



School of Law
UNIVERSITY OF GEORGIA

Digital Commons @ University of Georgia School of Law

Scholarly Works

Faculty Scholarship

1-1-2017

Jury Simulation Goals

Jonathan J. Koehler

Assistant Professor of Law *Northwestern Pritzker School of Law*, jay.koehler@northwestern.edu

John B. Meixner Jr.

Assistant Professor of Law *University of Georgia School of Law*, John.Meixner@uga.edu



Repository Citation

Jonathan J. Koehler and John B. Meixner Jr., *Jury Simulation Goals* (2017),
Available at: https://digitalcommons.law.uga.edu/fac_artchop/1559

This Book Chapter is brought to you for free and open access by the Faculty Scholarship at Digital Commons @ University of Georgia School of Law. It has been accepted for inclusion in Scholarly Works by an authorized administrator of Digital Commons @ University of Georgia School of Law. [Please share how you have benefited from this access](#) For more information, please contact tstriepe@uga.edu.



NORTHWESTERN UNIVERSITY SCHOOL OF LAW

PUBLIC LAW AND LEGAL THEORY SERIES • NO. 14-42

Jury Simulation Goals

in

The Psychology of Juries:
Current Knowledge and a
Research Agenda for the Future
(Margaret Bull Kovera, ed., forthcoming)

Jonathan J. Koehler

Northwestern University School of Law

John B. Meixner

Northwestern University School of Law

**Northwestern University Department of
Psychology**

Jury Simulation Goals

Jonathan J. Koehler and John B. Meixner*

DRAFT: 6-13-14 (no quotes)

Juries have long been of special interest to psychologists who study the way individuals and groups make decisions. Juries are charged with making some of the most important decisions in our society (including, in some cases, whether a criminal defendant lives or dies) and the secrecy of their deliberations adds to their mystique. For decades, psychologists interested in legal decision-making have conducted a type of controlled experiment, known as a *jury simulation* or mock trial. Jury simulations are experimental studies in which the researcher attempts to construct a setting that mirrors, to some extent, a jury decision-making environment. Jury simulations vary widely in terms of participants, materials, physical settings, realism, methods, independent variables, dependent measures, and other experimental features.

What are the goals that researchers who conduct jury simulations have or should have? Drawing on Pennington and Hastie (1981), we identify three primary goals: (1) develop theory, (2) describe how juries perform, and (3) improve the jury process. This short list of goals is not exhaustive (other goals are possible), nor are the goals themselves mutually exclusive. For example, many jury researchers initially seek to provide insight into basic human cognition by describing a systematic influence on the behavior of participants in a jury simulation (goal 1). Once the experimental demonstration is established, the researchers may then argue that real juries will behave similarly (goal 2), and that this result points the way toward improving some aspect of the legal system (goal 3). Having multiple goals is not itself a problem. But

* Jonathan J. Koehler is the Beatrice Kuhn Professor of Law at Northwestern University School of Law. He has a B.A. in philosophy from Pomona College and an M.A. and Ph.D. in behavioral sciences from the University of Chicago. Prior to joining Northwestern Law in 2010, he was a

researchers should be aware that the validity concerns associated with each goal are different and may even tug in opposite directions.

Validity Concerns

Cook and Campbell (1979) famously identified various threats to the validity (trustworthiness) of many types of experimental research. These include threats to *internal validity* (the causal connection between independent and dependent variables), *external validity* (the generalizability of results to other samples and settings), and *statistical conclusion validity* (the relationship between variables). *Construct validity* (how well the operationalized variables capture the underlying constructs that they are supposed to represent) and *ecological validity* (how well the experimental setting mimics real world settings of interest) may also be of more or less concern depending on how a simulation is designed. For our purposes here, internal validity, external validity and ecological validity are most relevant.

Internal validity: Researchers who conduct jury simulations largely to advance basic theory (goal 1) are generally interested in using research methods that maximize *internal validity*. Most people agree that properly conducted randomized experiments (including jury simulations) have relatively strong internal validity. In these studies, significant changes in the dependent variables of interest (e.g., damages awarded, subjective probabilities of guilt, or verdict) may be traced directly to the causal influence of one or more manipulated independent variables (e.g., use of emotional arguments, specific judicial instruction, etc.). As a general matter, internal validity is of concern in experimental designs, because without strong internal validity, researchers cannot draw valid conclusions about the result of their manipulations, and thus cannot proceed to the second question of whether those results are generalizable.

External validity: External validity is concerned with how well the results of a study generalize across various people, times, settings, and other specific elements in the study. Studies that find consistent results across different types of participants and experimental contexts and stimuli, have higher external validity than those that do not. Jury simulations commonly sacrifice external validity in order to gain more control, precision, and hence greater confidence about hypothesized causal relationships in the examined context. In other words, jury simulations commonly favor internal validity over external validity or generalizability. Field studies - in which the behaviors of jurors in various actual cases are observed – make the opposite tradeoff. That is, field studies commonly sacrifice control and experimental precision but their emphasis on naturally occurring settings gives them a leg up in terms of external validity. As others have observed, internal and external validity are often at odds, where efforts to increase one type of validity often come at the cost of reducing another type of validity.

Ecological validity: Ecological validity is related to external validity, but is more directly concerned with how well the experimental setting mimics the real world setting or settings of interest.¹ Jury simulations that include many key features of actual jury trials (e.g., live opening arguments from attorneys, cross-examination of witnesses, jury deliberation, etc.) have higher ecological validity than short, paper and pencil studies (Breau & Brook, 2007). Whereas there is general agreement that much jury research suffers from low ecological validity, there is less agreement about the extent to which this shortcoming matters (compare, for example, Vidmar, 1979, 2008; and Weiten & Diamond, 1979 with Penrod et al., 2011). Our view is that the negative impact of low ecological validity on a study depends largely on the investigators' goals and claims. Researchers interested primarily in goal 1 — advancing basic psychological theory

¹ This working definition of ecological validity is widely used and accepted (see e.g., Penrod, Kovera & Groscup, 2011; Robbennolt, 2002-2003). However, the term originally had a different meaning that some have argued should be retained (Hammond, 1998).

— need not worry as much about this issue. However, researchers interested in goals 2 and 3 — describing jury behavior and impacting legal policy — should be more concerned. In a sense, this point is self-evident: an investigator’s goals must play a large role in the design of his or her studies. At the same time, our view is that many, if not most, jury simulations would likely be conducted differently if investigators took this recommendation to heart at the earliest stages of their research programs.

