

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Publications, Department of Psychology

Psychology, Department of

December 2004

The Impact of Different Types of Expert Scientific Testimony on Mock Jurors' Liability Verdicts

Brian H. Bornstein

University of Nebraska-Lincoln, bbornstein2@unl.edu

Follow this and additional works at: <https://digitalcommons.unl.edu/psychfacpub>



Part of the [Psychiatry and Psychology Commons](#)

Bornstein, Brian H., "The Impact of Different Types of Expert Scientific Testimony on Mock Jurors' Liability Verdicts" (2004). *Faculty Publications, Department of Psychology*. 161.

<https://digitalcommons.unl.edu/psychfacpub/161>

This Article is brought to you for free and open access by the Psychology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, Department of Psychology by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

The Impact of Different Types of Expert Scientific Testimony on Mock Jurors' Liability Verdicts

Brian H. Bornstein

Department of Psychology, University of Nebraska–Lincoln

Abstract: Participants in two experiments acted as jurors for a personal-injury case containing different types of expert testimony. In both experiments, the defendant was more likely to obtain a verdict in his favor when his expert presented anecdotal case histories than when the expert presented experimental data. Participants' liability judgments were correlated with their perceptions of the experts' credibility (experiments 1 and 2) and were moderated somewhat by their need for cognition and preference for numerical information (experiment 2). The results are discussed in terms of reasoning heuristics such as the base-rate fallacy.

Keywords: Liability, Juror Decision Making, Expert Testimony

INTRODUCTION

Expert testimony has become increasingly common in both civil and criminal trials (Gross, 1991; Landsman, 1995; Faigman *et al.*, 1997; Faigman, 1999). For example, in a review of 529 civil trials, Gross and Syverud (1991) found that 86% utilized expert testimony. In some areas of civil law, such as products liability and medical malpractice, expert testimony is now virtually ubiquitous (Gross, 1991; Gross and Syverud, 1991). Rule 702 of the Federal Rules of Evidence defines expert testimony as the “scientific, technical, or other specialized knowledge” of “a witness qualified as an expert by knowledge, skill, experience, training, or education.” Such testimony is admissible if it is reliable, would help the trier of fact decide the case correctly, and if the subject matter is beyond lay comprehension.

The matter of whether expert testimony is relevant in a particular case is left to judicial discretion. The Supreme Court's recent decisions regarding the admissibility standard for scientific (*Daubert v. Merrell Dow Pharmaceuticals*, 1993; *General Electric Co. v. Joiner*, 1997) and technical evidence (*Kumho Tire Co. v. Carmichael*, 1999) eliminated general acceptance in the field as the sole requirement (Faigman, 1995; Berger, 2000; Gatowski *et al.*, 2001). Instead, the courts themselves must act as “gatekeepers” in determining whether the evidence was derived by scientifically valid research methods. Consequently, the courts are scrutinizing scientific evidence more closely than in the past, with the result that the standard is looser for some evidence yet stricter for other evidence (Faigman, 1995, 1999; Gatowski *et al.*, 2001).

A key issue in debating the appropriateness of any expert testimony is determining the effect that it is likely to have. Most psychological research on expert testimony has, not surprisingly, focused on the employment of psychological expertise (*e.g.* Faust and Ziskin, 1988; Saks, 1992; Faigman *et al.*, 1997). To a lesser extent, researchers have also investigated the effects on jurors of non-psychological expert testimony (McAuliff *et al.*, 2003). In civil cases, medical experts providing clinical opinion testimony on individual cases appear to be the prototype (Landsman, 1995). In Gross's (1991; Gross and Syverud, 1991) review of California civil cases, half of the expert witnesses were medical doctors, 9% were other medical professionals (*e.g.* clinical psychologists, dentists), 20% were engineers and other scientists, 11% were financial or business experts, and 8% were experts in reconstruction and investigation. Thus, in at least one state, the emphasis on psychological expert witnesses that exists in the psychological literature does not accurately reflect the proportion of mental health experts among the expert population as a whole.

In criminal cases, one of the more common types of expert testimony is testimony about probabilistic incidence rates (Fienberg, 1989). As forensic science and other technological disciplines become ever more sophisticated, such evidence (*e.g.* DNA matching, statistical evidence of discrimination, epidemiological data in toxic tort cases) has been, and is likely to continue to be, on the rise at trial (Fienberg, 1989; Faigman, 2000). Loftus (1980) showed that mock jurors underweigh such abstract, scientific testimony, especially compared to highly salient, individuating information such as an eyewitness identification. This finding is consistent with several studies that have found that mock jurors are reluctant to base verdicts on statistical evidence alone (*e.g.* Niedermeier *et al.*, 1999), underutilize expert probabilistic testimony compared to Bayesian norms (Faigman and Baglioni, 1988; Kaye and Koehler, 1991; Koehler, 1992), and are influenced by how that testimony is presented (Thompson and Schumann, 1987; Cooper *et al.*, 1996; Koehler, 2001). More generally, it reflects people's tendency to attend preferentially to anecdotal information and to underweigh summary base rates (Kahneman and Tversky, 1973; Borgida and Brekke, 1981; Stanovich and West, 1998).

Base rates can be defined as the "relative frequency with which an event occurs or an attribute is present in a population" (Koehler, 1996, p. 16). For example, evidence that 10% of the population shares a criminal defendant's blood type, which was found at the scene of a crime, would constitute a base rate relevant to the fact-in-issue of whether the defendant was present at the scene. Although people tend to be insensitive to base rates, they are incorporated into predictive judgments to a greater extent if their relevance to the task at hand is made more salient (Borgida and Brekke, 1981; Kruglanski *et al.*, 1984; Koehler, 1996), as by highlighting the causal relationship between the base rate and the predicted criterion (Ajzen, 1977; Tversky and Kahneman, 1980).

This general pattern appears to hold true for the effectiveness of different types of expert testimony in the trial context. For example, Schuller (1992) has distinguished between "general expert" testimony, which summarizes research findings that are only indirectly relevant to the particular case at hand, and "specific expert" testimony, which explicitly links the background research to the case under current consideration. This distinction is not the same as the base-rate/anecdotal distinction, but they are analogous in contrasting general background with case-specific evidence. The general testimony can be thought of as providing base-rate information, while specific testimony provides more individuating information. Consistent with research on the base-rate fallacy, expert psychological testimony is more influential when it provides a concrete example of how the

relevant body of research is linked to the case at hand than when it merely summarizes abstract research findings (Schuller, 1992; Gabora *et al.*, 1993; Kovera *et al.*, 1994, 1997; Krauss and Sales, 2001).

Several studies have found that mock jurors are sensitive to variations in the content of scientific evidence (Brekke *et al.*, 1991; Koehler, 2001, experiment 3; Smith *et al.*, 1996). For example, Smith *et al.* presented mock jurors with a videotape of a simulated rape trial, in which an expert forensic serologist testified regarding crucial blood-and enzyme-type evidence. Participants who were told that 20% of the population had an enzyme type that was shared by the assailant and the defendant judged the defendant as more guilty than participants who learned that 80% of the population had the same enzyme type. On the other hand, Kovera *et al.* (1999) found that mock jurors were not sensitive to more subtle manipulations of expert testimony content. In a simulated hostile work environment case, they varied characteristics of the experimental methodology used in research described by the expert witness (*i.e.* the research's general acceptance, ecological validity, and construct validity). Although these cues had some effect on mock jurors' perceptions of the expert's credibility, they exerted no effect on their verdicts.

