JURIES AND EXPERT EVIDENCE: SOCIAL FRAMEWORK TESTIMONY

NEIL J. VIDMAR* AND REGINA A. SCHULLER†

I
INTRODUCTION

During the past decade and a half, courts have been faced with an increasing number of attempts to have social science experts testify about such matters as eyewitness unreliability, post-traumatic stress disorders, or cross-cultural differences in the meaning of behavior. The purpose of the testimony, to which Walker and Monahan have given the generic label of "social framework evidence" in order to distinguish it from other forms of social science evidence,1 is to provide the factfinder, usually a jury, with information about the social and psychological context in which contested adjudicative facts occurred. It is presumed that knowledge about the context will help the factfinder interpret the contested adjudicative facts.

Courts dealing with the various types of social framework evidence have raised common concerns about both the extent to which such evidence may be helpful to the jury and the extent to which any probative assistance it provides may be outweighed by its prejudicial impact.2 Central to these

Copyright © 1989 by Law and Contemporary Problems
The authors wish to thank Robert Mosteller for comments on parts of the manuscript. Regina Schuller’s contribution was supported, in part, by a predoctoral fellowship in the Law and Social Sciences Program at Northwestern University and by a postdoctoral fellowship from the Social Sciences and Humanities Research Council of Canada.

* Professor of Social Science and Law, Duke University; Professor of Psychology and Professor of Law (1989-90), University of Western Ontario, London, Canada.
† Postdoctoral Fellow, Department of Psychology, Northwestern University.


2. For cases dealing with eyewitness reliability, see United States v. Langford, 802 F.2d 1176, 1181-85 (9th Cir. 1986) (Ferguson, C.J., dissenting); United States v. Amaral, 488 F.2d 1148, 1153 (9th Cir. 1973); Nelson v. State, 362 So. 2d 1017, 1021 (Fla. Dist. Ct. App. 1978) ("[I]t is within the common knowledge of the jury that a person being attacked and beaten undergoes stress that might cloud a subsequent identification of the assailant . . . . As such, the subject matter was not properly within the realm of expert testimony."); Jones v. State, 232 Ga. 762, 764, 208 S.E.2d 850, 852-53 (1974); Pankey v. Commonwealth, 485 S.W.2d 513, 522 (Ky. 1972) ([E]yewitness expert testimony "constitutes an invasion of the province of the jury in assessing the credibility of the witness."); State v. Stucke, 419 So. 2d 939, 945 (La. 1982) ("Such testimony invades the province of the jury and usurps its function."); Porter v. State, 94 Nev. 142, 148, 576 P.2d 275, 278-79 (1978) (noting "a substantial risk that the potential persuasive appearance of [the expert] would have had a greater influence on the jury than the evidence presented at trial, thereby interfering with the province of the
concerns are empirical questions about how juries receive and process the evidence given to them.

This article reviews empirical research bearing on the impact of three types of social framework evidence on jury decisionmaking: battered woman syndrome, rape trauma syndrome, and eyewitness unreliability. Despite the diversity of these substantive topics, some common patterns of jury behavior in response to social framework evidence emerge from the surveyed research. These patterns point to a tentative conclusion that juries can utilize social framework evidence in a legally appropriate manner. There are, however, some inconsistencies in the data and deficiencies in the research literature.

Part II of this article discusses the nature of social framework evidence and provides a broader conceptualization of it than that offered by Walker and Monahan. Part III discusses legal concerns about the evidence and conceptualizes them for empirical analysis. Part III also provides a brief overview of the social psychology of jury decisionmaking and how it may be affected by the introduction of social framework evidence. Part IV reviews research on the impact of evidence about battered woman syndrome, rape trauma syndrome, and eyewitness unreliability, as well as other research about how juries respond to expert evidence. Part V discusses the limitations of the research and offers suggestions for improving our knowledge about the impact of social framework evidence on juries.

II

The Nature of Social Framework Evidence

Laurens Walker and John Monahan developed the concept of "social framework" to capture the similarities among certain types of social science evidence and to distinguish social framework evidence from other uses of social science in law. A partial list of evidence that would qualify as social framework evidence would include testimony on eyewitness reliability, predictions of dangerousness, battered woman syn-

[Vol. 52: No. 4]


3. See Walker & Monahan, supra note 1, at 559.


5. For a discussion of specific social framework evidence used to predict dangerousness, see Walker & Monahan, supra note 1, at 559, 566.
drome, rape trauma syndrome, behavioral and psychological symptoms of sexually abused children, battered child syndrome, "brainwashed" defendants, Vietnam veterans syndrome, discriminatory behavior, effects of cultural stress, and psycholinguistic meaning.

