JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, March 2011, Vol. 56, No. 2
doi: 10.1111/j.1556-4029.2010.01671.x
Available online at: onlinelibrary.wiley.com

PAPER

JURISPRUDENCE

Lisa L. Smith,[1] *M.Sc.; Ray Bull,*[1] *D.Sc.; and Robyn Holliday,*[1] *Ph.D.*

Understanding Juror Perceptions of Forensic Evidence: Investigating the Impact of Case Context on Perceptions of Forensic Evidence Strength

ABSTRACT: The most widely accepted model of juror decision making acknowledges the importance of both the case-specific information presented in the courtroom, as well as the prior general knowledge and beliefs held by each juror. The studies presented in this paper investigated whether mock jurors could differentiate between evidence of varying strengths in the absence of case information and then followed on to determine the influence that case context (and therefore the story model) has on judgments made about the strength of forensic DNA evidence. The results illustrated that mock jurors correctly identified various strengths of evidence when it was not presented with case information; however, the perceived strength of evidence was significantly inflated when presented in the context of a criminal case, particularly when the evidence was of a weak or ambiguous standard. These findings are discussed in relation to the story model, and the potential implications for real juries.

KEYWORDS: forensic science, forensic psychology, jury decision making, evidence, probative value

In recent decades, courtrooms have seen an astonishing increase in the application of scientific techniques to the law. The development of fingerprint techniques for identifying individuals in 1901 was integral in the development of identification techniques and was closely followed in 1910 by Edmund Locard's principle of trace evidence that revolutionized the way crime detection was approached by both the police and the scientific community (1). Current forensic science techniques are capable of recovering and analyzing a wide range of materials (e.g., glass, fibers, paint, gunshot residue) that can be used to establish a connection between a source and a criminal act or crime scene.

As well as increasing the range of types of material that can be recovered and analyzed for forensic purposes, technological advances have also decreased the quantity of trace material required to conduct useful comparative analyses. Material samples smaller than one nanogram (one millionth of a gram) can be detected and analyzed using current forensic science technology (1). Although these extremely sensitive detection methods can be a great benefit to the investigation of crime, the interpretation of such evidence requires increased consideration of alternative explanations for the presence of tiny traces of material (2).

The final contextual interpretation of forensic science evidence presented in the courtroom is the job of the jury; therefore, improving our understanding of how jurors perceive and evaluate forensic evidence is particularly important in light of technological advances in evidence analysis techniques as these have introduced more ambiguity into the interpretation of probative value (3). The complicated information presented to jurors in the courtroom is inevitably integrated with the individual juror's prior beliefs and knowledge about forensic science and the Criminal Justice System in general, which may come from a variety of sources, and be of varying degrees of accuracy (4).

The most widely accepted model of juror decision making, the story model, proposes that jurors organize information presented during a trial into narrative representations which they then use to evaluate the evidence and eventually reach a verdict (5). The construction of a story representation of the events presented at trial relies on both the evidence concerning the event and the people involved, as well as general knowledge about similar situations held by the juror. The use of general knowledge, in the form of episode schemas, in constructing a narrative can result in different jurors constructing different stories (and perhaps reaching different verdicts) even though they have heard the same items of evidence (4,5). The reliance on episode schemas can also result in inferences being made about information which is not stated in the evidence or evidence which appears to be missing altogether (6).

The story model does acknowledge the importance of episode schemas in allowing a juror to construct a narrative of the events leading up to the crime in question (4). The use of schemas and also heuristics in decision making is not limited to that of jurors making judgments of guilt or innocence, as people use these cognitive tools to assist many complex decision tasks (7). Schoemaker (8) describes the story model process as a means "to connect the new stimuli to the mental models in our heads" (p. 278). This incorporation of new stimuli (which is evidence in the case of trial information) into the existing mental models can help jurors make sense of a large amount of information that is often quite ambiguous. This reliance on mental models can, however, also result in sub-optimal reasoning and decision making. Jurors can become over-reliant on these schemas or frames, and this can result in an

[1]University of Leicester, Leicester, U.K.