Thus, the findings of previous research on jurors' use of expert scientific testimony are somewhat mixed. On the one hand, they tend to underweigh such testimony compared to a normative, Bayesian standard (Thompson and Schumann, 1987; Faigman and Baglioni, 1988; Kaye and Koehler, 1991; Smith *et al.*, 1996), to be unaffected by research methodology (Kovera *et al.*, 1999), and to be influenced by how the evidence is presented (Thompson and Schumann, 1987; Cooper *et al.*, 1996; Koehler, 2001); yet, on the other hand, they are properly sensitive to variations in expert testimony's probative value (Saks and Wissler, 1984; Brekke *et al.*, 1991; Smith *et al.*, 1996).

One possible reason for these disparate findings is variability in how individual jurors will respond to complex scientific testimony. Individual jurors naturally vary in their personal biases and the sophistication of their judgmental processes, and a variety of attitudes can influence jurors' decisions (Kaplan, 1982; Graziano *et al.*, 1990). In complex cases where scientific evidence is prominent, jurors' attitudes and beliefs toward scientific research and complex information (*e.g.* technical, statistical, medical, or other forms of expert testimony) are likely to be especially relevant. The goal of the present research was to explore the impact on mock jurors of different types of expert testimony, particularly for a case in which the expert testimony addressed the sole fact-in-issue. The research also assessed the relationship between mock jurors' attitudes toward scientific and complex information and their evaluation of expert scientific testimony.

EXPERIMENT 1

Prior research has shown that psychological experts are more influential when testifying about concrete cases than about general, impersonal research findings (Schuller, 1992; Gabora *et al.*, 1993; Kovera *et al.*, 1994, 1997; Krauss and Sales, 2001), but a similar comparison has not been made with non-psychological expert testimony. The main comparison in the present experiment is between concrete and abstract testimony provided by a non-psychological expert, testifying for the defense, in a personal-injury case revolving around the connection between ingestion of a potentially harmful substance and deleterious health effects. Specifically, the expert presents testimony that either focuses on individual cases ("anecdotal" evidence) or summarizes abstract research findings ("experimental" evi-

dence). Because the experimental evidence is akin to base rates, it is predicted that the anecdotal evidence would have a greater impact on mock jurors' judgments.

Experiment 1 also addressed the predicted insensitivity to base rates by including two additional variables that manipulated the content of the plaintiff's expert's testimony, provided by a statistician. People's statistical reasoning is subject to a number of shortcomings (*e.g.* Nisbett *et al.*, 1983). In particular, they are typically not sensitive to sample size and fail to apply the law of large numbers (Tversky and Kahneman, 1974). If jurors tend to be relatively uninfluenced by general (*i.e.* non-case-specific) research findings, then they should not be affected by characteristics of the sample on which that research is based.

Sample characteristics were manipulated in two ways: first, by varying the size of the sample used in the epidemiological research described by the plaintiff's statistical expert; and second, by varying the frequency of occurrence, or incidence, of the target characteristic under study. This type of evidence has become increasingly common, especially in products liability, toxic tort, and DNA matching cases (Cecil *et al.*, 1991; Gross, 1991; Green, 1992; Koehler, 2001). It is predicted that despite the relevance of sampling characteristics to the conclusions that can legitimately be drawn from scientific research, mock jurors' tendency to underutilize base rates (*e.g.* Koehler, 1992) and to be insensitive to variations in scientific evidence quality (Kovera *et al.*, 1999) would lead them to be insensitive to variations in characteristics of the base-rate sample. Finally, mock jurors answered questions designed to measure their attitudes toward scientific research in general.

Method

Participants

Participants were 152 undergraduates (38 males and 114 females) enrolled in psychology courses who received extra course credit.

Materials and design

The trial scenario was a one-page, single-spaced summary (540–575 words, depending on condition) of a lawsuit in which the plaintiff alleged that lead contained in refuse stored by a paint manufacturer had seeped into his well water and caused his 12-year-old son to develop rheumatoid arthritis. The case was modified from an actual lawsuit (*Backes v. The Valspar Corporation*, 1986) and contained a description of the plaintiff and defendant, the plaintiff's claim (*i.e.* that drinking contaminated water caused the boy's illness), and the known facts. For example, there was no dispute that "the lead content of the family's well water was found to be higher than the level fit for human consumption", or that the lead came from the company's waste products. The summary described the boy as having had several unsuccessful operations to relieve the arthritis in his hands and feet, great difficulty walking and writing, and unremitting pain. The family sought compensation for the boy's pain and suffering, medical costs, and lost future income.

In such cases, the major, and often sole, issue at trial concerns causality (Green, 1992; Hart and Honore, 1985); hence the sole fact-in-issue in the present trial was whether the lead caused the boy's arthritis. The summary informed participants that lead contributes to other health problems, such as learning disabilities, but that its role in causing arthritis had not yet been determined by the scientific community. Expert testimony offered by both the

TABLE 1 Examples of expert testimony. Italicized figures were altered depending on condition. Alternative versions of the expert testimony are described in the text.

Epidemiological Evidence (for Plaintiff; small sample/low incidence condition). [The witness] compared people in two cities from the same state, each with a population of about 50,000 people: one whose water supply had an unusually high level of lead, and one whose water supply did not. Approximately 40 children in the city with lead-contaminated water had arthritis, while only 13 in the city with lead-free water did. He concludes that exposure to lead poses a substantial health risk, particularly for the developing of arthritis in children.

Anecdotal Evidence (for Defendant). [The witness] presents the case histories of three boys who are the same age as [the plaintiff] and have lived in [the plaintiff's] neighborhood for as long as he has. None of them has ever had any signs of arthritis, or any other abnormal health problems. Each of the three boys appears in court to testify that he has the same water supply and drinks a normal amount of water. The witness has examined the boys and testifies that none of them has any symptoms of arthritis. He concludes that arthritis is not caused by environmental factors like lead poisoning.

Experimental Evidence (for Defendant). [The witness] compared two groups of young rats, 100 rats per group: one that was given water containing more lead than the EPA's acceptable level, and one that was given water with a normal amount of lead. The number of rats in the two groups that developed arthritis was exactly the same, five rats in both groups. Based on his findings, the witness concludes that arthritis is not caused by environmental factors like lead poisoning.

plaintiff and the defendant addressed this key issue by focusing on whether lead in general is capable of causing arthritis.

The nature of the expert testimony was manipulated by varying three factors, two concerning the content of the plaintiff's expert's testimony, and one concerning the type of defendant's expert (see Table 1). The plaintiff's expert witness was a statistician who presented a comparison of the incidence of arthritis in two cities, one where the water supply was contaminated with lead, and one where it was not. One variable was the size of the sample on which the statistician's findings were based: in the *small sample* condition, each city contained 50,000 people, while in the *large sample* condition, each city contained 500,000 people ("small" and "large" here are meant purely in relative terms). A second variable within the statistician's testimony was the number of people who had the disease. In the *high incidence* condition, 40 people per 10,000 in the city exposed to lead-contaminated water had arthritis (*i.e.* 200/50,000 or 2,000/500,000); in the *low incidence* condition, eight people per 10,000 in the city exposed to lead-contaminated water had arthritis (*i.e.* 40/50,000 or 400/500,000). Thus, the incidence of arthritis was five times greater in the high incidence than in the low incidence condition. The ratio of arthritis cases in the exposed sample, compared to the non-exposed sample, was held constant across conditions at approximately 3:1. This ratio was the basis of the plaintiff's expert's conclusion (in all conditions) that exposure to lead increased the risk of arthritis.