Despite the wide variety of substantive topics and the use of the evidence for different litigation goals, all this evidence shares the common characteristic of employing general conclusions drawn from social science research to help evaluate factual issues in a specific case. In each instance, the evidence involves general assertions about some aspect of human behavior that is intended to help the trier of fact, usually a jury, to interpret the meaning of some disputed fact or testimony at issue in the suit.

In theory, social framework testimony does not bear directly on the ultimate fact to be decided by the trier; rather, it provides a social and psychological context in which the trier can understand and evaluate claims about the ultimate fact. Thus, a psychologist called to testify about potential eyewitness unreliability does not assert that the eyewitness in the particular case is wrong, only that studies of perception and recall indicate that people in

---


15. The potential uses of social framework testimony are as varied as inventive lawyers can devise within the constraints of evidentiary rules and policies. The testimony might be used to provide guidance about whether a future, present, or past event at issue at trial was more or less probable; it might be employed by either party in a criminal or civil suit; and it might pertain to plaintiffs, complainants, victims, defendants, or witnesses. See generally J. Monahan & L. Walker, supra note 7.

16. See Walker & Monahan, supra note 1, at 570.

17. Judges might equally benefit from the expert evidence if they are ignorant of certain social science findings about human behavior.

18. See Mosteller, supra note 8, at text accompanying note 85; Walker & Monahan, supra note 1, at 563-71, 580-83; see also Mosteller, supra note 8, at text in Parts IV.A.2 and IV. B.
similar contexts are often wrong. To cite another example, an expert called to testify about rape trauma syndrome, in a case where the defendant disputes the complainant’s testimony that she did not consent to sexual intercourse, describes studies on the characteristic reactions and behavior patterns of women who have been sexually assaulted. Although in some instances the expert may also be allowed to indicate that the complainant herself exhibits these characteristic symptoms, the testimony revolves around characteristics of a class of persons, sexual assault victims. The factfinder may then use this information to evaluate case-specific evidence.

Several matters regarding the conceptualization of social framework evidence need to be discussed at this point. First, as Walker and Monahan observe, not only is the basic thrust of social framework evidence probabilistic in nature, it may also address past, present, or future events. Where eyewitness reliability is in dispute, the issue is whether the eyewitness is mistaken about a past event. Testimony about dangerousness, of course, represents a judgment about future behavior. The expert testimony on allegations of sexual abuse of a child may relate to the validity of the child’s current retraction of previous allegations. Social framework evidence, therefore, is not bounded by time constraints.

A second characteristic of social framework evidence is that the research is often, though not invariably, “off the rack.” The studies relied upon by the experts are usually conducted and published well before the disputed events occurred, and involve persons having no connection to the legal proceeding. Often the research was never intended for use in a legal setting, but addressed a problem in another context. For example, the data on rape trauma syndrome were derived from studies originally intended to help devise therapeutic measures for rape victims. Many of the studies that experts cite in testimony about the unreliability of eyewitnesses were conducted in settings and contexts removed from the conditions involved in the identification offered in the specific case. In our view, however, Walker and Monahan place too much emphasis on the “off the rack” nature of social framework evidence. Research intended for social framework applications can be undertaken with the express purpose of use in a particular case. To cite one example, in cases where it is suspected that a line-up was biased, research bearing on the line-up procedure using groups of people having no connection to the case can be generated specifically for that trial. One can

---

19. Walker & Monahan, supra note 1, at 563-64.
20. See Mosteller, supra note 8, at text accompanying footnotes 182-86.
22. Walker & Monahan, supra note 1, at 568.
23. See Burgess & Holstrom, supra note 7.
easily conceive of litigation-driven social framework research with respect to other issues. In short, "off the rack" research is not a necessary characteristic of social framework evidence.