Received 14 Jan. 2010; and in revised form 3 Mar. 2010; accepted 13 Mar. 2010.

unwillingness to change the story in light of evidence that does not fit the preferred schema (8).

The story model has been used to test and explain jury decision making in the research literature and has been quite successful at offering a descriptive account of the cognitive process of reaching a verdict in criminal trials (5,6). However, this research literature has not been applied more recently with regards to using the story model as a theoretical basis for understanding how jurors evaluate more complex scientific evidence types, such as DNA evidence. Instead, the focus of recent research has been the use of probability theory models to determine juror comprehension of statistical testimony that often accompanies these types of evidence.

Probability theory models are considered by some researchers to be especially useful for evaluating jury decision making, particularly in cases where jurors are presented with statistical or probabilistic evidence in the form of expert witness testimony (9–11). Typically, the evidence is associative in nature (also referred to as identification evidence), which refers to its ability to be matched to a comparison sample originating from the defendant or another person relevant to the case (9). A jury is therefore presented with the physical evidence and also hears testimony from an expert that provides the court with relevant probabilities that are meant to assist the jurors with evaluating the probative value of the evidence (12,13). Although some researchers suggest that this approach is potentially well suited for a range of evidence types, its reliance on quantifiable and valid probabilities and base rates limits its use in the courtroom. However, courtrooms have adopted this method to use with DNA evidence presented at trial, which is typically accompanied by the probability of the sample matching the defendant as a result of coincidence (referred to as the random match probability—RMP) (14).

One approach in the research literature has used Bayes' theorem as a normative calculation to be compared to mock jurors' decision making concerning the probability of guilt of a suspect. The goal of these studies has been to investigate the relative weights assigned by jurors to various pieces of evidence, and the subsequent impact that this has on the jurors' final perception of suspect guilt. The most common finding is that jurors tend to under-value the probabilistic evidence when compared to a Bayesian calculation (9–11,15), although this is not a consistent finding across all studies (16).

This previous research focus on the interpretation of probabilistic information presented in the courtroom provides some explanation for one potential source of error for jurors evaluating forensic evidence. Another potential source of error that is not addressed in this previous research is the extent to which jurors understand the probative value of various types of physical evidence. This type of error in judgment may be of particular concern in light of recent technological advances, and the resulting sensitivity of evidence extraction and analyses in practice today.

The probative value of forensic evidence is determined by numerous characteristics of the evidence within the context of a particular crime scene. Location, orientation, uniqueness, and mobility are examples of factors that can contribute to the potential probative value of evidence (17,18). Advances in forensic evidence recovery and analysis techniques have further highlighted the importance of mobility in particular, as innocent contamination of evidence becomes an increasing concern as the size of samples required for analysis decreases (19,20).

The two studies reported in this paper used systematic variations in two characteristics of evidence, mobility, and relevance, to quantify the probative value of various types of forensic evidence. By defining probative value in terms of these two factors, it is possible to determine whether these aspects of evidence are considered by potential jurors when judgments of probative value are made. In study one, potential jurors evaluated the strength of forensic evidence in the absence of case context to determine whether mobility and relevance play a role in the perception of probative value. Study two aimed to determine the impact of case context on the evaluation of evidence strength, and therefore how the proposed integration of case information into a narrative (as described by the story model) affects the perception of evidential strength.

Study One

Method

Participants—The sample consisted of jury-eligible members of the general public (N = 163) and included 118 female and 45 male respondents, aged 18–65 years (M = 32.5 years, SD = 11.8 years). The basic criteria for jury eligibility were met if participants indicated they were at least 18 years old and had no previous criminal convictions. Seventy-seven participants (47%) indicated that secondary school was the highest level of education they had achieved, 37 participants (23%) had completed an undergraduate degree, 49 (30%) held postgraduate qualifications.