The third variable was the type of expert witness called by the defense. The defendant's expert witness was either a scientist presenting experimental findings from laboratory research on animals (*experimental* condition), or a doctor presenting anecdotal evidence from three case histories (*anecdotal* condition). The defendant's expert always concluded that lead does not increase the risk of developing arthritis (see Table 1).

The design was therefore a 2 (Size of plaintiff's expert's sample: large vs small) \times 2 (Incidence of arthritis in plaintiff's expert's sample: high vs low) \times 2 (Type of defense expert: experimental vs anecdotal) between-groups design, with 19 participants per condition.

Procedure

Participants were assigned randomly to one of the eight conditions. The written instructions informed them that the experiment's purpose was to address how people make decisions about legal liability and defined important terminology (e.g. the preponderance of the evidence standard, compensatory damages, etc.). They were also told "to determine the facts solely from the evidence presented in the case. This evidence will consist of testimony by experts."

After reading the trial summary, participants gave a verdict; an estimate of the probability that the defendant caused the plaintiff's injury (on a 0–100 scale); and a compensation award, if they first found the defendant liable. To provide a rough proxy of each expert's credibility, participants also responded to three questions assessing their perception of each expert witness's testimony (e.g. "How important was the expert's evidence in your decision?"), using seven-point Likert scales, and rated the difficulty of each expert's testimony. Finally, participants used seven-point scales to respond to three questions designed to measure their general attitude toward scientific research: (1) "What is your overall opinion (extremely negative–extremely positive) about scientists and the scientific process?" (2) "How comfortable are you (extremely uncomfortable–extremely comfortable) with the scientific process and scientific research results?" (3) "How comfortable are you with math and numbers?"

Participants completed the questionnaire individually, with up to 10 participants in a single session. The entire procedure took 20–30 minutes to complete.

Results

Liability Judgments

The two measures of liability were participants' dichotomous verdicts and their estimates of the probability that the defendant caused the plaintiff's injury (see Table 2). Participants' verdicts were regressed onto the variables of Plaintiff's Sample Size, Incidence in Plaintiff's Sample, and Defense Expert Type in a logistic regression model; causality ratings were used as the dependent measure in a three-way ANOVA.

Participants were more likely to find the defendant liable when his expert presented experimental evidence (59% liable verdicts) than when he presented anecdotal evidence

TABLE 2 Percentage of liable verdicts and mean causality ratings, experiment 1.

	Defense expert type					
	Anecdotal			Experimental		
	% Liable	Mean causality	SD	% Liable	Mean causality	SD
Plaintiff's expert's small sample						
Low incidence	31.6	41.3	16.8	42.1	43.5	24.6
High incidence	21.1	33.1	21.8	68.4	53.9	22.0
Plaintiff's expert's large sample						
Low incidence	47.4	46.9	24.1	47.4	50.7	26.3
High incidence	26.3	38.1	24.8	78.9	63.2	22.9

n = 152 (19 per cell). Causality ratings reflect participants' subjective estimate that the defendant caused the plaintiff's injury (0–100). Incidence refers to the relative number of people in the plaintiff's expert's sample who displayed the target characteristic (i.e. children with arthritis).

(32% liable verdicts), $B = 0.61$, $SE = 0.18$, $p < 0.001$. This effect was qualified by a significant Defense Expert Type \times Incidence interaction, $B = 0.50$, $SE = 0.18$, $p < 0.005$, indicating that the effect of Expert Type was greater when the incidence in the plaintiff's sample was high (74% liable verdicts in the experimental condition vs 24% in the anecdotal condition) than when it was low (45% and 39%, respectively).

Analyses of the causality ratings yielded similar results. Participants rated the defendant as more likely to have caused the plaintiff's injury when the defendant's expert witness presented experimental evidence (mean (M) = 52.83, $SD = 24.6$) than anecdotal evidence ($M = 39.84$, $SD = 22.2$), $F(1, 144) = 12.04$, $p < 0.001$, $\eta^2 = 0.077$; in other words, the defendant's expert had a greater impact when presenting anecdotal evidence. Mirroring the verdict data, this effect was especially pronounced when the incidence of arthritis in the plaintiff's expert's sample was high, as indicated by a significant interaction between Defense Expert Type and Incidence, $F(1, 144) = 7.12$, $p < 0.01$, $\eta^2 = 0.047$.

Compensation Awards

Compensation was analyzed only for participants who first found the defendant liable ($n = 69$), resulting in cell sizes that ranged from 4 to 13 across the eight conditions. None of the variables had a significant effect on how much compensation was awarded to the plaintiff, $F_s < 1.7$ (the result was the same whether analyzing raw or logarithmically transformed values).

Perception of Expert Witnesses

Responses to the three credibility questions were highly intercorrelated, so they were averaged to create a single credibility index (cf. Brekke *et al.*, 1991), which ranged from 1 to 7 ($\alpha = 0.71$ for the plaintiff's expert and 0.70 for the defendant's expert). Credibility ratings for the plaintiff's and the defendant's experts were not significantly correlated, $r = -0.13$, $p > 0.1$; consequently, participants' perceptions of the testimony offered by each expert were analyzed separately, as a function of Plaintiff's Sample Size, Incidence in Plaintiff's Sample, and Defense Expert Type.

When the defendant's expert presented anecdotal evidence, he was perceived as more credible ($M = 5.20$, $SD = 1.11$), $F(1, 144) = 5.27$, $p < 0.05$, $\eta^2 = 0.035$, and marginally less difficult to understand ($M = 1.88$, $SD = 1.03$), $F(1, 144) = 3.21$, $p = 0.075$, $\eta^2 = 0.022$, than when the expert presented experimental evidence (credibility: $M = 4.79$, $SD = 1.13$; difficulty: $M = 2.21$, $SD = 1.26$). There was also a significant interaction between Plaintiff's Sample Size and Incidence in the Plaintiff's Sample on ratings of the defendant's expert's credibility, $F(1, 144) = 5.04$, $p < 0.05$, $\eta^2 = 0.034$. Post hoc comparisons showed that when the plaintiff's expert's evidence dealt with a relatively large sample, the defendant's expert was perceived as more credible for low incidence ($M = 5.32$, $SD = 0.90$) than for high incidence ($M = 4.79$, $SD = 1.13$) testimony, $t(74) = 2.82$, $p < 0.05$, $d = 0.66$. There was no effect of incidence in the small sample condition, $t(74) < 1$. In other words, when there were relatively few incidents in the plaintiff's expert's research of similar injuries, the defendant's expert was perceived as more credible than when there were relatively many incidents, but only if the plaintiff's expert's research involved a large sample.