Social framework evidence may be based on large sets of systematically analyzed data bases or simply on clinical experience. Walker and Monahan cite *State v. Davis* as an example of a large-scale data set. In *Davis*, a defendant who pleaded guilty to murder offered the testimony of a sociologist in the penalty trial, which was held to determine whether the defendant should be executed or receive a mandatory thirty-year minimum sentence. The sociologist’s opinion was not based upon a personal evaluation of Davis. Rather, it was based upon published studies and government statistics indicating that murderers have low rates of recidivism and that persons of the age Davis would be after serving the mandatory sentence are unlikely to commit violent crimes. In contrast to these kinds of data, which are usually summarized by means of computer processing with disaggregation and probability projections according to type of offense or demographic characteristics of the offenders, consider *Barefoot v. Estelle*. In *Barefoot*, the defendant had been found guilty of murder; in the penalty trial, the prosecution called two psychiatrists to testify about the probability that someone like Barefoot would constitute a continuing danger to society. As in *Davis*, neither of these experts personally interviewed Barefoot. Unlike *Davis*, neither psychiatrist relied on aggregated statistical information. Rather, both were asked to testify about their own professional experience with convicted felons and were asked hypothetical questions about the likelihood that someone with Barefoot’s characteristics would commit violence in the future. The data bases were statistical only in the sense that they involved an aggregation of the psychiatrists’ own clinical experience with similar persons. Yet, in both *Davis* and *Barefoot*, the evidence referred to empirical relationships beyond the confines of the particular case.

Thus, the data bases from which social framework evidence is drawn vary along continuums of rigor and size. The type of data, the specific profession of the expert who presents the data, and other factors may affect how the jury

26. For social framework evidence to be useful, the case might actually require generating data where none currently exist—for example, rates of recidivism in groups of persons with certain demographic characteristics, or cultural meanings of behavior to particular members of a subculture. The possibilities for special data collection are, in fact, constrained only by the limits of time, money, and expertise.

27. Walker & Monahan, supra note 1, at 565, 569-70. Id. at 568 n.23.


receives and uses the evidence. Conceptually, however, social framework evidence can be defined solely by its function: to supply the triers with information about some aspect of human behavior to aid in interpreting disputed facts.  

III

LEGAL CONCERNS ABOUT THE IMPACT OF SOCIAL FRAMEWORK EVIDENCE ON THE JURY

Although social framework evidence involves different substantive topics and is tendered for different litigation goals, some common concerns regarding its impact on the jury have emerged. Two specific concerns are the extent to which the evidence may be helpful to the jury and the extent to which its impact may be more prejudicial than probative. Robert Mosteller's article, which appears elsewhere in this issue, thoroughly discusses the legal doctrines bearing on these concerns and the rationales underlying them. Therefore, the first section of Part III summarizes those concerns that speak directly to the jury's factfinding functions, then conceptualizes them for empirical analysis. The second section discusses research strategies for evaluating these issues empirically. The third section describes contemporary social psychological perspectives on jury decisionmaking processes that bear on the legal concerns. The ideas developed in this part provide a perspective for critiquing the research involving juries and social framework evidence, which is the subject of the final part of the article.

A. Evidentiary Concerns

Professor Mosteller points out that legal concerns with social framework evidence derive from the normative goals the jury trial is intended to serve. The different evidentiary rules and their underlying doctrines have evolved gradually, and address overlapping issues, which in themselves may be contradictory. Yet, in order to understand and examine the empirical assertions implicit in these evidentiary doctrines, researchers must spell them

32. As a practical matter, Davis and Barefoot probably define the extremes of a continuum. Most social framework testimony will combine the expert's professional experience and clinical observations of the relevant categories of persons with hard statistical data derived from his or her own research and the research of others.
34. Mosteller, supra note 8.
35. Other concerns may include the scientific reliability and validity of the evidence, the amount of trial time that might be absorbed by the testimony, the development of a "battle of experts," or concerns about the perception of fairness. See id. at Parts I and II. See also Goodman, Demographic Evidence in Capital Sentencing, 39 STAN. L. REV. 499 (1987); Lempert, Social Science in Court: On "Eyewitness Experts" and Other Issues, 10 LAW & HUM. BEHAV. 167 (1986); Tribe, supra note 21.
36. Mosteller, supra note 8, at note 86 and accompanying text.
out in such a way that they can be observed and measured. Then strategies of observation and measurement must be devised. The following discussion summarizes these evidentiary concerns and proposes means of observation and measurement.