I. Goal 1: Theory Development

As indicated above, theory development (goal 1) is a goal that many jury researchers have and should have. Psychologists and others who are interested in addressing basic questions about cognition, decision-making, and social dynamics will find that the jury simulation paradigm — in which participants act as mock jurors in a civil or criminal case — often offers an appropriate method for studying these questions.

Consider, for example, a study one of us performed that was designed to test a general theory for how people think about low probability events (exemplar cuing theory). Koehler and Macchi (2004) presented the results from two controlled, highly simplified simulations that included statistical evidence from a hypothetical criminal case. From an ecological validity standpoint, the study was lacking. Jurors did not receive judicial instructions, watch witnesses undergo cross-examination, hear objections, receive comprehensive judicial instructions, or deliberate in groups. Although the participants were selected from two countries (Italy and the United States), the external validity of the study was not particularly high either due to these same considerations as well as the use of a rather narrow range of range of stimulus materials (focused on statistical DNA evidence). In light of these ecological shortcomings, one might

legitimately question the authors' claim that the results could have implications for how lawyers may wish to present statistical evidence at trial. But the central purpose of the study was theory testing. The legal context was simply a convenient one in which to explore the notion that people may give more and less weight to low frequency statistical information as a function of the ease with which they are able to generate relevant exemplars.

Similar comments could be made about other legal decision-making studies. For example, James Davis's work on Social Decision Scheme (SDS) theory (see e.g., Davis, 1973; Stasser & Davis, 1981) was motivated by a desire to understand the way small groups aggregate the individual judgments and decisions of its individual members. Jury simulations that included deliberation provided an obvious context for Davis's work and related follow-ups in this area. Davis and other SDS researchers generally did not go to great lengths to create a realistic jury atmosphere, but this shortcoming did not interfere with the primary goals of the research program.

In some cases, psychologists might conduct their studies in a jury setting simply to show how a group (including a jury) *could be induced* to behave, rather than to show how they often *do* behave. In these situations, neither the study's ecological validity nor its external validity would be of great concern. Suppose, for example, that a researcher is interested in demonstrating that groups can be bullied by a single member into doing something that all of its members know to be illegal, inappropriate, or unethical. A jury simulation might be conducted in which the bully is introduced into a mock jury to convince the jury to find the defendant guilty of charges that are obviously false. Such a study, which may be quite artificial and may not have direct applicability to actual jury deliberations, might be informative simply as a way to test the null hypothesis that groups of laypeople would not behave in that manner (cf. Mook, 1983, p. 382).

These points are not new. Both critics and defenders of the current jury simulation paradigm have long noted that theory-testing is a legitimate purpose and that the reduced ecological and external validities that commonly accompany such research do not themselves undermine the value of these studies (Kerr & Bray, 2005; Weiten & Diamond, 1979). But as Vidmar (1979) and others have noted, theory testers commonly suggest that their research has significant implications for how real juries behave and what types of reforms should be implemented based on their results.

In many cases, the practical implications claimed for jury simulations reach well beyond anything that is justified from the research. Indeed, Thompson (1993) concluded that “[t]he credibility of the field as a whole has been damaged by researchers who have made sweeping and misleading generalizations about the real legal system based on” brief, unrealistic jury simulations (p. 205). Others counter that failure to seek insights from these studies for actual juries would undersell our own research efforts (Kerr & Bray, 2005).

Reasonable people may agree with one view or the other on this point. Our view is that those who wish to draw implications from jury simulations to real life jury trials should and could do more at the design stage to support their speculations regarding goal 2 (describing jury behavior) and goal 3 (impacting legal policy) matters. The fact that short, unrealistic simulations may have value beyond theory development and testing should not be used to justify failures to introduce study design features that increase the validity of goal 2 and goal 3 inferences. As we discuss in the following section, there is no question that introducing design features that enhance the ecological validity of jury simulations is a time-consuming and costly endeavor. But it will often be time and money well spent. If the effects that our brief studies identify persist in more realistic legal environments, our confidence in the importance of those effects for

jury decision-making increases. Of course, many effects identified in goal 1 studies will not affect relevant legal measures (e.g., verdicts) when realism is increased to make those studies more appropriate for goals 2 and 3. That is, some genuine effects may simply wash out in the more complex environment of a full-length trial (Dillehay & Nietzel, 1980). Other effects may interact in complex ways with various trial features. These things are important to know and are therefore worth the time and effort required to learn them.

II. Goal 2: Modeling Juries

As noted in Section I, jury simulations will often be a legitimate methodological option for researchers seeking to advance basic psychological theory. But we also noted that, in these types of studies, the jury context might be of secondary importance. Researchers conducting studies that focus on basic theory will often be more interested in demonstrating or testing a particular phenomenon, leaving it to subsequent studies to identify whether that phenomenon is broadly applicable or whether it persists in a particular applied setting (e.g., real jury trials).

Some studies, however, focus primarily on what might be called *jury modeling*. Researchers conducting jury modeling studies are expressly interested in borrowing from the theoretical frameworks and phenomena identified by past work, and using them to learn more about the decision processes of juries. Jury modeling is fundamentally different from theory development, and thus is best achieved by attending to a different set of considerations. By its nature, jury modeling is an *applied* goal that can only be met through applied means: if jury modelers want their studies to persuade scientists and policy makers about how real juries operate, they must model the aspects of the trial that are relevant to the decision-making of real

jurors. In our view, the central purpose of such studies is to allow extrapolation from the lab context to the real jury context.

While these separate goals and corresponding separate considerations might seem obvious, we think they are worth considering carefully when planning experiments, because jury modelers are faced with unique challenges not present in some other contexts. The jury and trial context are exceptionally difficult to model (Pennington & Hastie, 1981). A number of unique factors are present: the voir dire process winnows the jury-eligible population in case-specific (and perhaps attorney-specific) ways, and jury trials are notoriously lengthy and complex, and the stakes of the jury's decisions are high. The difficulties of simulating a jury trial in a controlled laboratory setting have been discussed in various reviews (e.g., Bornstein, 1999; Bray & Kerr, 1979; Diamond, 1997; Weiner, Krauss, & Lieberman, 2011; Weiten & Diamond, 1979). Below, we provide a brief overview of what we regard to be some of the most significant aspects of the jury decision-making context that are commonly not accounted for in jury simulations.

Stakes & Duration

Several aspects of the jury's role are likely almost impossible to simulate in the lab. First, real trials are typically *high stakes* affairs. The decisions juries make can implicate large amounts of money in civil cases and determine the freedom of defendants in criminal trials. In contrast, mock jurors in jury simulations know that no real consequences will result from the choices they make. Whereas jurors in real cases may agonize for hours or even days over their decisions, mock jurors contemplating their hypothetical decisions will likely not do the same.