Compared to experimental evidence, anecdotal evidence presented by the defense made the plaintiff's expert's testimony seem correspondingly less credible ($M_s = 4.53$ vs 4.94), $F(1, 144) = 4.98$, $p < 0.05$, $\eta^2 = 0.033$. This main effect was qualified by a significant interaction between Defense Expert Type and Incidence in the Plaintiff's Expert's Sample,

$F(1, 144) = 5.41, p < 0.05, \eta^2 = 0.036$. As with the liability judgments, there was an effect of Expert Type when Incidence was high, $t(74) = 3.77, p < 0.001, d = 0.88$, such that the plaintiff's expert seemed more credible when the defense expert presented experimental ($M = 5.22, SD = 0.93$) than anecdotal evidence ($M = 4.38, SD = 1.01$). There was no effect of Defense Expert Type when Incidence was low, $t(74) < 1$. Finally, the plaintiff's expert's testimony was perceived as significantly easier to understand when it dealt with a large sample ($M = 1.87, SD = 1.04$) than when the sample size was small ($M = 2.24, SD = 1.18$), $F(1, 144) = 4.13, p < 0.05, \eta^2 = 0.028$.

The relationship between participants' liability judgments and their perceptions of the experts was assessed in two ways: by comparing credibility ratings for participants who did or did not find the defendant liable, and by computing correlations between participants' credibility and causality ratings. Participants who found the defendant liable perceived the plaintiff's expert as more credible, $t(150) = 11.43, p < 0.001, d = 1.87$, and the defendant's expert as correspondingly less credible, $t(150) = 5.73, p < 0.001, d = 0.94$, than participants who did not find the defendant liable. The correlational analyses yielded the same pattern of findings: there was a positive correlation between participants' judgments of the probability that the defendant caused the plaintiff's injury and their perception of the plaintiff's expert's credibility, $r = 0.63, p < 0.001$, while there was a negative correlation between participants' causality ratings and the perceived credibility of the defendant's expert, $r = -0.39, p < 0.001$.

Scientific Attitudes

Responses to the three questions that assessed participants' general attitude toward the scientific process were significantly intercorrelated, so they were summed to form a single Scientific Attitude index (with higher scores indicating a more positive attitude toward scientific research in general). Scores on this index had a mean of 15.81 ($SD = 2.72; \alpha = 0.61$).

In order to assess the relationship between participants' attitudes toward science and their liability judgments, participants' Scientific Attitude scores were compared for those who found the defendant liable and those who did not. Participants' Scientific Attitude scores did not differ as a function of their verdict, $t(149) = 1.37, p = 0.1$, nor did they correlate significantly with participants' estimates that the defendant caused the plaintiff's injury, $r = -0.11$. Participants with a more positive scientific attitude perceived the defendant's expert as more credible, $r = 0.26, p < 0.001$, but Scientific Attitude was not significantly correlated with either perceptions of the plaintiff's expert's credibility, $r = 0.04$, or the perceived difficulty of either expert's testimony, $-0.1 < rs < 0$.

Discussion

The two main hypotheses for experiment 1 were supported: first, the type of scientific expert testimony affected participants' liability judgments, with judgments more favorable to the defendant when his expert's testimony was anecdotal than when it was experimental. This advantage for anecdotal over experimental evidence is consistent with research showing that psychological expert testimony is more influential when it deals with salient, individual cases than with abstract research findings (Schuller, 1992; Gabora *et al.*, 1993; Kovera *et al.*, 1994, 1997; Krauss and Sales, 2001). It is also consistent with more general decision-making research, indicating that people tend to underutilize base rates relative to more salient, individuating information (*e.g.* Kahneman and Tversky, 1973).

Second, variations in the statistical expert testimony failed to affect participants' liability judgments. Specifically, their verdicts and causality judgments were largely uninfluenced by either the size of the expert's sample or the incidence in the sample of the target characteristic (*i.e.* childhood arthritis). This insensitivity to sample characteristics in expert testimony is consistent with research showing that people are insensitive to sampling variables in nonlegal contexts (Tversky and Kahneman, 1974). Mock jurors may be sensitive to overt variations in the probative value, or conclusiveness, of statistical testimony (Smith *et al.*, 1996; Koehler, 2001), but they appear to be less responsive to more subtle manipulations that hold the testimony's stated conclusiveness constant while varying the quality of the evidence on which that testimony is based (Kovera *et al.*, 1999).

Although the same observation of a low-frequency event is statistically more reliable as sample size and event incidence increase, participants in the present study were not responsive to such variations. In order to keep the testimony's conclusiveness roughly equal across conditions, the manipulations of the statistician's evidence were relatively subtle: a 10-fold difference in sample size and a five-fold difference in incidence rate, while the ratio of arthritis in the exposed to the unexposed groups was held constant at 3:1. It is likely that participants would have been more responsive to variations in the magnitude of the effect leading to the expert's conclusion (*e.g.* Koehler, 2001, experiment 3).

Finally, mock jurors' attitudes toward science in general were correlated with their perceptions of one, but not both, experts, and they were not predictive of their verdicts. This finding suggests that overall attitudes toward the scientific process might be too abstract a construct to be of much use in predicting mock jurors' verdicts in a specific case involving scientific evidence.

EXPERIMENT 2

The major finding of experiment 1 was that mock jurors were influenced more by the defendant's expert when he presented anecdotal evidence than when he presented experimental evidence. However, the type of expert testimony for the defense was confounded with the expert witness's credentials. The trial summary described the expert witness who presented anecdotal evidence as "a doctor", while it described the expert who presented experimental evidence simply as "a scientist." As expert witnesses' credibility varies depending on their discipline (Saks and Wissler, 1984) as well as their credentials (*e.g.* Bank and Poythress, 1982; Cooper *et al.*, 1996), it is possible that anecdotal evidence exerted a greater effect than experimental evidence because of who the expert was rather than what the expert said. The primary aim of experiment 2 was to compare anecdotal and experimental evidence while controlling for the expert's identity and credentials. In addition, although experiment 1 compared anecdotal and experimental expert testimony to each other, it did not assess the overall effectiveness of either testimony type. Thus, experiment 2 also includes a control condition with no expert testimony on behalf of the defense.

In light of the failure of participants' general attitudes toward science to predict their verdicts, a secondary aim of experiment 2 was to explore further potential individual difference variables that might predict mock jurors' receptivity to scientific testimony. Rather than relying on intuitive, face valid measures like those used in experiment 1, two measures with known psychometric properties were chosen: Need for Cognition (NFC; Cioppo *et al.*, 1996) and Preference for Numerical Information (PNI; Viswanathan, 1993).

The NFC scale measures individual differences in “people’s tendency to engage in and enjoy effortful cognitive activity” (Cacioppo *et al.*, 1996, p. 197) and is positively correlated with open-mindedness and general intelligence. Four previous studies have addressed the relationship between NFC and potential jurors’ decisions. Graziano *et al.* (1990) found that NFC was related to how mock jurors reasoned their way to a verdict (*e.g.* number and type of reasons listed) in a case concerning unlawful transportation of a stolen car across state lines, although it did not predict their ultimate verdicts. Lassiter *et al.* (1992), while not conducting a simulated trial, found nonetheless that participants high in NFC judged a criminal suspect who made a questionable confession as having a greater likelihood of guilt than did participants low in NFC. DeWitt *et al.* (1997) also found an association between NFC and mock jurors’ verdicts, with high-NFC mock jurors being more defense-oriented in a civil “brainwashing” case. Thus, there is some indication that NFC can predict jurors’ response to trial evidence. The relationship between NFC and jurors’ judgments should be even greater in cases where the evidence is high in cognitive complexity, such as those relying substantially on expert testimony (McAuliff *et al.*, 2003). Indeed, McAuliff and Kovera (2001) found that mock jurors high and low in NFC differed in their sensitivity to variations in some aspects of expert testimony, although they were equally insensitive to other evidentiary manipulations (see also DeWitt *et al.*, 1997).

