,[124] and that risk-taking is not affected at all by a consequentiality manipulation.[125] In particular, in the area of Expected Utility Theory, the issue remains whether providing people with actual rather than imagined incentives would eliminate routinely observed behaviors showing that people fail to maximize the expected utility of their decisions. For example, research on framing effects assumes that the use of hypothetical situations adequately measures what decisions would be like in actual situations. The classic framing task asks people to imagine, for instance, that a disease is expected to kill 600 people.[126] One group is then asked to decide between two positively framed choices: either saving 200 people for sure or taking a one-in-three chance of saving all 600.[127] Another group chooses between negatively framed choices: either 400 people dying for sure or a two-in-three chance of all 600 dying.[128] In the positive-frame condition, people tend to prefer the first (sure) choice over the latter (risky) option, whereas in the negative-frame condition, people tend to choose the risky option over the sure choice.[129]

Does this effect hold true when real consequences are at stake? Wiseman and Levin varied risks between framed monetary gambles and time investments that were presented first as a hypothetical decision, then later as a real decision.[130] The results indicated no significant differences in participants' decision behavior between real and hypothetical situations;[131] but because the hypothetical choices preceded the real choices, the findings may simply be due to carryover effects, where participants might have just tried to appear consistent in their choices.[132]

Kühberger et al. also note that most experiments contrasting real and hypothetical outcomes suffer from a lack of nontrivial real outcomes and that the framing effect may disappear when real and hy-

---

124. *See* Terence Lafferty & Kenneth L. Higbee, *Realism and Risk Taking*, 34 PSYCHOL. REP. 827, 827-29 (1974); Paul Slovic, *Differential Effects of Real Versus Hypothetical Payoffs on Choices Among Gambles*, 80 J. EXPERIMENTAL PSYCHOL. 434, 434 (1969).

125. *See* Julie R. Irwin et al., *Hypothetical and Real Consequences in Experimental Auctions for Insurance Against Low-Probability Risks*, 5 J. BEHAV. DECISION MAKING 107 (1992); Paul E. Spector et al., *The Effects of Real vs. Hypothetical Risk on Group Choice-Shifts*, 2 PERSONALITY & SOC. PSYCHOL. BULL. 290 (1976). This variability of empirical findings could reflect the wide range of experimental methodologies used in psychological research on judgment and decisionmaking. Hertwig & Ortmann, *supra* note 2.

126. Amos Tversky & Daniel Kahneman, *The Framing of Decisions and the Psychology of Choice*, 211 SCIENCE 453, 453 (1981).

127. *Id.*

128. *Id.*

129. *See id.*

130. Wiseman & Levin, *supra* note 123.

131. *Id.* at 248-49.

132. Kühberger et al., *supra* note 7, at 1166.

pothetical decisions with small and large payoffs are compared.[133] To test this possibility, they assigned participants to either a positive or a negative framing condition; the participants then made both hypothetical and real decisions with both small and large payoffs. They found that the payoff size had a significant effect, such that participants more often chose the sure option when payoffs were large rather than small.[134] They also found the expected framing-by-payoff-size interaction—with small payoffs, there was no framing effect, but with large payoffs, the typical framing effect appeared.[135] Of greatest relevance to the present discussion, there were no observed effects for real versus hypothetical payoffs.[136] In sum, Kühberger et al. found that when people are presented with an actual risk-decision task, their decisions are similar to when they are presented with a hypothetical risk-decision task: risks are taken with small payoffs regardless of frame and with large payoffs when negatively framed.

So what does this mean for jury simulations? On their face, the results of the Kühberger et al. study give some hope for the use of jury simulation studies, in that the decision process appeared not to depend on whether the decision task was real or hypothetical.[137] However, there are several fundamental differences between the two tasks that make the application of the results not as widely applicable as one might hope. First, there is the nature of the task. Unlike the risk-decision studies where participants are asked to make a decision between similar monetary gambles that result in statistically equal outcomes, juries are given the task of weighing evidence and deciding facts in arriving at a well-reasoned decision with quite divergent alternative outcomes. Secondly, although there are consequences involved in juror decisionmaking for the jurors themselves, the principal consequences are for the litigants in the case;[138] whereas the consequences involved in making decisions about risk or gambles will primarily (and sometimes exclusively) affect the person

---

133. *Id.*

134. *Id.* at 1169-70.

135. *Id.*

136. *Id.* at 1169. There were also no interactions between consequentiality and the other variables.

137. In a more extensive review of the literature, Hertwig and Ortmann reached a different conclusion, finding that although real consequences in the form of actual payments "do not guarantee optimal decisions, in many cases they bring decisions closer to the predictions of the normative models." Hertwig & Ortmann, *supra* note 2, at 395. However, they noted the considerable variability in the effects of consequentiality across various decisionmaking contexts. *Id.* at 395-96. They also excluded from consideration studies in which there was no clear standard for optimal performance; whereas in jury trials, the optimal or "correct" outcome is typically unknown. *See id. See generally* Colin F. Camerer & Robin M. Hogarth, *The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework*, 19 J. RISK & UNCERTAINTY 7 (1999).

138. *See supra* note 6 and accompanying text.

making the decision. This distinction could induce different decision processes in the two types of situations.

V. ALTERNATIVES AND RECOMMENDATIONS

Is the enterprise of studying juries through experiments absurd? In light of the concerns presented here, are we (that is, experimental jury researchers) wasting our time and efforts? As with virtually all literature reviews, this one concludes that more research is needed on real versus hypothetical jury decisions. We found only five direct comparisons in the jury domain, and these studies produced inconsistent findings. The question has been addressed to some extent in other areas of research, such as the broader decisionmaking literature, yet without shedding much light on the differences between real and simulated decisionmaking in terms of either decision process or outcome. In designing such research, future scholars should carefully consider whether consequentiality matters from a legal or a theoretical perspective.