The recruitment of participants was achieved mainly via a press release issued by the researchers' institution (U.K. university). Participants were able to access and participate in the study by visiting a website address provided in the press release. A snowball technique was also employed by encouraging participants to forward the study hyperlink to other eligible participants after they had completed the task (21).

Materials—The online questionnaire used for this study was developed using the Bristol Survey Tool that enables an online questionnaire to be built and offers a number of options in question and answer format and also collects and stores the data electronically which can then be exported into Excel or SPSS for analysis. Once the questionnaire had been designed, a web address was assigned for the project to which participants could then be directed to complete the questionnaire.

A series of statements describing various types and sources of forensic evidence were presented to participants (see Appendix A for the list of evidence items). A five-point Likert scale was used to measure the participants' opinions in response to the question "How useful would this evidence be in determining the guilt of a suspect?" The responses on the scale ranged from "irrelevant/useless evidence" to "conclusive/definite evidence" and were coded such that a high numeric score indicated evidence that was perceived to be strong.

The evidential items were not only varied in terms of the type of forensic evidence (fingerprints, DNA, footwear, etc.) but also in terms of the investigative relevance of the evidence and the mobility of the sample. The mobility of a sample is an important factor to consider when evaluating evidence because the more mobile a sample is, the less evidential value it has (17). This is because of the fact that mobile samples can be more easily innocently explained as it is plausible that they could have been deposited at a crime scene without requiring the associated person to be involved in the offense (e.g., a discarded cigarette end found at a scene which could have been innocently transported there from somewhere else) (20).

The investigative relevance of a sample is determined by a number of factors, which are largely dependent on the nature of the sample. Factors that were considered to determine investigative

relevance of a sample included the location of the sample (e.g., on a suspect or victim's body, or in a room, etc.) and the uniqueness of the sample (20). In other words, investigative relevance was conceptualized as the extent to which guilt could be directly inferred from the evidence. The result of these manipulations is summarized in Fig. 1.

As illustrated in Fig. 1, investigative relevance is considered the more salient factor in the determination of overall probative value, as the strongest categories of evidence must have high levels of investigative relevance. This is because of the fact that investigative relevance is a necessary condition for evidence to meet in order for it to be considered admissible in a court of law. Typically, evidence that is of low investigative relevance (regardless of its level of mobility) is not considered to be admissible. However, in the case of mobility, both levels (high and low) may be considered admissible if they are of investigative relevance.

Procedure—Once the questionnaire was developed and adapted into an online format, it was piloted on a group of students participating in a course on advanced online research methods ($N = 28$). The feedback received from the pilot participants was incorporated into the final version where appropriate, for example some of the questions contained forensic-related language that was not well understood by participants, and in other items, descriptions of evidence were considered to be ambiguous and were made clearer in the final version of the questionnaire.

The online questionnaire web address remained active for a period of 3 months at the end of which the responses were exported to SPSS for coding and analysis.

Results

There are a number of occasions in the analyses of the data collected in which a repeated-measures analysis of variance (ANOVA) was used to determine within-subjects differences on various measures. In all of these instances, the Mauchly test of sphericity was used, and where this assumption had been violated by these data, the degrees of freedom used in the ANOVA were corrected using the Greenhouse–Geisser estimate of sphericity (22).

The evidence presented to participants was manipulated on two factors, each of which had two levels: mobility of the evidence sample (low or high) and investigative relevance of the evidence (low or high). This manipulation resulted in a 2×2 factorial design with varying relative strengths as highlighted in Fig. 1.

A factorial repeated-measures ANOVA was used to investigate whether the independent variables "mobility" and "relevance" affected the strength ratings for the evidence. Descriptive statistics of the mean strength ratings for each category of evidence are summarized in Table 1. The mean strength scores for each category follow the trend that would be expected as a result of the manipulations; therefore, it seems that participants were considering both mobility and relevance as important when rating strength of evidence.