Even putting aside concerns about effort, there may be something about the sheer enormity of the stakes that influences real jurors' decisions but cannot be modeled in the lab. In 2013, a jury found in favor of the plaintiffs in a class-action price-fixing case against Dow

Chemical for \$400 million. In accordance with federal antitrust law, this verdict was then trebled, yielding a \$1.2 billion final verdict (*In re Urethane Antitrust Litigation*, M.D. Kan. 2013). In 2005, actor Robert Blake was acquitted of murder by a jury that deliberated for nearly nine days (Associated Press, 2005). The stakes and intensity associated with such civil and criminal decisions surely played a role in the minds of jurors. Yet neither can be recreated with any degree of confidence in the laboratory.

Occasionally, researchers attempt to simulate the high stakes of a real-world trial (Diamond & Zeisel, 1974; see also Bornstein & McCabe, 2005). However, such efforts are rare and may raise ethical concerns. In one interesting study, Breau & Brock (2007) used deception to make half of their participants in a mock law school honor code hearing think that they were participating in an actual hearing. Students who thought the hearing was real spent more time deliberating and reached more lenient outcomes than students who knew the hearing was staged. This result, though not dispositive, suggests that our general inability to replicate the stakes associated with real trials may point to an important limitation on our ability to predict real jury behavior. Likewise, recently accumulated data that point to the importance of hard-to-replicate emotion on jurors' judgments and decisions (see e.g., Salerno & Bottoms, 2009) suggest that jury simulations not accounting for this element may miss an important aspect of jury decision-making.

The *duration and volume of information* presented to jurors at trial is a second set of considerations that is also difficult, if not impossible, to model in the lab. More than half of all civil trials last at least two days, and many cases last for weeks (Galanter, 2004). Most jury decision-making studies take much less than an hour from start to finish; yet the average jury in simple misdemeanor federal cases deliberates for more than three hours (Mize, Hannaford, &

Waters, 2007). Not only does this create a potentially important disparity between jury studies and jury experiences in terms of detail and complexity, but it also means that mock jurors are unlikely to experience the slow pace and potential boredom that may come with a long trial. This problem may not be trivial, particularly if the process of culling the most relevant evidence from the sea of information presented in actual trials requires a type of heuristic processing that is not required in shorter studies (e.g., Pennington & Hastie, 1986).

Other Trial Features

While the trial aspects of stakes, duration, and information volume are nearly impossible to simulate, other potentially important features of the trial may be modeled in the laboratories of researchers who have sufficient time, resources, and motivation. These features include (1) the demographics of the jurors, (2) the evidence presentation modality, (3) the judge—jury interaction, (4) group deliberation, and (5) the questions answered by the jury. Many of these characteristics, and the difficulty of modeling them, have been written about elsewhere (see e.g., Dillehay & Nietzel, 1980; Vidmar, 1979, 2008; Weiten & Diamond, 1979). Here, we review the topic briefly, and focus on what we see as the most pressing issues facing jury-modeling researchers.

Jurors: Perhaps the most visible difference between real-world juries and mock juries in a typical jury-modeling experiment is the sample of jurors themselves. Whereas real-world juries are at least theoretically constructed of a random sample of jury-eligible citizens,² the majority of jury-modeling studies use a more readily available population: students. This problem is well-known (for an early review, see Weiten & Diamond, 1979), though there is no

² There are data indicating that actual juries sometimes skew in various ways, such as by age and race (e.g., Walters, Marin, & Curridan, 2005), gender (for review, see Devine, 2012, p. 111-113), and authoritarianism (Narby, Cutler, & Moran, 1993), though gathering accurate data on jury skew can be difficult due to poor record keeping by the courts (Rose & Abramson, 2011).

consensus as to the significance of the problem. One review concluded that there is rarely a main effect of sample when comparing the verdicts of student and nonstudent mock-juror participants (e.g., Bornstein, 1999). However, other recent studies suggest that mock-jury population may *interact* with other experimental variables (for a review, see Weiner et al., 2011). For example, several studies published in a special issue of *Behavioral Sciences and the Law* demonstrated that the jury sample (student vs. nonstudent) interacted with such variables as the type of damages in a medical malpractice case (punitive vs. compensatory, Fox, Weingrove & Pfeifer, 2011) and ethnicity in a Title VII discrimination case (Schwartz & Hunt, 2011). More significantly, there are few available data about how juror type interacts with a variety of other variables, including an important but commonly neglected variable: jury deliberation (Nunez, McCrea, & Culhane, 2011). The conclusion, then, is an uncertain one. Whereas the makeup of a mock jury may not matter in certain situations or with regard to certain dependent measures, it may matter strongly in others, making it difficult for jury modelers to know whether drawing mock jurors from a student population is a reasonable way to learn about real jury behavior.

Evidence Presentation Modality: An area that has received less attention in the literature is the modality in which mock trial testimony is presented to participants in jury-modeling studies. In real trials, jurors are presented with stimuli almost entirely through visual and auditory methods—they *listen* to witnesses’ testimony and *watch* witnesses as they testify, assessing both the content and the demeanor of the witnesses. And there are visual and auditory cues at trial beyond the witness box—jurors also listen to the questions of attorneys and observe their demeanor. And there are visual and auditory “offstage” events occurring at trial as well: jurors may visually observe the parties seated with their counsel (including, notably, criminal

defendants)³ and other individuals in the courtroom, such as audience members (Rose, Diamond, & Baker, 2010). Despite the nearly exclusive visual and auditory stimulus presentation method at trial, and the increasing ease of conducting jury simulations that include visual and audio components, most jury research is still conducted instead using entirely written stimulus materials.

Is such a difference important? The data are not clear. Bornstein (1999) surveyed 11 studies that presented both written trial materials and more realistic materials, such as videotaped mock trials or live testimony. He noted that the presentation medium only affected the mock jurors' verdicts in three of the 11 studies, and among those in which main effects were found, the direction of the effects were inconsistent. Furthermore, Bornstein (1999) did not find that presentation medium interacted with other variables (such as the type of testimony presented), and recent data from Pezdek, Avila-Mora, and Sperry (2010) are consistent with this conclusion as well.