Viswanathan (1993), p. 749) defines PNI as “a preference or proclivity toward using numerical information and engaging in thinking involving numerical information.” Not surprisingly, PNI and NFC are positively correlated ($r = 0.30$; Viswanathan, 1993; study 7); however, PNI is a more specific construct, in that it taps individuals’ tendency to enjoy thinking about one type of information (*i.e.* numbers) rather than thinking in general (Viswanathan, 1993). Because scientific evidence, especially when it involves experimental or epidemiological findings, is heavily numerically based, this measure was included in the present study as well. It was predicted that participants high in NFC or PNI would be more responsive to the experimental evidence, and less responsive to the less probative anecdotal evidence, than those low in NFC or PNI.

Method

Participants

Participants were 141 undergraduates (80.6% female) enrolled in psychology courses who received extra course credit. Only one participant had previously served on a jury, while nine participants had taken part in a trial in some other capacity (*e.g.* litigant or witness).

Materials and Design

The same basic trial scenario was used as in experiment 1. Expert testimony was again presented on behalf of the plaintiff by a statistician, who described the results of an epidemiological comparison (the *small sample-low incidence* condition from experiment 1). The expert testimony presented on behalf of the defendant varied two factors: the presence/absence of anecdotal evidence, and the presence/absence of experimental evidence. When the expert presented anecdotal evidence, he testified that he had examined three boys who were exposed to the same water supply as the plaintiff, none of whom had developed arthritis. When the expert presented experimental evidence, he described the results of an animal study. In all conditions where an expert witness testified for the defense, he was de-

scribed as an MD, from the same university as the plaintiff's expert, whose research led him to conclude that lead poisoning did not cause arthritis. The overall design was therefore a 2×2 between-groups design that varied the content of the defense expert's testimony based on whether each of two types of evidence (anecdotal and experimental) was present. Participants were randomly assigned to one of the four conditions, with 33–37 participants per cell.

The procedure and main dependent variables were the same as those used in experiment 1. The only change was that instead of asking participants their attitudes toward science, they completed a "Personality Questionnaire" containing the short form of the Need for Cognition Scale (18 items; Cacioppo *et al.*, 1984) and the PNI (20 items), with items from the two scales mixed together randomly. Participants rated their agreement/disagreement with each item on a nine-point scale.

Results

Liability Judgments

The two measures of liability were participants' dichotomous verdicts and their estimates of the probability that the defendant caused the plaintiff's injury (see Table 3). For the verdict data, participants' verdicts were regressed onto the variables of Anecdotal Evidence (present/absent) and Experimental Evidence (present/absent). Participants were significantly less likely to find the defendant liable when he presented anecdotal evidence (47.1% liable verdicts) than when he did not (67.1%, $B = 0.83$, $p < 0.05$). Whether or not the defendant presented experimental evidence exerted no effect on participants' verdicts, $B = 0.04$, $p > 0.5$.

Participants' estimates of the probability that the defendant caused the plaintiff's injury were similar to their verdicts: they gave the defendant somewhat lower causality ratings when he presented anecdotal evidence ($M = 48.3$, $SD = 25.1$) than when he did not ($M = 55.6$, $SD = 25.4$), $F(1, 137) = 2.91$, $p = 0.09$, $\eta^2 = 0.021$. Neither the effect of Experimental Evidence nor the interaction was significant, $F_s < 1$.

Compensation Awards

Compensation was analyzed only for participants who first found the defendant liable ($n = 81$), resulting in cell sizes that ranged from 15 to 25 across the four conditions. Neither variable exerted an effect on compensation awards, $F_s < 1$ (the result was the same whether analyzing raw or logarithmically transformed values).

TABLE 3 Percentage of liable verdicts and mean causality ratings, experiment 2.

<i>Experimental evidence</i>	<i>Anecdotal evidence</i>					
	<i>Present</i>			<i>Absent</i>		
	<i>% Liable</i>	<i>Mean causality</i>	<i>SD</i>	<i>% Liable</i>	<i>Mean causality</i>	<i>SD</i>
Present	48.6	47.2	24.9	66.7	55.6	26.4
Absent	45.5	49.4	25.6	67.6	55.6	24.8
Total	47.1	48.3	25.1	67.1	55.6	25.4

$n = 141$ (33–37 per cell). Causality ratings reflect participants' subjective estimate that the defendant caused the plaintiff's injury (0–100).

Perceptions of Expert Testimony

Participants' perceptions of the testimony offered by the two experts were analyzed separately, for both their overall credibility (again an index variable averaging responses to three questions) and perceived difficulty of the testimony. Neither variable affected perceptions of the plaintiff's expert's testimony, $F_s(1, 137) < 3.58, p_s > 0.05$.

As participants did not evaluate the defendant's expert's testimony in the condition where no expert testified for the defense, perceptions of the defendant's expert were analyzed by a one-way ANOVA with three conditions (anecdotal evidence only, experimental evidence only, and anecdotal and experimental evidence together). Participants' perceptions of the defense expert's overall credibility varied as a function of testimony type, $F(2, 101) = 3.33, p < 0.05, \eta^2 = 0.062$. Tukey post hoc comparisons indicated that the defendant's expert was perceived as more credible when he presented both types of testimony ($M = 5.19, SD = 1.02$) than when he presented only experimental evidence ($M = 4.55, SD = 0.97$), $p < 0.05$. The mean rating when the defendant's expert presented only anecdotal evidence was intermediate ($M = 4.98, SD = 1.16$) and did not differ from the other two conditions. The perceived difficulty of the defense expert's testimony did not vary depending on the type of testimony presented, $F(2, 101) = 0.48$.

The relationship between participants' perceptions of the expert witnesses and their liability judgments was assessed by comparing the credibility ratings of participants who did or did not find the defendant liable, and by correlating participants' credibility ratings with their causality judgments. As in experiment 1, participants who found the defendant liable perceived the plaintiff's expert as more credible than participants who found for the defendant ($M_s = 5.68$ vs 4.38), $t(139) = 8.70, p < 0.001, d = 1.48$. Conversely, participants who found the defendant liable perceived the defendant's expert as significantly less credible than participants who found in favor of the defendant ($M_s = 4.59$ vs 5.28), $t(102) = 3.41, p < 0.001, d = 0.68$. Similarly, participants' estimates of the probability that the defendant caused the plaintiff's injury were positively correlated with ratings of the plaintiff's expert's credibility, $r = 0.57, p < 0.001$, but somewhat negatively correlated with ratings of the defendant's expert's credibility, $r = -0.14, p > 0.1$.