The results of the factorial repeated-measures ANOVA indicate that there is a significant main effect of the degree of mobility of the evidence, $F_{1,162} = 135.69$, $p < 0.05$, partial $\eta^2 = 0.46$ as well as of the level of investigative relevance, $F_{1,162} = 476.97$, $p < 0.05$, partial $\eta^2 = 0.75$. No significant interaction between mobility and relevance was found, $F_{1,162} = 1.25$, $p = 0.27$. This was interpreted to mean that regardless of the level of mobility, high relevance evidence was rated stronger than low relevance evidence; and regardless of the degree of relevance, low mobility evidence was rated stronger than high mobility evidence.

	Low Mobility	High Mobility
Low Relevance	Moderate/weak	Weakest evidence
High Relevance	Strongest evidence	Moderate/strong

FIG. 1—*Relative strengths of evidence as a function of mobility and investigative relevance.*

TABLE 1—*Mean strength ratings by mobility and relevance categories.*

Category of Evidence	Mean Strength Rating	SD
High mobility/low relevance (weakest)	7.49	1.69
Low mobility/low relevance (moderate/weak)	8.55	1.90
High mobility/high relevance (moderate/strong)	9.93	1.74
Low mobility/high relevance (strongest)	11.18	1.51

All mean scores differ significantly at $p < 0.05$.

Study Two

Method

Participants—This study had the same target population as study one and advertised for participation in a similar way. In total, 160 participants completed the online task, consisting of 113 female and 47 male respondents between the ages of 18 and 67 years ($M = 34.8$ years, SD = 12.1 years). Thirty participants (19%) indicated that their highest level of education was secondary school, 105 (66%) had completed an undergraduate degree, and 25 (15%) completed postgraduate study.

Materials—To allow random allocation of participants to one of the four conditions in this study, a single main webpage was assigned to the project, which contained a welcome message and brief overview of the research objectives. Participants were given this web address for participation in the study, and by visiting this site, they were then randomly forwarded to one of four web pages, each containing one of the experimental conditions. Each of these four sites included a further introductory message and more detailed description of the research and instructions for completing the task, as well as the informed consent details.

Control (No-Context) Condition—Participants allocated to the control (or no-context) condition responded to an evidence evaluation task containing different materials and formats than the other three conditions. This task was identical in design to the evidence evaluation task described in study one.

Case Context Conditions—To present participants with contextual information about a criminal case, and the associated evidence, a case summary and excerpts from court testimony were developed with reference to a fictional murder case. To control for a variety of case variables, such as source of forensic evidence and seriousness of offense, one case scenario was used with only the strength of the forensic DNA evidence being varied across the three conditions. The presentation of the DNA evidence in the transcript followed the format used in previous research in jury decision making (16). The case information included testimony from an eyewitness (victim's neighbor), the arresting police officer, a forensic pathologist,

a forensic laboratory analyst as well as a summary of the prosecution and defense arguments in the case.

Initially, participants were presented with a brief summary of the case which did not include any detail about evidence recovered from the scene or witnesses who testified in the trial. After reading this summary, participants indicated how probable they thought it was that the suspect was guilty of the murder as described in the summary (expressed as a percentage between 0 and 100%). This measure provided a baseline perception of guilt for each participant against which comparisons can be made with later measures in the study.

Following this summary, participants read through the testimony of key witnesses in the murder trial. The material presented in the witness statements was identical in each of the three conditions, with the exception of the scenes of crime examination information which varied in terms of the mobility and relevance of the DNA evidence recovered from the crime scene. Following each witness statement, participants were asked to indicate how useful they felt the evidence was for determining the guilt of the suspect, using the same five-point Likert scale described in study one.