Nevertheless, we think the data collected to date are not sufficient to justify a conclusion that presentation mode is irrelevant to jury modeling efforts. First, presentation modality (e.g., written vs. video stimuli) has not yet been tested in a variety of situations where there is reason to suspect that modality may matter. Substantial research indicates that visual cues often affect judgments, including those made in a legal context. For example, the confidence level of mock expert witnesses (manipulated by tone of voice, postural awkwardness, eye contact, etc.) affects the perceived credibility of those witnesses as well as the verdicts that mock jurors render, even when the experts use identical words (Cramer, Brodsky, & DeCoster, 2009). Such an influence could not be captured in studies that use written stimuli. Similarly, Brodsky (2009) showed that

³ Indeed, the Supreme Court has recognized that requiring criminal defendants to be tried while wearing identifiable prison clothes violates the prisoners' due process guarantees (*Estelle v. Williams*, 425 U.S. 501 (1976)).

the perceived likability of witnesses as influenced by demeanor and appearance could affect their perceived credibility. Because demeanor is impossible to capture in written materials, modality would seem to matter in at least some contexts. Also, real jurors commonly discuss the “offstage” behavior of parties and other individuals in deliberations, though it is not clear whether those behaviors affect liability outcomes (Rose, Diamond, & Baker, 2010). Finally, a few studies have detected presentation modality effects in limited contexts. Heath, Grannemann, & Peacock (2004) showed that presentation modality interacts with the perceived honesty and emotionality of witnesses, and Nietzel, McCarthy, & Kern (1999) showed that modality can affect verdicts by interacting with jurors’ attitudes (for further discussion, see Penrod et al., 2011). For all of these reasons, we think it would be premature to accept the convenient conclusion that presentation modality is unimportant to jury modeling, though we agree that the data are currently inconclusive.

The Judge—Jury Interaction: Another relatively unexamined aspect of jury decision-making at trial is the judge—jury relationship. Unlike the parties and their representatives, judges are widely viewed as disinterested, fair-minded, and authoritative. As such, judicial words and behaviors may have a significant impact on jurors. For example, judges who frequently overrule one party’s objections while sustaining the other’s may inadvertently provide a signal to jurors about the relative credibility of the arguments offered by the two parties. Although empirical data on judge-jury interactions are sparse, Blanck, Rosenthal and Cordell (1985) showed that judges’ expectations about trial outcomes affected their verbal and nonverbal behavior, and these behaviors in turn influenced jurors’ decisions. Other research indicates that jurors may not always respond to the judge’s instructions and interventions as intended. For example, mock jurors have difficulty disregarding inadmissible evidence following an

instruction from the judge to do so (for a meta-analysis, see Steblay, Hosch, Colhane, & McWethy, 2006). There is also some evidence that mock jurors may not heed a judge's instruction to use evidence for one purpose but not another (Wissler & Saks, 1985). While these types of effects do not tell us whether omission of the judge from jury simulations harms the external validity of the research, they do hint that the judge—jury relationship may be a complex one, and that judicial actions and inactions may affect juries in complex ways.

Group Deliberation: A more well-explored and frequently simulated aspect of the jury is its group nature, especially at the deliberative stage. Numerous studies have examined the effects of deliberation on jury decision-making (for reviews, see Diamond, 1997; Nuñez, McCrea, & Culhane, 2011; Devine, 2012, chapter 7), and it appears that deliberation impacts some of the decisions jurors make. Here we touch on a few ways in which deliberation likely matters, particularly with respect to the potential for group deliberation to mitigate or exacerbate *errors* in jurors' understanding.

Although intuition might suggest that groups should be better than individuals at avoiding factual errors (such as mathematical computation errors or use of an improper evidentiary standard), the data do not clearly support this view. Indeed, a number of studies have demonstrated that groups often make more errors, and make more extreme errors, than individuals. For example, Tindale (1993) demonstrated that groups might make more conjunction errors—probability errors in which decision makers improperly assume that more specific conditions are more probable than a single general condition—than individuals. Drawing on a larger set of findings, Kerr and Tindale (2004) concluded that individual-level decision biases that promote suboptimal decision strategies are generally exacerbated in groups and lead to even worse decisions. Relatedly, some studies have found that group deliberation

increases jurors' confidence in their verdicts, without any corresponding increase in their understanding of the relevant facts. For example, Kaye, Hans, & Dann, et al. (2007) asked mock jurors to watch a videotaped trial involving Mitochondrial DNA evidence, after which they made a series of judgments about the evidence and the case before and after deliberating in groups. They found that group deliberation increased jurors' confidence in their own judgments. Post deliberation, the proportion of jurors who reported a very high degree of uncertainty about the defendant's guilt (i.e., $P(\text{Guilt}) = 50\%$) dropped from 16% to 10%. Correspondingly, the proportion of jurors who were highly certain about the defendant's guilt (i.e., $P(\text{Guilt}) = 100\%$) increased from 8% to 17% following deliberation. Despite this increase in confidence, deliberating jurors did not appear to have a better grasp of the evidence than they had before deliberating. For one thing, their overall scores on a DNA evidence comprehension test increased by just 3% following deliberation (from 70% to 73%). More strikingly, a whopping 40% of jurors – both before and after group deliberation – mistakenly believed that the DNA evidence was “completely irrelevant” because it was possible that people other than the defendant contributed the hairs in question. This result fits well with Shari Diamond's speculation that when most jurors are confused, “deliberation may simply reinforce the inaccuracies of the majority” (Diamond, 1997, p. 565).

Deliberation may exert different types of influence, some positive, in other contexts. For example, deliberation may produce more nuanced reasoning among jurors (McCoy, Nunez, & Dammeyer, 1999), reduce the biasing effects of inadmissible evidence (London & Nunez, 2000; Wheatman & Shaffer, 2001), and may induce jurors to give larger damage awards to plaintiffs than they otherwise would give (Diamond & Casper, 1992; Schkade, Sunstein, & Kahneman, 2000). Deliberation may impact legal judgments in many other ways as well (for a review, see

Salerno & Diamond, 2010). Our point here is a fairly modest one: in some contexts that jury-modeling researchers likely care about, deliberation may impact the judgments and decisions juries make. Deliberation may also interact with other aspects of the trial. If true, then it would seem unwise for jury modelers to ignore deliberation when planning their studies, except perhaps in those limited situations in which there is good reason to believe that deliberation will not impact the independent or dependent variables of interest.

Questions Answered: Finally, we note that many of the questions that mock jurors and juries answer in jury modeling research are different from those answered by real juries. Researchers commonly ask individual mock jurors to provide a variety of judgments pertaining to such matters as witness credibility, evidentiary strength, probability of guilt, and confidence regarding various decisions. Although real jurors and juries do not make such judgments (at least not explicitly), these continuous measures often provide a more statistically powerful way to detect small effects than a binary verdict offered by a group entity. But one danger associated with asking for such judgments is that researchers may inadvertently direct jurors' attention to matters they might otherwise have ignored. By manipulating attention in this way, researchers may be generating results that have limited external validity. We are not aware of data that address the issue, but we merely raise the point that researchers' efforts to obtain more data may itself distort the data they receive.