Most courts have demonstrated a clear reluctance to base legal holdings on experimental research findings. Therefore, if psycholegal researchers wish to produce legally relevant research, it seems imperative that they (that is, *we*) determine whether, and how, consequentiality affects juror decisions.[139] The corpus of studies is as yet insufficient to reach any definitive conclusions. Adding to this corpus is complicated, however, by the numerous practical difficulties described in Part III of this Article. What alternatives then exist?

One alternative is to conduct research that, while not manipulating consequentiality directly, manipulates variables analogous to consequentiality. This is the approach taken in the research on penalty severity.[140] Because of the difficulties inherent in experimentally manipulating the consequences for the litigants, another approach would be to manipulate the consequences for the mock jurors. As described above, the notion of consequentiality subsumes the jurors as well as the litigants. One could vary the consequences for the mock jurors by, for example, varying their task motivation (perhaps by offering some sort of reward for good performance, such as a financial incentive)[141] or accountability for their decision.[142] If such variables

---

139. This is by no means to suggest that psycholegal research *must* have legal applications for it to be worthwhile. Jury simulations can provide a useful context within which to test psychological theories and explore fundamental processes. In such "Stage One" research, the apparent realism of the legal situation is not critical. *See* Diamond, *supra* note 4, at 563. As Diamond notes, the problem arises when one seeks to make policy recommendations from that research. *Id.* According to Diamond, "Stage Two" research, which more closely approximates real legal situations, is then much more likely to be effective. *Id.*

140. *See supra* notes 109-19 and accompanying text.

141. Motivation and incentive have robust effects on behavior across a variety of domains. *See, e.g.*, Camerer & Hogarth, *supra* note 137; G. Douglas Jenkins, Jr. et al., *Are*

affect either the process or the outcome of mock jurors' decisions, then we need to be cautious about generalizing from mock jurors to real jurors.

Another alternative is to conduct experimentation on real juries in field settings. From a scientific perspective, there are trade-offs in conducting field research in lieu of experimental research. Most prominently, field research sacrifices experimental control, which is essential for making causal inferences, and it is usually more difficult and expensive to conduct.[143] Doing field research on juries carries the additional complication that, with isolated exceptions,[144] it is illegal to observe juries during deliberation; and even if deliberation is not of primary interest, it still requires, at a minimum, the approval of the participating judges.

Nonetheless, jury researchers have occasionally succeeded at conducting field studies. Professors Heuer and Penrod conducted several field studies on the effects of procedural innovations such as allowing jurors to ask questions and take notes;[145] more recently, the Arizona Jury Project addressed these and similar questions.[146] In both cases, participating judges allowed juries to be randomly assigned to different experimental conditions. Both projects have led to policy changes being adopted by the courts.[147] Calls for more field experiments on juries have been issued from prominent psychological and legal scholars.[148] Because of the high external validity of field studies, their findings are harder for courts to dismiss on methodological

---

Financial Incentives Related to Performance? A Meta-Analytic Review of Empirical Research, 83 J. APPLIED PSYCHOL. 777 (1998).

142. *See generally* Lerner & Tetlock, *supra* note 11.

143. *See generally* Aronson et al., *supra* note 2; THOMAS D. COOK & DONALD T. CAMPBELL, QUASI-EXPERIMENTATION: DESIGN & ANALYSIS ISSUES FOR FIELD SETTINGS (1979).

144. For example, the Arizona Jury Project collected data on deliberating juries and tested the effectiveness of a number of procedural reforms, such as allowing jurors to discuss the case during trial. *See, e.g.*, Shari Seidman Diamond, The Impact of Juror Discussions During Trial: The Arizona Jury Project, Address at the American Psychology-Law Society Meeting (Mar. 9, 2002).

145. *See* 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); Larry Heuer & Steven Penrod, *Juror Notetaking and Question Asking During Trials: A National Field Experiment*, 18 LAW & HUM. BEHAV. 121 (1994); Larry Heuer & Steven Penrod, *Trial Complexity: A Field Investigation of Its Meaning and Its Effects*, 18 LAW & HUM. BEHAV. 29 (1994).

146. *See supra* note 144.

147. *See, e.g.*, Leonard Post, *Study Endorses Midtrial Juror Chat: Arizona's New Rule Is Seen as Helpful*, NAT'L L.J., Jan. 19, 2004, at 1.

148. *See, e.g.*, Paul E. Meehl, *Law and the Fireside Inductions: Some Reflections of a Clinical Psychologist*, *in* LAW, JUSTICE, AND THE INDIVIDUAL IN SOCIETY: PSYCHOLOGICAL AND LEGAL ISSUES, *supra* note 4, at 10; Laurens Walker, *Perfecting Federal Civil Rules: A Proposal for Restricted Field Experiments*, LAW & CONTEMP. PROBS., Summer 1988, at 67.

grounds.[149] As the field of jury research continues to mature, we hope to see more of them.

149. Unless, of course, the increased external validity comes, as it sometimes does, at the cost of decreased internal validity, in which case the field studies could simply be dismissed on other grounds. *See, e.g.*, Aronson et al., *supra* note 2, at 129-33; Saks, *Jury Experiments*, *supra* note 4, at 4-9; Walker, *supra* note 148, at 73-74. With few exceptions, however (for example, Free v. Peters, 12 F.3d 700 (7th Cir. 1993)), appellate courts have tended to emphasize concerns of external validity. *See, e.g.*, Acker, *supra* note 23, at 80-82; Diamond, *supra* note 4, at 569; Tanford, *supra* note 21, at 144-48. This suggests that a greater proportion of field experiments would substantially further the goal of getting courts to take experimental research seriously.