After all of the witness statements were read, participants were once again asked to estimate the perceived probability that the suspect was guilty of the murder. The final piece of trial information presented at this stage was a summary of the case presented by the prosecution and then by the defense lawyers. These summaries were presented as brief narratives which accounted for the evidence presented, but either supported the guilt (prosecution summary) or innocence (defense summary) of the accused. Following these summaries, participants were asked to estimate the probability of guilt one final time, as well as make a verdict decision (guilty or not-guilty of murder).

Results

Table 2 summarizes the mean strength ratings for the evidence evaluated in the no-context control condition. These findings replicate the results from study one and suggest that potential jurors can correctly assess the relative weight of evidence in the absence of case context. A one-way repeated-measures ANOVA was found to be significant ($p < 0.05$), and post hoc tests confirmed that the mean strength ratings were significantly different between each category of evidence strength, $F_{2,78} = 64.37$, $p < 0.001$.

The mean strength ratings for the DNA evidence evaluated in the case scenario conditions are also summarized in Table 2. A one-way ANOVA was used to compare the strength ratings of the DNA evidence between the three conditions. The results of the ANOVA indicated that there was a significant difference in the strength ratings, $F_{2,119} = 11.20$, $p < 0.001$, and post hoc tests confirmed that the DNA evidence in the weakest evidence condition was rated significantly lower than the DNA evidence in the other two conditions.

TABLE 2—*Mean strength ratings for all evidence strength conditions.*

Category of Evidence	Mean Strength Rating (no-context)	SD	Mean Strength Rating (with case context)	SD
High mobility/low relevance (weakest)	2.52a	0.67	3.10d	1.00
High mobility/high relevance (moderate)	3.28b	0.69	3.85e	0.74
Low mobility/high relevance (strongest)	3.73c	0.62	3.88e	0.69

Column means with different subscripts differ significantly at $p < 0.05$.

The impact of presenting case context on the mean strength ratings of the various strengths of DNA evidence is illustrated in Fig. 2. All categories of evidence are rated as stronger on average when accompanied by case context (compared to the control condition), and this difference is most prominent when the evidence is of a weak or moderate standard. It is also worth noting that the only condition in which evidence was rated as "weak" (e.g., below a mid-scale score of 3) is in the weakest evidence category, but only when no case context is presented.

Probability of Guilt Estimates—Participants estimated the probability that the defendant was guilty of murder at three stages in the trial scenario: The prior probability (after a brief summary, but before any evidence was presented), postevidence probability (immediately after all of the evidence/testimony was heard), and the final probability (after the prosecution and defense summaries).

The mean prior probability of guilt estimates did not differ between conditions, which is to be expected as this is before any evidence is heard in the case. To compare the postevidence and final probability of guilt estimates between conditions, a multivariate analysis of variance (MANOVA) was used. Table 3 summarizes the mean probability of guilt estimates for each condition, and Fig. 3 represents these relationships graphically.

Using Pillai's Trace, it was confirmed that there was an overall effect of strength of evidence condition on probability of guilt estimates, $F_{4,234} = 3.43$, $p < 0.05$. However, separate univariate ANOVAs for the measures of probability of guilt revealed that the only significant difference was between the weakest evidence and the other two evidence strength categories ($p < 0.05$).

Verdict Decisions—At the end of the trial scenario, participants had to render a verdict decision which could be guilty or not-guilty of murder. Table 4 summarizes the frequencies of each verdict for each of the three conditions.

There was a significant overall association between DNA evidence strength condition and verdict decision, χ^2 (2) = 6.87, $p < 0.05$; however, further analyses confirmed that only the strongest evidence category differed significantly in verdict decision from the two other evidence conditions (which did not differ significantly from one another).