* * *

The discussion to this point may seem to paint a gloomy picture of goal 2 (jury modeling) research. However, we do not mean to suggest that jury simulations that do not incorporate every feature of real jury trials are worthless for goal 2 purposes. Studies that lack ecological validity may very well still yield results that generalize across a range of actual trials if those

aspects of actual trials that are not modeled in the simulation are not relevant to the question being studied. For example, if the stakes of a trial do not affect the many in which jurors interpret DNA random match probabilities (RMPs), then a failure to model the high-stakes nature of the trial is not a shortcoming to a study that sets out to examine the limited question of whether some other variable – such as the way in which RMPs are presented to jurors – affects the impact of this evidence. The general problem, however, is that it is difficult to know *which* aspects of the trial are relevant in a given context. Research may provide some clues over time, though answers will be slow in coming given the many potential interactions among trial variables and the current lack of theory as to *why* certain trial variables may not influence decision-making or may only influence it in certain contexts (Diamond, 1997, p. 563). Consequently, we think researchers should *err on the side of caution rather than convenience when conducting jury simulations for purposes of modeling jury decision-making*. While recognizing that improving the ecological validity of our studies will take a toll on resources – and may ultimately limit the quantity of studies we conduct – our view is the field would do well to push more in this direction, particularly when we wish to make claims about real jury behavior.

At the very least, we suggest that jury researchers who have goal 2 in mind should generally be extremely cautious when making claims based on data from individual jurors who did not deliberate in groups and who based their judgments entirely on short written materials. There may be cases in which an exception is justified, such as when there are affirmative reasons to believe that visual evidence and group deliberation would have no influence on the results. But we reject the notion that the high cost involved in designing and conducting more ecologically valid studies justifies the status quo approach to goal 2 research. The cost of doing

unpersuasive research is even higher. For those interested in goal 3 research – i.e., persuading policy makers to make reforms – it is even more imperative that to conduct ecologically valid studies. As we discuss below, policy makers want to see research that they can trust and that is broadly applicable.

III. Goal 3: Jury Improvement & Policy Changes

This third goal among jury researchers - improving the jury process - is different from the other two goals in that the intended audience is not other jury researchers. Instead, the focal audience is policy makers – a group including judges, legislators, attorneys general, prosecutors, rules committees, and the like. These are the people who have the power to decide, for example, whether a procedural rule or judicial instruction should be modified. Policy makers may likewise be empowered to permit jurors to ask questions in open court, allow jurors to engage in informal deliberations prior to the end of a case, or recommend that deliberating jurors refrain from conducting straw polls on the ultimate issue until all jurors have expressed their initial views. Though policy decisions should be grounded in methodologically sound empirical facts, few policy makers are trained in such methodological concepts as construct validity, internal validity, or experimental design. Indeed, these decision makers may not have *any* background in scientific and methodological matters. Moreover, it is not clear that even those policy makers who have technical backgrounds would be persuaded to take corrective actions based on the results of the typical jury simulation.

Judiciary's Response to Jury Simulations

Scholars who have looked at the judiciary's response to social science evidence, including evidence from jury simulations, generally report that the courts are not impressed

(Bornstein & McCabe, 2005; Caprathe, 2011; Fradella, 2003; Marder, 2006). One famous case out of the Seventh Circuit illustrates what can happen when a brilliant social science scholar meets a brilliant legal mind. In *Free v. Peters* 12 F.3d 700 (7th Cir. 1993), Judge Richard Posner lambasted a jury simulation conducted by social science luminary Professor Hans Zeisel.

In the case, defendant Free appealed his death penalty conviction on grounds that the instructions the jurors received were confusing and misleading. In support, Free produced a study by Professor Zeisel showing that mock jurors who heard the very same instructions that Free's jurors received misinterpreted much of what they were told. Based on their answers to 18 true-false questions, Zeisel concluded that nearly half of the mock jurors misunderstood key questions, and a significant proportion of these individuals thought that the instructions actually conveyed the very *opposite* points that they were intended to convey.

A District Court concluded that Zeisel's study "should be taken seriously" and ordered a new sentencing hearing for Free (*Free v. Peters*, 778 F. Supp. 431 (N.D. Ill. 1991)). But the 7th Circuit Court of Appeals reversed on grounds that Zeisel's study was "deficient" and could not be "taken seriously in light of the extraordinary vulnerability of his method" (p. 706). Writing for the majority, Judge Posner dismissed Zeisel's study as fatally flawed because there was a "lack of comparability between the test setting and the setting of the sentencing hearing," and because the study did not include "a control group consisting of persons administered a test containing what Zeisel (or Free's lawyers) would consider adequately clear instructions" (p. 705).⁴

⁴ Regarding the second alleged fatal flaw in *Free v. Peters* (1993) (lack of control group), the court argued that Professor Zeisel's simulation is uninformative because it failed to demonstrate that mock jurors who received clear instructions understood those directions. We offer two comments about this argument. First, what the court asks for is not a "control group" but an additional treatment group. Second, and more significantly, the objective of Zeisel's study wasn't to demonstrate that one group of instructions is superior to another. It was merely to demonstrate that the instructions used in Free's case were likely misunderstood. For this purpose, a comparison group is

Regarding the first alleged fatal flaw (lack of comparability in test and real jury settings), the majority reasoned as follows: “There is little a priori reason to think that the results of such an examination offer insight into the ability of a real jury, which has spent days or weeks becoming familiar with the case and has had the benefit of oral presentations by witnesses, lawyers, and the judge, and which renders a verdict after discussion rather than in the isolation of an examination setting” (p. 705-06). In other words, the majority felt that a simulation that did not map onto a target case in terms of length, oral presentation, and deliberation was unlikely to provide “insight” into what a real jury did. A concurring opinion also noted that the lack of voir dire reduced the value of simulation results because voir dire excuses jurors who are “easily confused or easily swayed by non-significant matters” (Bauer, J., concurring, p. 707). As we noted previously, whether variables such as length, oral presentation, jury deliberation, and voir dire actually matter in jury decisions are empirical questions.