Personality Variables

NFC and PNI scores did not differ for participants who did and did not find the defendant liable, $t_s(139) < 1$. Correlations were also computed between NFC, PNI, and participants' causality ratings and perceptions of the expert witnesses. NFC and PNI were positively correlated with each other, $r = 0.40, p < 0.001$. However, the two personality measures were largely uncorrelated with participants' judgments and perceptions. NFC was negatively correlated with the plaintiff's expert's perceived credibility, $r = -0.19, p < 0.05$, but otherwise NFC and PNI were not significantly correlated with participants' perceptions or their causality ratings, $-0.10 < r_s < 0.06, p_s > 0.2$.

To assess whether the personality variables interacted with the experimental manipulations, participants were categorized as high or low on NFC and PNI (by median split), and the principal analyses were rerun including these factors. NFC interacted with Anecdotal Evidence on participants' causality judgments, $F(1, 130) = 4.62, p < 0.05$. Participants high in NFC were uninfluenced by whether the defense expert did ($M = 54.54, SD = 21.96$) or did not ($M = 52.65, SD = 27.02$) present anecdotal evidence; whereas participants low in NFC judged the defendant as less likely to have caused the plaintiff's injury when his

expert did present anecdotal evidence ($M = 44.24$, $SD = 26.37$) than when he did not ($M = 60.47$, $SD = 22.51$). There were no other significant effects involving NFC, aside from an uninterpretable three-way interaction on causality ratings and perception of the plaintiff's expert, $F_s(1, 130) > 5.4$, $p_s < 0.05$.

The only significant effect involving PNI was an interaction with Experimental Evidence on participants' causality ratings, $F(1, 131) = 5.66$, $p < 0.05$. Participants high in PNI judged the defendant as more likely to have caused the plaintiff's injury when his expert did not present experimental evidence ($M_s = 56.16$ absent vs 43.59 present), whereas participants low in PNI judged the defendant as more likely to have caused the plaintiff's injury when his expert did present experimental evidence ($M_s = 57.97$ present vs 49.30 absent).

Discussion

The major finding of experiment 2 was that there was a sizable impact on mock jurors' liability judgments of anecdotal expert testimony. In addition, the effectiveness of anecdotal evidence was the same whether or not it was accompanied by testimony about experimental research on animals. The effect of experimental findings alone was no different from a control condition lacking any expert testimony for the defendant. These findings suggest that, as with expert psychological testimony (Schuller, 1992; Gabora *et al.*, 1993; Kovera *et al.*, 1994), non-psychological experts are also more influential when presenting concrete cases than when summarizing general research findings.

As in experiment 1, there was a strong relationship between the experts' perceived credibility and participants' liability judgments. Across all types of experts, participants who reached a verdict for the plaintiff perceived the plaintiff's expert as more credible, and the defendant's expert as correspondingly less credible, than participants who found for the defendant. Thus, their verdicts were consistent with their evaluation of the evidence formally presented at trial. There was also some indication that the amount of evidence presented by the defense expert affected his credibility, in that participants perceived him to be most credible when he presented both anecdotal and experimental evidence. However, he was no more credible in this condition than when presenting anecdotal evidence alone.

As in previous research (*e.g.* McAuliff and Kovera, 2001), there was little direct relationship between personality variables relating to individuals' predilection for complex information (NFC and PNI) and their judgments. The inability of the present measures to predict mock jurors' judgments is consistent with other research showing that despite a couple of notable exceptions (*e.g.* authoritarianism, death penalty attitudes), there is little correlation between personality attributes and jurors' verdicts (*e.g.* Kaplan, 1982; Fule-ro and Penrod, 1990). However, there was some evidence that the personality variables employed in the present study were important moderators. Specifically, low-NFC (but not high-NFC) participants' causality ratings were influenced by the presence or absence of anecdotal evidence. Similarly, the presence of experimental evidence had a positive influence on high-PNI participants (*i.e.* it moved them in the direction of the side presenting that evidence), but a negative influence on low-PNI participants (*i.e.* it moved them away from the side presenting that evidence). Thus, consistent with the findings of McAuliff and Kovera (2001), mock jurors' attitudes toward complex information have an effect on how they process various aspects of expert testimony.

GENERAL DISCUSSION

The present findings support previous research (*e.g.* Thompson and Schumann, 1987; Smith *et al.*, 1996) showing that mock jurors respond to variations in the expert testimony presented in a simulated trial. Expert psychological testimony about specific instances has a greater effect on mock jurors' judgments than testimony about general research findings (Schuller, 1992; Gabora *et al.*, 1993; Kovera *et al.*, 1997; Krauss and Sales, 2001), and this effect generalizes to non-psychological expert testimony as well. Different types of expert testimony can have a profound effect on jurors' judgments, especially when the expert testimony is central to the case (*cf.* Brekke *et al.*, 1991). However, the variable effects of different types of expert testimony are not necessarily normative. Participants in the present experiments were non-normative in giving greater weight to anecdotal evidence based on a very small sample than to experimental findings based on much larger samples in both experiments.

Given the rarity of the plaintiff's illness in the present simulation, it should hardly be surprising—or very informative—that the defendant in such a case could find individuals similar to the plaintiff who did not develop adverse health effects (it would only be surprising if he could *not* find such individuals). An argument can therefore be made that the defense expert's anecdotal testimony was virtually worthless with respect to its probative value, yet it nevertheless exerted a greater effect than the scientifically more probative experimental evidence. Despite the adoption by many jurisdictions of the somewhat more stringent *Daubert* standard, clinical opinion expert testimony, like the anecdotal evidence presented here, is often held to a lower standard (Krauss and Sales, 2001), and examples of "junk science" in the courtroom abound (Huber, 1991). Thus, although one might hope that probatively worthless anecdotal expert testimony would not be admissible in the first place, such a hope is unlikely to be realized any time soon (Faigman, 2000).

Participants also perceived experts presenting anecdotal evidence as more credible than experts presenting non-anecdotal evidence. Of course, the content of the message is only one factor in determining the effectiveness of expert testimony (McAuliff *et al.*, 2003); characteristics of the source play a part as well (Bank and Poythress, 1982) and become particularly influential as the complexity of the expert testimony increases (Cooper *et al.*, 1996). Variations in experts' status may explain why different types of experts vary in terms of their perceived honesty and competence (Saks and Wissler, 1984). Although the type of expert (*i.e.* the expert's credentials) was confounded with the nature of the expert's testimony in experiment 1, experiment 2 controlled for the expert's identity and still found a stronger effect of anecdotal compared to experimental evidence.

There are several possible explanations for why jurors tend to weigh anecdotal expert testimony more heavily than general research findings. First, they may not understand the more abstract research. The present experiments did not assess participants' comprehension directly. However, there was only a marginally significant tendency in one of two experiments for participants to perceive anecdotal evidence as easier to understand than experimental evidence. Although mock jurors' perception of evidence difficulty does not necessarily reflect how well they actually understood it, this finding suggests that they do not invariably experience general research findings as hard to comprehend.

A second reason for jurors' underutilization of general research findings may be that people harbor a mistrust of scientific research (Saks, 1992). The increasing use of expert testimony and changing standards for its admissibility have been accompanied by a grow-

ing awareness that scientific methods are not infallible (Faigman, 1995, 1999). This skepticism might lead jurors to give scientific testimony less weight than it deserves, especially when that testimony is relatively abstract. However, participants' judgments were largely uncorrelated with their attitudes toward the scientific process, need for cognition, or preference for numerical information, suggesting that this explanation is also not wholly satisfactory. Instead, consistent with previous research (e.g. Graziano *et al.*, 1990; DeWitt *et al.*, 1997; McAuliff and Kovera, 2001), participants' attitudes appeared to be associated, to some extent, with how they interpreted different kinds of expert testimony in evaluating the main fact-in-issue, whether the defendant caused the plaintiff's injury. This raises the possibility that *evidence-specific* attitudes, as well as *case-specific* attitudes (Moran *et al.*, 1994), are capable of influencing jurors' verdicts.