Discussion

The findings in study one regarding participants' ability to evaluate the strength of various pieces of evidence seem promising as

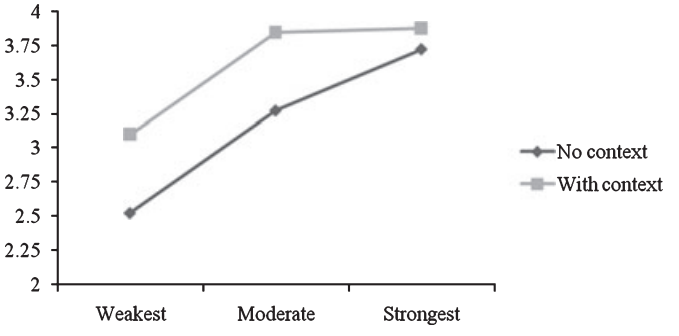

FIG. 2—*Effect of context on mean strength ratings for evidence categories (*N = *160).*

TABLE 3—*Mean probability of guilt estimations for each DNA evidence strength condition.*

Category of DNA Evidence	Prior Probability of Guilt	Postevidence Probability of Guilt	Final Probability of Guilt
High mobility/low relevance (weakest)	49.40a	55.28b	52.55d
High mobility/high relevance (moderate)	50.42a	67.95c	64.48e
Low mobility/high relevance (strongest)	50.75a	68.48c	68.50e

Means with different subscripts differ significantly at $p < 0.05$.

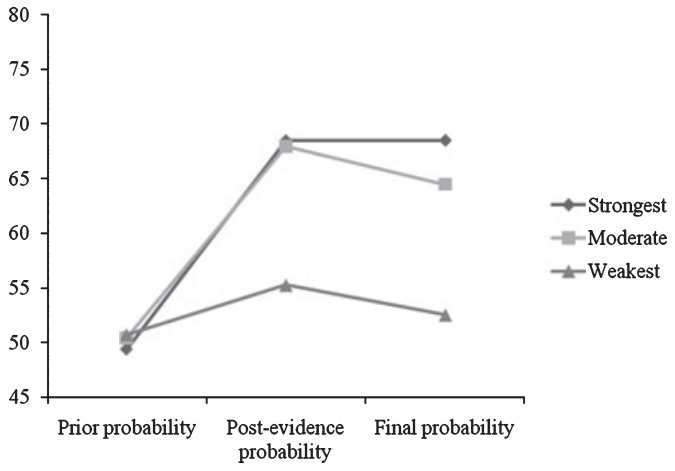

FIG. 3—*Mean probability of guilt estimates by DNA evidence condition (N = 120).*

TABLE 4—*Frequency and percentage of guilty and not-guilty verdicts by evidence condition.*

Category of DNA Evidence	Guilty Verdict	Not-guilty Verdict
High mobility/low relevance (weakest)	6 (15%)	34 (85%)
High mobility/high relevance (moderate)	9 (23%)	31 (77%)
Low mobility/high relevance (strongest)	16 (40%)	24 (60%)

they do suggest that mobility and investigative relevance of evidence are important factors in the evaluation process. The fact that participants were not given any contextual information about the evidence in the evaluation task suggests that mobility and relevance were indeed responsible for the strength ratings obtained in this study. The fundamental importance of relevance, above that of mobility, was also demonstrated in participants' ranking of the evidential strengths and the fact that there was no interaction between these two main effects. This finding suggests that potential jurors do understand the need for evidence to be indicative of guilt in order for it to be useful, regardless of its level of mobility.

By demonstrating that potential jurors are sensitive to two important factors of evidence strength (mobility and relevance) in the absence of any contextual information and that potential jurors in this study were not making strength judgments based solely on the type of evidence (e.g., DNA, fingerprints, footwear), the findings suggest that jurors are capable of understanding some of the intricacies which theoretically determine the probative value of forensic evidence. This seems contrary to some previous research, which has suggested that mock jurors do not reason effectively about evidence in courtroom situations (12,15,11,10), and study two addressed this issue further by investigating decisions made about the same strengths of evidence in the context of a criminal case.