Whether the criticism is empirically justified or not, many judges will likely take the position that empirical studies that contain obvious ecological imperfections should not be given much weight. Consider, for example, *State v. Deck*, 994 S.W.2d 527 (Mo. 1999). Like *Free v. Peters*, *State v. Deck* involved a criminal defendant’s introduction of an empirical study calling into question how well jurors understood judicial instructions. Similar to *Free v. Peters*, the *Deck* court the Supreme Court of Missouri rejected the study “because the people interviewed for the study . . . 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” (p. 542). Although not all judges will reach the same conclusions, the point is that even in the absence of data indicating that poor ecological validity results in poor external

unnecessary. Interestingly, subsequent studies actually did show that a revised set of instructions improved comprehension levels (Diamond & Levi, 1996; Wiener, Pritchard, & Weston, 1995).

validity in jury simulations, judges may not be willing to assume that ecologically invalid simulations can be trusted to describe the behavior (or likely behavior) of real jurors.

Greater Attention to Ecological Validity

The larger point of these admittedly selective cases is that if jury simulation research is to have an impact on the judiciary and other legal policy makers, researchers will have to pay close attention to matters of ecological validity. They should do so for at least two reasons. One reason is simply because policy makers care about how well the laboratory setting captures various potentially influential factors in real trials. Studies that use students (as opposed to “real” jurors), short written stimuli (as opposed to detailed trial videos), or lack group deliberation will likely be viewed with suspicion by many policy makers. Even when jury simulation studies have high degrees of internal validity, policy makers may have a hard time getting past their artificiality. A second and related reason jury researchers should place more attention on ecological validity is that doing so will enhance what Dillehay and Nietzel (1980) refer to as the *applied explanatory power* of the effects their studies uncover. Effects that are identified from internally valid studies – including those that appear to generalize across a variety of settings and populations – may not tell us much about whether those effects explain a meaningful proportion of variance in the more complex, applied setting of interest, namely, jury trials. For example, suppose that a series of well-controlled studies finds that the precise language that a forensic science expert uses to describe a match impacts jury verdicts (as compared to the use of slightly different language conveying the same general concept). If the ecological validity of those studies is low, we may not assume that that this language effect captures a meaningful proportion of verdict variance in complex real cases in which the forensic

experts may be examined and cross-examined for hours, and in which jurors hear testimony from other witness, arguments from attorneys, and instructions from a judge.

Some jury researchers will reject this perspective. For example, Kerr and Bray (2005) suggest that social scientists should pay relatively little attention to the wishes of policy makers and matters of ecological validity and instead focus on conducting high quality, internally valid studies. Similarly, Kerr and Bray (2005) argue that rather than simply giving policy makers the kind of studies that they want, jury researchers should try to educate policy makers (e.g., as expert witnesses and through *amicus curiae* briefs) about the value of highly controlled, albeit highly artificial, jury simulations (p. 356-57). Relatedly, one might also argue that social scientists should inform policymakers that reviews such as Bornstein (1999), Desmarais and Read (2011), and Camerer and Hogarth (1999) suggest that intuitions about the importance of running high ecological validity studies, including those that have high stakes, are wrong. In other words, according to this argument, we should try to persuade policy makers that practical implications for real juries and ideas for reform *can* be extracted from highly artificial simulations.

We agree that social scientists should never stop trying to educate policy makers, many of whom are lawyers, about matters of science and inference. For this reason, we support recent efforts to increase social science training in law schools by offering courses in empirical research methods, science and law, jury decision-making, statistics, quantitative reasoning, and the like. At the same time we think it is unrealistic to expect that policy makers will ignore ecological validity shortcomings, nor do we think they should do so until a body of research, specifically designed to address these issues, emerges that shows such concerns to be misguided.

Practical Considerations

Introducing greater realism into jury simulations will often be difficult and costly (Penrod et al., 2011). This fact provides jury simulation researchers with at least some incentive to dig in and challenge the wisdom of what we suggest here. But here, once again, is where the investigator's goals should be consulted. If jury researchers wish to advance basic scientific theory, then traditional, unrealistic, randomized, controlled laboratory studies that place a premium on internal validity are fine. But researchers seeking to have an impact on trial policy cannot assume that the legitimate excuses they provide for not producing realistic simulations will persuade judges, policy makers, and advisory groups (e.g., a diverse National Research Council panel) to give greater weight to their studies. Instead, these excuses may simply cause policy makers to remain skeptical about the utility of experimental results and conclusions.

We also suggest that those who are interested in conducting goal 3 research should take greater care in matters of problem selection. Thompson (1993) has suggested that jury researchers too often address questions that “lawyers consider obvious, trivial, unimportant or uninteresting” (p. 204; see also Vidmar, 1979). Similarly, Lempert (1991) noted that jury researchers often study questions that already been decided by the courts, and therefore have little practical impact. This gulf between the questions that social scientists ask and the questions that policy makers want answered may also widen when social scientists who are really most interested in goal 3 research identify and design studies with goal 1 in mind. We therefore recommend that jury simulation researchers who wish to affect policy try putting themselves in the shoes of policy makers when selecting research questions. What are the problems that the system is currently grappling with, and how might an empirical study point the way toward a cost-effective and reliable solution? Researchers who think along these lines may ultimately find themselves in the best position to impact legal policy.

IV. Conclusion

We have suggested that there are at least three goals that drive jury simulation research: (1) theory development, (2) jury modeling, and (3) jury reform. We have also suggested that researchers should think hard about which goals are of primary interest before locking into a research question and design. Where basic theory (goal 1) matters most, internal validity should be stressed. Where describing the behaviors of real juries (goal 2) or persuading policy makers about changes that should be made (goal 3), external and ecological validity must be stressed as well.

We pushed this idea a bit further by offering the (admittedly untested) empirical claim that most jury researchers are more interested in describing actual jury behavior and impacting legal policy than they are in advancing basic psychological theory. This is not to say that jury researchers are uninterested in identifying general psychological phenomena. But we suspect that the jury setting is of special interest to most of those who conduct jury simulations, as opposed to being just a handy setting in which to test psychological theory.

If our suspicions are correct, then jury researchers should focus more energy on matters related to ecological validity and external validity in their experimental designs. Such focus will likely increase the amount of time and money needed to conduct jury research, and may even lead to fewer published studies (Penrod et al., 2011, p. 196). Kerr and Bray (2005) reject this strategy as too costly (p. 356), and Penrod et al. (2011) point out that the impact of low ecological validity on the external validity of jury simulation studies remains an unanswered empirical question. Still, we throw our lot in with critics like Vidmar (2008), Diamond (1997), Dillehay and Nietzel (1980), Nunez et al. (2011), and Wiener (2010) who have suggested that the

poor ecological validity of jury simulations is a major problem for the field and one that requires corrective action. We also agree with Wells (2005), who suggests that jury researchers should think more about how to communicate effectively with legal policy makers who are unlikely to be satisfied by evidence of internal validity. As Wells (2005), Lempert (1991), and others note, if psychologists want to have a meaningful impact on the legal system, they will need to maintain a sharp focus on the goals and concerns of the legal policy makers they seek to persuade: “[P]olicymakers are not going to find useful a body of research that undermines their current policies and practices, unless there are clear demonstrations of better policies and practices to take their place” (Wells, 2005, p. 497).