At the outset, one significant assumption about social framework evidence should be made explicit: that a specific body of knowledge exists and that at least some persons, considered to be expert in the field, concur about its basic nature. This is an arguable assumption, and indeed some critiques have contested its validity in relation to each of the three areas reviewed in the next section. We justify our assumption here on the grounds that it is necessary to our analysis of potential jury behavior, and that court decisions and empirical research frequently point to the existence of such a body of knowledge.

Our analytical scheme draws upon the two legal concerns—the "helpfulness" of the evidence and whether it is "more probative than prejudicial"—and discusses them in terms of five issues. Two issues relate to the "helpfulness" concern: (1) whether the expert evidence provides knowledge or corrects mistaken ideas, and (2) whether the jurors integrate the information with their existing knowledge and utilize it. The remaining three issues address potential prejudicial effects: (3) juror responses to the "aura of science," (4) impermissible spillover effects on other trial evidence, and (5) excessive infringement upon the jury's role in determining witness credibility.

1. Providing "Knowledge" to the Jury. The purpose of social framework evidence, as with any expert evidence, is to assist the trier of fact by providing information that is either unknown to the trier or potentially at variance with what the trier believes to be true. A frequently expressed objection to social framework evidence is that it contributes little or no new information about human behavior to the pooled experiences of the jurors. In situations in which social framework evidence is used, the concern is often that jurors may hold misconceptions about some aspect of human behavior rather than that they have no knowledge. For instance, it is asserted that jurors are inclined to believe that there is a positive correlation between an eyewitness' confidence and the accuracy of an eyewitness identification. In fact, research shows that this assumption is often questionable. See Lindsay, Wells & O'Connell, Mock-Juror Belief of Accurate and Inaccurate Eyewitnesses: A Replication and Extension, 13 Law & Hum. Behav. 333 (1989). Similarly, it is said that laypersons often have misconceptions about rape victims. Burt, Cultural Myths and Supports For Rape, 38 J. Personality & Soc. Psychology 217 (1980); Tetreault, Rape Myth Acceptance: A Case for Providing Educational Expert Testimony in Rape Jury Trials, 7 Behavioral Sci. & L. 245 (1989). Jurors also have misconceptions about the credibility of child witnesses. See Goodman, Golding & Haith, Jurors' Reactions to Child Witnesses, J. Soc. Issues, Apr.-June 1984, at 139.

---

37. This is known as developing an "operational definition" and is essential for testing research hypotheses. For a general discussion of this process, see F. KERLINGER, FOUNDATIONS OF BEHAVIORAL RESEARCH 30-35 (1973); P. RUNKEL & J. McGrath, RESEARCH ON HUMAN BEHAVIOR 150-53 (1972).
38. See infra notes 87-89, 105, 121 and accompanying text.
39. In situations in which social framework evidence is used, the concern is often that jurors may hold misconceptions about some aspect of human behavior rather than that they have no knowledge. For instance, it is asserted that jurors are inclined to believe that there is a positive correlation between an eyewitness' confidence and the accuracy of an eyewitness identification. In fact, research shows that this assumption is often questionable. See Lindsay, Wells & O'Connell, Mock-Juror Belief of Accurate and Inaccurate Eyewitnesses: A Replication and Extension, 13 Law & Hum. Behav. 333 (1989). Similarly, it is said that laypersons often have misconceptions about rape victims. Burt, Cultural Myths and Supports For Rape, 38 J. Personality & Soc. Psychology 217 (1980); Tetreault, Rape Myth Acceptance: A Case for Providing Educational Expert Testimony in Rape Jury Trials, 7 Behavioral Sci. & L. 245 (1989). Jurors also have misconceptions about the credibility of child witnesses. See Goodman, Golding & Haith, Jurors' Reactions to Child Witnesses, J. Soc. Issues, Apr.-June 1984, at 139.
between the body of expert knowledge and the knowledge or assumptions made by lay persons, and, if so, the magnitude of the discrepancy.

Whatever discrepancies are uncovered between lay knowledge and the specific body of expert knowledge do not address the normative task of determining the threshold below which laypersons should be adjudged ignorant or misinformed. The quantification, however, does provide data around which the normative questions may be debated.