Study two introduced elements of jury decision-making theory, namely the story model, by providing participants with enough case information to form a narrative representation of the evidence to evaluate its probative value. This narrative structure of the case information was also enhanced in this study by providing participants with the prosecution and defense summaries, which were presented as stories accounting for the evidence from two competing perspectives. The results suggest that the consideration of the forensic evidence in the context of a case served to enhance the perceived probative value of the evidence in all of the experimental conditions. This increase in perceived probative value was most prominent when the DNA evidence was of a moderate or weak standard, which is supported by the literature that takes an interactionist model approach to juror bias. This approach views juror behavior as influenced by both the situation (e.g., the specific trial and accompanying evidence) and the individual (e.g., juror personality, beliefs, attitudes, etc.) (23). In the case of a trial with very strong evidence, the situational factors will be especially salient, and therefore, this approach would predict that individual juror differences would play a relatively insignificant role in the decisions made by the jury. However, in cases where the trial evidence is weak or ambiguous, the lack of clear situational cues will result in a greater influence of juror biases and heuristics.

It is also important to discuss the impact of these findings on the verdict decisions made by participants, as this is the ultimate decision made by real jurors in the courtroom. The findings of this research suggest that although the weight of the moderate and weak forensic evidence was enhanced when accompanied by case context, this did not result in a corresponding increase in guilty verdicts in these conditions. This is an interesting result particularly considering that the reported probability of the defendant's guilt was similar in both the moderate and the strong evidence conditions; however, only when the forensic evidence was strong were participants more likely to vote guilty. This does seem to suggest that there is some appreciation of the difference between the moderate and strong evidence, despite the fact that this was not evident in the probability of guilt estimations or the strength of evidence ratings. One possibility is that the defense summary of the events was perceived as more plausible in the moderate evidence condition, and less convincing in the strong evidence condition.

This research, and indeed much of the literature investigating jury decision making, has some significant limitations that are important to acknowledge when interpreting the findings. The decision-making processes of real juries deciding on real cases are largely off limits to researchers, and this requires the use of simulated decision-making tasks in the research methodology. The use of artificial situations inevitably leads to issues with generalizing the findings to real juries. This is potentially exacerbated by the fact that participants in research projects often do not represent the characteristics of citizens who actually serve on juries. In many cases, jury research is carried out using student populations, which are often criticized for being substantially different in age and educational level from real juries (24), and the research presented in this paper attempted to improve on this particular problem using jury-eligible members of the public as participants. It would be ideal if future research of this type could be carried out with people who have acted as real jurors to improve the generalizability of the findings.