This conclusion may not be a popular one among jury researchers. Many will continue to believe that a well-designed study can simultaneously achieve all three goals (i.e., advance basic theory, describe real jury behavior, and reform legal policy). Indeed, it can (for a sample of such studies, see reviews by Bornstein, 1999; Diamond, 1997; and Wiener et al., 2011). At the same time, however, with the rise of Mechanical Turk, Qualtrics, and various related Internet-based survey research platforms, the temptation to conduct short, cheap, highly artificial legal studies is greater than ever. So long as we, as a field, fail to take ecological validity seriously, we must be prepared to accept that our impact on the courts and other policy makers will be small.

References

Associated Press (3/16/2005). Blake found not guilty.

http://www.today.com/id/7211376/ns/today-today_entertainment/t/blake-found-not-guilty/#.U5nkLy-vy3k

Blanck, P. D., Rosenthal, R., & Cordell, L. D. H. (1985). The appearance of justice: Judges’

- verbal and nonverbal behavior in criminal jury trials. Stanford Law Review, 38, 89-164.
- Bornstein, B. H. (1999). The Ecological Validity of Jury Simulations: Is the Jury Still Out? Law and Human Behavior, 23, 75-91.
- Bornstein, B. H., & McCabe, S. G. (2005). Jurors of the Absurd? The Role of Consequentiality in Jury Simulation Research. Florida State University Law Review, 32, 443-467.
- Bray, R. M., & Kerr, N. L. (1979). Use of the Simulation Method in the Study of Jury Behavior: Some Methodological Considerations. Law and Human Behavior, 3, 107-119.
- Brodsky, S. L., (2009). Credibility in the Courtroom: How Likeable Should an Expert Witness Be? Journal of the American Academy of Psychiatry and Law, 37, 525-532.
- Breau, D. L., & Brook, B. (2007). "Mock" mock juries: A field experiment on the ecological validity of jury simulations. Law and Psychology Review, 31, 77-92.
- Camerer, C. F. & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. Journal of Risk and Uncertainty, 19, 7-42.
- Caprathe, W. J. (2011). Commentary: Participant Differences and Validity of Jury Studies. Behavioral Sciences and the Law, 29, 328-330.
- Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design and Analysis Issues for Field Settings. Boston, MA: Houghton Mifflin.
- Cramer, R. J., Brodsky, S. L., & DeCoster, J. (2009). Expert witness confidence and juror personality: Their impact on credibility and persuasion in the courtroom. Journal of the American Academy of Psychiatry and Law, 37, 63-74.

- Davis, J. H. (1973). Group Decision and Social Interaction: A Theory of Social Decision Schemes. Psychological Review, 80, 97-106.
- Desmarais, S. L., & Read, D. (2011). After 30 Years, What Do We Know about What Jurors Know? A Meta-Analytic Review of Lay Knowledge Regarding Eyewitness Factors. Law and Human Behavior, 35, 200–210.
- Devine, D. J. (2012). Jury Decision Making: State of the Science. New York: New York University Press.
- Diamond, S. S. (1997). Illuminations and shadows from jury simulations. Law and Human Behavior, 21, 561–571.
- Diamond, S. S. & Casper, J. D. (1992). Blindfolding the jury to verdict consequences: Damages, experts, and the civil jury. Law and Society Review, 26, 513-563.
- Diamond, S. S., & Zeisel, H. (1974). A courtroom experiment on juror selection and decision-making. Personality and Social Psychology Bulletin, 1, 276-277.
- Diamond, S. S., & Levi, J. N. (1995-96). Improving decisions on death by revising and testing jury instructions. Judicature, 79, 224-232.
- Dillehay, R. C., & Nietzel, M. T. (1980). Constructing a science of jury behavior. In Review of Personality and Social Psychology, Vol.1 (Ed: Ladd Wheeler). Sage: Beverly Hills, CA.
- Estelle v. Williams*, 425 U.S. 501 (1976).
- Fox, P., Winegrove, T., & Pfeifer, C. (2011). A comparison of students' and jury panelists' decision-making in split recovery cases. Behavioral Sciences and the Law, 29, 358-375.
- Galantner, M. (2004). The vanishing trial: An examination of trials and related matters in federal and state courts. Journal of Empirical Legal Studies, 1, 459-470.
- Fradella, H. F. (2003). A content analysis of federal judicial views of the social science

- “researcher’s black arts.” *Rutgers Law Journal*, 35, 103-170.
- Free v. Peters*, 778 F. Supp. 431 (N.D. Ill. 1991).
- Free v. Peters*, 12 F.3d 700 (7th Cir. 1993).
- Hammond, K. R. (1998). Ecological validity: Then and now. Retrieved June 12, 2014.
<http://www.brunswik.org/notes/essay2.html#Now>
- Heath, W. P., Grannemann, B. D., & Peacock, M. A. (2004). How the defendant’s emotion level affects mock jurors’ decisions when presentation mode and evidence strength are varied. *Journal of Applied Social Psychology*, 34, 624-664.
- In re Urethane Antitrust Litigation* MDL No. 1616, Case No. 04-1616-JWL., 08-2617-JWL, 09-2026-JWL (M.D. Kan 2013).
- Kaye, D. H., Hans, V. P., Dann, B. M., Farley, E. J., & Albertson, S. (2007). Statistics in the Jury Box: How Jurors Respond to Mitochondrial DNA Probabilities. *Journal of Empirical Legal Studies*, 4, 797-834.
- Kerr, N. L., & Bray, R. M. (2005). Simulation, realism, and the study of the jury. In N. Brewer & K. D. Williams (Eds.). *Psychology and Law: An Empirical Perspective* (pp. 322-364). New York, NY: The Guilford Press.
- Kerr, N. L., & Tindale, R. S. (2004). Group Performance and Decision Making. *Annual Review of Psychology*, 55, 623-55.
- Koehler, J. J., & Macchi, L. (2004). Thinking about low-probability events: An exemplar cuing theory. *Psychological Science*, 15, 540-546.
- Lempert, R. O. (1991). Why do jury research? In R. Hastie (Ed.). *Inside the Juror* (pp. 242–254). Cambridge: Cambridge Univ. Press.
- London, K. Y., & Nuñez, N. (2000). The effect of jury deliberation on jurors’ propensity to