It is likely that inasmuch as the participants were college students, their scores on the attitudinal measures were higher, and had a more restricted range, than would be true of the population as a whole. Although jury simulation research has detected little difference between the performance of student mock jurors and more diverse samples (Bornstein, 1999), differences confounded with student status—such as NFC and PNI—seem especially relevant to how jurors would process complex information in cases turning on expert testimony. Thus, research employing more ecologically valid samples is called for before ruling out the possibility that attitudes toward science and complex information can predict jurors' verdicts in complex cases.

A third interpretation of the differential effectiveness of anecdotal versus general expert testimony is to view it as a demonstration within the trial context of the base-rate fallacy, according to which people are influenced more by salient, individual cases than by base rates drawn from larger samples (Kahneman and Tversky, 1973; Koehler, 1996). People tend not to rely sufficiently on base rates because of the difficulty in generalizing from abstract background information to specific, concrete cases (Borgida and Brekke, 1981). Conversely, anecdotal evidence, by nature, tends to be more elaborate, vivid, and salient than base-rate information. Vivid information presented at trial may receive more weight than relatively pallid information for a number of reasons, such as greater attention or memorability (Bell and Loftus, 1985). These effects are magnified under conditions of high overall information complexity (Wilson *et al.*, 1989), which characterizes most trials, especially those involving expert testimony. This interpretation is consistent with research indicating that expert psychological testimony that specifically links research findings to the case at hand is more effective than testimony that merely summarizes the relevant research (Schuller, 1992; Gabora *et al.*, 1993; Kovera *et al.*, 1997). General expert testimony, such as base-rate evidence, fails to prove or disprove the "causal nexus" with regard to the individual plaintiff (Cecil *et al.*, 1991).

One final explanation of the present findings is that the anecdotal evidence might have been perceived as more relevant because of essential differences between the subjects of the anecdotal and experimental testimony. For example, the anecdotal testimony in the present trial simulation described boys of the same age as the plaintiff, whose health status would likely be perceived as more relevant to the case at hand than the animals described in the experimental testimony, by virtue of their greater similarity to the plaintiff (*i.e.* demonstrating the "representativeness" bias; Tversky and Kahneman, 1974). The anecdotal evidence might also have been perceived as more relevant because of the subjects' greater physical proximity to the plaintiff—that is, boys from the same neighborhood—than animals in a laboratory environment many miles away.

Such confounds are an inherent limitation in studies that strive to be ecologically valid, and future research should, to the extent possible, manipulate these variables orthogonally in order to tease apart their effects. However, it is in the nature of anecdotal and experimental evidence that they will inevitably differ in these and other respects, owing to the fact that anecdotal evidence based on non-human animals would likely fail to exceed the relevance threshold, while ethical considerations preclude rigorous experimental work exposing humans to high levels of lead or other potentially toxic substances. The courts are often reluctant to admit expert testimony based solely on animal studies. For example, in dismissing the claims of Agent Orange plaintiffs, a federal court held that “animal studies are not helpful in the instant case because they involve different biological species. They are of so little probative value and are so potentially misleading as to be inadmissible” (*In re “Agent Orange”*, 1985; see Green, 1992). Nonetheless, the scientific and medical communities view animal studies as extremely relevant to health issues in humans, and they are a frequent element of scientific testimony in various types of trials (Faigman *et al.*, 1997; Faigman, 1999). Future research should investigate ways of making such testimony, as well as other forms of abstract (*e.g.* epidemiological) expert testimony, more palatable to jurors.

References

- Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology*, 35, 303–314.
- Backes v. The Valspar Corporation, 783 F.2d 77 (7th Cir. 1986).
- Bank, S. C. and Poythress, N. G. (1982). The elements of persuasion in expert testimony. *Journal of Psychiatry and Law*, 10, 173–204.
- Bell, B. E. and Loftus, E. F. (1985). Vivid persuasion in the courtroom. *Journal of Personality Assessment*, 49, 659–664.
- Berger, M. A. (2000). The Supreme Court’s trilogy on the admissibility of expert testimony. In F. M. Smith (Ed.), *Reference Manual on Scientific Evidence*, 2nd edn (pp. 9–38). Washington, DC: Federal Judicial Center.
- Borgida, E. and Brekke, N. (1981). The base rate fallacy in attribution and prediction. In J. H. Harvey, W. Ickes and R. F. Kidd (Eds.), *New Directions in Attribution Research*, vol. 3 (pp. 63–95). Hillsdale, NJ: Erlbaum.
- Bornstein, B. H. (1999). The ecological validity of jury simulations: is the jury still out? *Law and Human Behavior*, 23, 75–91.
- Brekke, N. J., Enko, P. J., Clavet, G. and Seelau, E. (1991). Of juries and court-appointed experts: the impact of nonadversarial versus adversarial expert testimony. *Law and Human Behavior*, 15, 451–475.
- Cacioppo, J. T., Petty, R. E. and Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48, 306–307.
- Cacioppo, J. T., Petty, R. E., Feinstein, J. A. and Jarvis, W. (1996). Dispositional differences in cognitive motivation: the life and times of individuals varying in need for cognition. *Psychological Bulletin*, 119, 197–253.
- Cecil, J. S., Hans, V. P. and Wiggins, E. C. (1991). Citizen comprehension of difficult issues: lessons from civil jury trials. *American University Law Review*, 40, 727–774.
- Cooper, J., Bennett, E. A. and Sukel, H. L. (1996). Complex scientific testimony: how do jurors make decisions? *Law and Human Behavior*, 20, 379–394.
- Daubert v. Merrell Dow Pharmaceuticals, 509 US 579, 113 S. Ct. 2786 (1993).
- DeWitt, J. S., Richardson, J. T. and Warner, L. G. (1997). Novel scientific evidence and controversial cases: a social psychological examination. *Law and Psychology Review*, 21, 1–27.