References

1. Caddy B, Cobb P. Forensic science. In: White PC, editor. Crime scene to court: the essentials of forensic science. Cambridge, U.K.: The Royal Society of Chemistry, 2004;1–20.
2. Gallop A, Stockdale R. Trace and contact evidence. In: White PC, editor. Crime scene to court: the essentials of forensic science. Cambridge, U.K.: The Royal Society of Chemistry, 2004;56–81.
3. Duncan CD. The future of DNA evidence. J Forensic Identif 2008;58: 283–96.
4. Pennington N, Hastie R. The story model for juror decision making. In: Hastie R, editor. Inside the juror: the psychology of juror decision making. Cambridge, U.K.: Cambridge University Press, 1993;192–221.
5. Pennington N, Hastie R. Evidence evaluation in complex decision making. J Pers Soc Psychol 1986;51:242–58.
6. Pennington N, Hastie R. Explanation-based decision making: effects of memory structure on judgement. J Exp Psychol Learn Mem Cogn 1988; 14:521–33.
7. Gigerenzer G. Fast and frugal heuristics: the tools of bounded rationality. In: Koehler D, Harvey N, editors. Blackwell handbook of judgement and decision making. Oxford, U.K.: Blackwell Publishing Ltd, 2004;62–88.
8. Schoemaker P. Forecasting and scenario planning: the challenges of uncertainty and complexity. In: Koehler D, Harvey N, editors. Blackwell handbook of judgement and decision making. Oxford, U.K.: Blackwell Publishing Ltd, 2004;274–96.
9. Thompson W, Schumann E. Interpretation of statistical evidence in criminal trials: the prosecutor's and the defense attorney's fallacy. Law Hum Behav 1987;11:167–87.
10. Smith B, Penrod S, Otto A, Park R. Jurors' use of probabilistic evidence. Law Hum Behav 1996;20:49–82.
11. Goodman J. Jurors' comprehension and assessment of probabilistic evidence. Am J Trial Advocacy 1992;16:361–89.
12. Nance D, Morris S. Juror understanding of DNA evidence. J Leg Stud 2005;34:395–444.
13. Schklar J, Diamond S. Juror reactions to DNA evidence: errors and expectancies. Law Hum Behav 1999;23:159–84.
14. Aitken C, Stoney D. The use of statistics in forensic science. West Sussex, U.K.: Ellis Horwood Limited, 1991.
15. Faigman D, Baglioni A. Bayes' theorem in the trial process. Law Hum Behav 1988;12:1–17.
16. Koehler J. When are people persuaded by the DNA match statistics? Law Hum Behav 2001;25:493–513.
17. Bond J. Value of DNA evidence in detecting crime. J Forensic Sci 2007;52:128–36.
18. White PC. Crime scene to court: the essentials of forensic science, 2nd edn. Cambridge, U.K.: The Royal Society of Chemistry, 2004.
19. Champod C, Lennard C, Margot P, Stoilovic M. Fingerprints and other ridge skin impressions. London, U.K.: CRC Press, 2004.
20. Evett I. Establishing the evidential value of a small quantity of material found at a crime scene. J Forensic Sci Soc 1992;33:83–6.
21. Robson C. Real world research, 2nd edn. Oxford, U.K.: Blackwell Publishers, 2002.
22. Field A. Discovering statistics using SPSS, 3rd edn. London, U.K.: Sage Publications, 2009.
23. Kassin S, Wrightsman L. The American jury on trial: psychological perspectives. Washington, DC: Hemisphere Publishing Corporation, 1988.
24. Bray R, Kerr N. Methodological considerations in the study of the psychology of the courtroom. In: Kerr N, Bray R, editors. The psychology of the courtroom. London, U.K.: Academic Press, 1982;287–324.

Additional information and reprint requests:
Lisa L. Smith, M.Sc., M.F.S.Soc.
University of Leicester
Department of Criminology
The Friars, 154 Upper New Walk
Leicester LE1 7QA
U.K.
E-mail: ls149@le.ac.uk

Appendix

Evidence Evaluation Items

Low Mobility/High Relevance Items (Strongest Evidence Category)

- DNA from a bloodstain at a crime scene matches a suspect's DNA.
- A suspect's fingerprint is found on a table in a burglary scene.
- Shoeprints are left in blood at the scene of a murder.

High Mobility/High Relevance Items (Somewhat Strong Evidence Category)

- A DNA sample from saliva on a beer bottle found inside a burglary scene
- DNA from sweat recovered from a mask left at a burglary scene
- Fingerprints are recovered from a discarded crisp pack in a stolen vehicle.

Low Mobility/Low Relevance Items (Somewhat Weak Evidence Category)

- A husband is suspected of killing his wife, and his fingerprints are found on the murder weapon, which is his hunting knife.
- Fingerprints are found on the outside of the door of a stolen vehicle.
- Shoeprints are found outside a burglary scene.

High Mobility/Low Relevance Items (Weakest Evidence Category)

- DNA recovered from discarded chewing gum found on the pavement outside a burglary scene
- DNA recovered from a discarded cigarette end found outside a stolen car
- Fingerprints are recovered from a tool discarded near the scene of a burglary (the tool may have been used during the burglary).