2. Acceptance, Integration, and Utilization. Assume that jurors do lack social context information known by experts, or that they hold misinformation about it, and that the expert is allowed to testify at trial. There are still questions about whether jurors are capable of accepting the evidence, of integrating it with the other information they possess, and of utilizing it in their decisionmaking. While this issue seldom receives much explicit attention in legal commentary or in psychological research on jury decisionmaking, there are a number of possible reasons why a jury might not use the expert testimony. The jurors simply may not understand the evidence and, therefore, ignore it entirely. Alternatively, the jurors may understand the evidence but not see its relevance. Thus, they might hear an expert testify about the typical reactions of victims of sexual abuse, but not grasp how it might apply to the complainant in the case they are deciding. This possibility is recognized in arguments that favor allowing the expert to establish a link between the general framework evidence and the witness in question. The arguments suggest that this link be forged by using hypothetical fact examples or by tendering an opinion that the subject witness exhibits characteristics corresponding to the group profile data that form the basis of the expert testimony.

A third possibility is that even though the jurors might understand the testimony and see its relevance, they nevertheless may reject it because it is in direct conflict with their pre-existing beliefs. Thus, for example, a juror who believed that “women usually provoke rape” might refuse even to consider contrary evidence bearing on sexual assault.

Still another possibility is that the expert evidence could be rejected because it is tendered after the juror has already formed an opinion about the witness who is the focus of the social framework testimony. The dynamics underlying the rejection of the evidence in this instance are similar to those

41. See Penrod & Cutler, supra note 4, at Parts VI and VII.
42. See, e.g., Mosteller, supra note 8, at notes 109-14 and accompanying text; see also Fox & Walters, The Impact of General Versus Specific Expert Testimony and Eyewitness Confidence upon Mock Juror Judgment, 10 LAW & HUM. BEHAV. 215 (1986).
43. See Penrod & Cutler, supra note 4, at Parts V.C.4, V.D. The phenomenon whereby people may reject valid information that is at odds with their pre-existing beliefs has been studied for over five decades and is well documented. See, e.g., M. SHERIF & C. HOVELAND, SOCIAL JUDGMENT: ASSIMILATION AND CONTRAST EFFECTS IN COMMUNICATION AND ATTITUDE CHANGE (1961).
involved with jurors who have pretrial biases, except that in this case the preformed opinion arises out of the trial process itself.\textsuperscript{44}

Finally, the jury might ignore or reject the evidence because of the form in which it is delivered to the jury. Social framework evidence may be transmitted to jurors through expert testimony or through judicial instructions. While the issue of whether a judge should instruct the jury on matters relating to witness evidence is not new, Walker and Monahan's proposal that social framework evidence should be delivered by a judge will certainly increase debate on this matter.\textsuperscript{45} For purposes of this discussion, the crucial point is that the form of presentation may affect a juror's willingness to accept or make use of social framework evidence.

3. The Impact of Science. Social framework evidence is derived from attempts to systematize and quantify aspects of human behavior. Its use therefore evokes more general legal concerns about the impact of scientific information on the jury. These concerns raise questions about the jury's ability to evaluate the merits of the body of knowledge and about the "misleading aura of science," or, in other words, the jury's proclivity to overvalue scientific testimony.\textsuperscript{46}

In circumstances where the assessment of the reliability and validity of the scientific information is left to the jury rather than to the court, a basic issue is whether the jury has the intellectual capacity to evaluate and apply scientific methodology to the forensic setting.\textsuperscript{47} The potential for confusion may be compounded when the expert is cross-examined on the underlying basis of her judgment, when a "battle of experts" results, or when the technique involves "fancy" instrumentation.\textsuperscript{48} Assessing jury capabilities in this regard requires determining how juries respond to questions about the methodological adequacy of the evidence.

\textsuperscript{44} Schwarzer, Communicating With Juries: Problems and Remedies, 69 Calif. L. Rev. 731, 754-55 (1981).


\textsuperscript{48} Imwinkelried, supra note 46, at 598-606. Social framework evidence will be unlikely to involve hardware other than a computer for processing the data. However, the measurement theories underlying the research, the statistical analyses of the data, and other methodological factors can be difficult to understand for persons not trained in the social sciences. See The Evolving Role of Statistical Assessments as Evidence in the Courts (S. Fienberg ed. 1989).
A second, related concern is whether the jury will evaluate the merits of the testimony using what Rosenthal has called the expert's "paramessage."