- Faigman, D. L. (1995). The evidentiary status of social science under Daubert: is it 'scientific,' 'technical,' or 'other' knowledge? *Psychology, Public Policy and Law*, 1, 960–979.
- Faigman, D. L. (1999). *Legal Alchemy: the Use and Misuse of Science in the Law*. New York: Freeman.
- Faigman, D. L. (2000). The law's scientific revolution: reflections and ruminations on the law's use of experts in year 7 of the revolution. *Washington and Lee Law Review*, 57, 661–684.
- Faigman, D. L. and Baglioni, A. J. (1988). Bayes' Theorem in the trial process: instructing jurors on the value of statistical evidence. *Law and Human Behavior*, 12, 1–17.
- Faigman, D. L., Kaye, D. H., Saks, M. J. and Sanders, J. (1997). *Modern Scientific Evidence: the Law and Science of Expert Testimony*, vol. 1. St Paul, MN: West.
- Faust, D. and Ziskin, J. (1988). The expert witness in psychology and psychiatry. *Science*, 241, 31–35.
- Fienberg, S. E. (1989). *The Evolving Role of Statistical Assessments As Evidence in the Courts*. New York: Springer-Verlag.
- Fulero, S. M. and Penrod, S. D. (1990). Attorney jury selection folklore: what do they think and how can psychologists help? *Forensic Reports*, 3, 233–259.
- Gabora, N. J., Spanos, N. P. and Joab, A. (1993). The effects of complainant age and expert psychological testimony in a simulated child sexual abuse trial. *Law and Human Behavior*, 17, 103–119.
- Gatowski, S. I., Dobbin, S. A., Richardson, J. T., Ginsburg, G. P., Merlino, M. L. and Dahir, V. (2001). Asking the gatekeepers: a national survey of judges on judging expert evidence in a post-Daubert world. *Law and Human Behavior*, 25, 433–458.
- General Electric Co. v. Joiner, 522 US 136, 118 S. Ct. 512 (1997).
- Graziano, S. J., Panter, A. T. and Tanaka, J. S. (1990). Individual differences in information processing strategies and their role in juror decision making and selection. *Forensic Reports*, 3, 279–301.
- Green, M. D. (1992). Expert witnesses and sufficiency of evidence in toxic substances litigation: the legacy of Agent Orange and Bendectin litigation. *Northwestern University Law Review*, 86, 643–699.
- Gross, S. R. (1991). Expert evidence. *Wisconsin Law Review*, 1991, 1113–1232.
- Gross, S. R. and Syverud, K. D. (1991). Getting to no: a study of settlement negotiations and the selection of cases for trial. *Michigan Law Review*, 90, 319–393.
- Hart, H. L. A. and Honore, T. (1985). *Causation in the Law*, 2nd edn. Oxford: Clarendon Press.
- Huber, P. W. (1991). *Galileo's Revenge: Junk Science in the Courtroom*. New York: Basic Books.
- In re "Agent Orange" Product Liability Litigation, 611 F. Supp. 1223, 1241 (EDNY 1985).
- Kahneman, D. and Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237–251.
- Kaplan, M. F. (1982). Cognitive processes in the individual juror. In N. L. Kerr and R. M. Bray (Eds.), *The Psychology of the Courtroom* (pp. 197–219). New York: Academic Press.
- Kaye, D. H. and Koehler, J. J. (1991). Can jurors understand probabilistic evidence? *Journal of the Royal Statistical Society A*, 154, 75–81.
- Koehler, J. J. (1992). Probabilities in the courtroom: an evaluation of the objections and policies. In D. K. Kagehiro and W. S. Laufer (Eds.), *Handbook of Psychology and Law* (pp. 167–84). New York: Springer-Verlag.
- Koehler, J. J. (1996). The base rate fallacy reconsidered: descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences*, 1–53.
- Koehler, J. J. (2001). When are people persuaded by DNA match statistics? *Law and Human Behavior*, 25, 493–513.
- Kovera, M. B., Levy, R. J., Borgida, E. and Penrod, S. D. (1994). Expert testimony in child sexual abuse cases: effects of expert evidence type and cross-examination. *Law and Human Behavior*, 18, 653–674.
- Kovera, M. B., Gresham, A. W., Borgida, E., Gray, E. and Regan, P. C. (1997). Does expert psychological testimony inform or influence juror decision making? A social cognitive analysis. *Journal of Applied Psychology*, 82, 178–191.
- Kovera, M. B., McAuliff, B. D. and Hebert, K. S. (1999). Reasoning about scientific evidence: effects of juror gender and evidence quality on juror decisions in a hostile work environment case. *Journal of Applied Psychology*, 84, 362–375.

- Krauss, D. A. and Sales, B. D. (2001). The effects of clinical and scientific expert testimony on juror decision making in capital sentencing. *Psychology, Public Policy and Law*, 7, 267–310.
- Kruglanski, A. W., Friedland, N. and Farkash, E. (1984). Lay persons' sensitivity to statistical information: the case of high perceived applicability. *Journal of Personality and Social Psychology*, 46, 503–518.
- Kumho Tire Co. v. Carmichael, 526 US 137, 119 S. Ct. 1167 (1999).
- Landsman, S. (1995). Of witches, madmen, and products liability: an historical survey of the use of expert testimony. *Behavioral Sciences and the Law*, 13, 131–157.
- Lassiter, G. D., Slaw, R. D., Briggs, M. A. and Scanlan, C. R. (1992). The potential for bias in videotaped confessions. *Journal of Applied Social Psychology*, 22, 1838–1851.
- Loftus, E. F. (1980). Psychological aspects of courtroom testimony. *Annals of the New York Academy of Sciences*, 347, 27–37.
- McAuliff, B. D. and Kovera, M. B. (2001). Need for cognition and juror sensitivity to methodological flaws in psychological science. Unpublished manuscript.
- McAuliff, B. D., Nemeth, R. J., Bornstein, B. H. and Penrod, S. D. (2003). Juror decision making in the 21st century: confronting science and technology in court. In D. Carson and R. Bull (Eds.), *Handbook of Psychology in Legal Contexts*, 2nd edn (pp. 303–327). Chichester: Wiley.
- Moran, G., Cutler, B. L. and De Lisa, A. (1994). Attitudes toward tort reform, scientific jury selection, and juror bias: verdict inclination in criminal and civil trials. *Law & Psychology Review*, 18, 309–328.
- Niedermeier, K. E., Kerr, N. L. and Messé, L. A. (1999). Jurors' use of naked statistical evidence: exploring bases and implications of the Wells effect. *Journal of Personality and Social Psychology*, 76, 533–542.
- Nisbett, R. E., Krantz, D. H., Jepson, C. and Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review*, 90, 339–363.
- Saks, M. J. (1992). Normative and empirical issues about the role of expert witnesses. In D. K. Kagehiro and W. S. Laufer (Eds.), *Handbook of Psychology and Law* (pp. 185–203). New York: Springer-Verlag.
- Saks, M. J. and Wissler, R. L. (1984). Legal and psychological bases of expert testimony: surveys of the law and of jurors. *Behavioral Sciences & the Law*, 2, 435–449.
- Schuller, R. A. (1992). The impact of battered woman syndrome evidence on jury decision processes. *Law and Human Behavior*, 16, 597–620.
- Smith, B. C., Penrod, S. D., Otto, A. L. and Park, R. C. (1996). Jurors' use of probabilistic evidence. *Law and Human Behavior*, 20, 49–82.
- Stanovich, K. E. and West, R. F. (1998). Who uses base rates and P(D/H)? An analysis of individual differences. *Memory & Cognition*, 26, 161–179.
- Thompson, W. C. and Schumann, E. L. (1987). Interpretation of statistical evidence in criminal trials: the prosecutor's fallacy and the defense attorney's fallacy. *Law and Human Behavior*, 11, 167–187.
- Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science*, 185, 1124–1131.
- Tversky, A. and Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in Social Psychology*, vol. 1 (pp. 49–72). Hillsdale, NJ: Erlbaum.
- Viswanathan, M. (1993). Measurement of individual differences in preference for numerical information. *Journal of Applied Psychology*, 78, 741–752.
- Wilson, M. G., Northcraft, G. B. and Neale, M. A. (1989). Information competition and vividness effects in on-line judgments. *Organizational Behavior and Human Decision Processes*, 44, 132–139.