- disregard inadmissible evidence. Journal of Applied Psychology, 85, 932-939.
- Marder, N. S. (2006). Bringing jury instructions into the twenty-first century. Notre Dame Law Review, 81, 449-511.
- McCoy, M. L., Nuñez, N., & Dammeyer, M. D. (1999). The effect of jury deliberations on jurors' reasoning skills. Law and Human Behavior, 23, 557-575.
- Mize, G. E., Hannaford-Agor, P., & Waters, N. L. (2007). The State-of-the-States Survey of Jury Improvement Efforts: A Compendium Report. Williamsburg, VA: National Center for State Courts. Retrieved from http://www.ncsconline.org/D_Research/cjs/pdf/SOSCompendiumFinal.pdf
- Mook, D. G. (April 1983). In Defense of External Invalidity. American Psychologist, 38, 379-387.
- Narby, D., Cutler, B. L., & Moran, G. (1993). A meta-analysis of the association between authoritarianism and jurors' perceptions of defendant culpability. Journal of Applied Psychology, 78, 34-42.
- Nietzel, M. T., McCarthy, D. M., & Kern, M. J. (1999). Juries: The current state of the empirical literature. In R. Roesch, S. D. Hart, & J. R. P. Ogloff (Eds.). Psychology and Law: The State of the Discipline (pp. 23–42). New York: Kluwer/Plenum.
- Nunez, N., McCrea, S. M., & Culhane, S. M. (2011). Jury decision making research: Are researchers focusing on the mouse and not the elephant in the room? Behavioral Sciences and the Law, 29, 439-451.
- Pennington, N. & Hastie, R. (1981). Juror decision making models: The generalization gap. Psychological Bulletin, 89, 246-287.
- Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision making. Journal

- of Personality and Social Psychology, 51, 651–672.
- Penrod, S. D., Kovera, M. B., & Groscup, J. (2011). Jury Research Methods. In Rosenfeld, B. and Penrod, S. D. (Eds.). Research Methods in Forensic Psychology (pp. 191-214). Hoboken, NJ: John Wiley & Sons, Inc.,
- Pezdek, K. Avila-Mora, E., & Sperry, K. (2010). Does trial presentation medium matter in Jury simulation research? Evaluating the effectiveness of eyewitness expert testimony. Applied Cognitive Psychology, 24, 673-690.
- Robbennolt, J. (2002-2003). Evaluating Empirical Research Methods: Using Empirical Research in Law and Policy, Nebraska Law Review, 81, 777-804.
- Rose, M. R., & Abrahamson, J. B. (2011). Data, Race, and the Courts: Some Lessons on Empiricism from Jury Representation Cases. Michigan State Law Review, 2011, 911-963.
- Rose, M. R., Diamond, S. S., & Baker, K. M. (2010). Goffman on the Jury: Real Jurors' Attention to the "Offstage" of Trials. Law and Human Behavior, 34, 310-323.
- Salerno, J. M., & Bottoms, B. L. (2009). Emotional evidence and jurors' judgments: The promise of neuroscience for informing psychology and law. Behavioral Sciences and the Law: Special Issue: The Neuroscience of Decision Making and Law, 27, 273-296.
- Salerno, J. M., & Diamond, S. S. (2010). The promise of a cognitive perspective on jury deliberation. Psychonomic Bulletin and Review, 17, 174-179.
- Schkade, D., Sunstein, C. R. & Kahneman, D. (2000). Deliberating about dollars: The severity shift. Columbia Law Review, 100, 1139-1176.
- Schwartz, S. L., & Hunt, J. S. (2011). Considering her circumstances: How ethnicity and cultural relativist arguments affect sexual harassment judgments by undergraduate

- and community mock jurors. Behavioral Sciences & the Law, 29, 419-438.
- Stasser, G., & Davis, J. H. (1981). Group Decision Making and Social Influence: A Social Interaction Sequence Model. Psychological Review, 88, 523-551.
- State v. Deck*, 994 S.W.2d 527 (Mo. 1999).
- Stebay, N., Hosch, H. M., Colhane, S. E., & McWethy, A. (2006). The Impact of Juror Verdicts on Judicial Instruction to Disregard Inadmissible Evidence: A Meta-Analysis. Law and Human Behavior, 30, 469-492.
- Thompson, William C. (1993). Research on jury decision-making: The state of the science. In N. J. Castellan (Ed.). Individual and Group Decision Making (pp. 203-218). Hillsdale, NJ: Erlbaum.
- Tindale, R. S. (1993). Decision errors made by individuals and groups. In N. J. Castellan (Ed.). Individual and Group Decision Making (pp. 109-124). Hillsdale, NJ: Erlbaum.
- Vidmar, N. (1979). The Other Issues in Jury Simulation Research: A Commentary with Particular Reference to Defendant Character Studies. Law and Human Behavior, 3, 95-106.
- Vidmar, N. (2008). Civil Juries in Ecological Context: Methodological Implications for Research. In Bornstein, B. H., Wiener, R. L., Schopp, R., Willborn, S. L. (Eds.). Civil Juries and Civil Justice: Psychological & Legal Perspectives (pp. 35-65). New York, NY: Springer.
- Weiten, W., & Diamond, S. S. (1979). A Critical Review of the Jury Simulation Paradigm: The Case of Defendant Characteristics. Law and Human Behavior, 3, 71-93.
- Wells, G. L. (2005). Helping experimental psychology affect legal policy. In N. Brewer & K. D. Williams (Eds.). Psychology and Law: An Empirical Perspective (pp. 483-

- 500). New York, NY: The Guilford Press.
- Wheatman, S. R., & Shaffer, D. R. (2001). On finding for defendants who plead insanity: The crucial impact of dispositional instructions and opportunity to deliberate. Law and Human Behavior, 25, 167-183.
- Wiener, R. L. (2010). What is the Study of Jury Decision Making About and What Should it be About? In Bornstein, B. H., Wiener, R. L., Schopp, R., Willborn, S. L. (Eds.). Civil Juries and Civil Justice: Psychological & Legal Perspectives (pp. 67-76). New York, NY: Springer.
- Wiener, R. L., Krauss, D. A., & Lieberman, J. D. (2011). Mock Jury Research: Where Do We Go from Here? Behavioral Sciences and the Law, 29, 467-479.
- Wiener, R. L., Pritchard, C. C., & Weston, M. (1995). Comprehensibility of approved jury instructions in capital murder cases. Journal of Applied Psychology, 80, 455-467.
- Wissler, R. L., & Saks, M. J. (1985). On the Inefficacy of Limiting Instructions: When Jurors Use Prior Conviction Evidence to Decide on Guilt. Law and Human Behavior, 9, 37-48.