Simply put, do jurors ignore the evidence itself and instead rely on the credentials and demeanor of the expert and what they perceive to be her or his conclusions? A third concern is whether the jury will attach so much weight to the expert testimony that it will not give adequate weight to its own collective, intuitive experience, or, worse, will interpret the expert's testimony as dispositive of the trial issues.

4. Impact on Other Evidence and Decision Standards. Even if the jury appropriately understands the scientific basis of the evidence and approaches it with caution, social framework evidence could have inappropriate consequences for other evidence or for the standards against which the evidence is to be judged. This could occur in a number of ways.

The testimony might cause the jury to ignore or discount other evidence that is worthy of consideration. This phenomenon would result if the expert evidence were used as the primary standard against which to evaluate other evidence. For example, testimony bearing on the reactions of the complainant in a sexual assault case might cause the jurors to negatively evaluate evidence supporting the defendant's version of events. Similarly, expert testimony about eyewitness unreliability might cause the jurors to conclude that all eyewitnesses are unreliable, without considering other factors that point toward accuracy in the present case.

The evidence might also cause the jury to draw legally inappropriate conclusions about character. The Federal Rules of Evidence strictly proscribe the use of character evidence as a means of proving action in conformity with a dispositional tendency. This proscription is based on the theory that such evidence is likely to be irrelevant. Since social framework evidence offers conclusions about propensities of a category of persons, if jurors decide that the witness is a member of that category, they may impermissibly conclude that the witness' action or state of mind was a result of a dispositional propensity. Used in this way, social framework evidence is very similar to evidence of specific character traits used to prove action in conformity therewith.

The evidence might also confuse the jurors on matters pertaining to the legal standards. As Professor Mosteller points out, in a sexual assault case where the accused argues that the victim consented to sexual intercourse, the

---


50. Mosteller, supra note 8, at Part II.C.

51. Id.; see also Penrod & Cutler, supra note 4, at Part VI.C.

52. FED. R. EVID. 608.

53. Id. at Note.

54. Mosteller, supra note 8, at Part II.D; Walker & Monahan, supra note 1, at 580-82.

55. Mosteller, supra note 8.
legal issue in many states centers upon the defendant's state of mind—namely, did he reasonably believe that the complainant was consenting to sexual intercourse? However, the expert testimony might cause the jury to believe instead that the crime is legally defined as being concerned with the complainant's state of mind rather than the defendant's reasonable, good faith perception.56

5. Excessive Impact on the Determination of Credibility. Social framework evidence is intended to be used for a limited purpose, namely to provide a context for evaluating the testimony of a witness. It should not be used to assess witness credibility, which is solely the task of the jury.57 Concern about infringing excessively upon the jury's function in this regard forms the basis for prohibiting the expert from using terminology that might imply that a legal act did occur (for example, by requiring that the expert discuss post-traumatic stress disorder rather than the rape trauma syndrome58). This concern also underlies the limitations on how far the testimony may go in providing a cognitive link between the normative data and the witness to whom the data have relevance.59 In other words, if the jury relies too heavily or exclusively on the expert evidence to determine credibility, the impact of the testimony would be considered improper.

B. Research Strategies

At least five different strategies can be devised to assess empirically the discrepancy between lay knowledge or belief and the body of expert knowledge.60 One strategy is to cull the studies in the relevant social science area, select the findings upon which there is general agreement, and develop a test in multiple choice or some other format. The test can be given to groups of laypersons, and their responses can be scored against the "correct" answers derived from the literature.61 A variation on this strategy is to

56. Id. at notes 46-48 and accompanying text.  
57. Id. at Part II.E; see infra notes 79-105 and accompanying text.  
58. See People v. Bledsoe, 36 Cal. 3d 236, 681 P.2d 291, 203 Cal. Rptr. 450 (1984); State v. Allewalt, 308 Md. 89, 517 A.2d 741 (1986); State v. Taylor, 663 S.W.2d 235 (Mo. 1984); Mosteller, supra note 8, at note 46.  
59. See, e.g., People v. Poddar, 26 Cal. App. 3d 438, 103 Cal. Rptr. 84 (1974), vacated on other grounds, 10 Cal. 3d 750, 518 P.2d 342, 111 Cal. Rptr. 910 (1974) (anthropologist permitted to testify about general facts of cross-cultural differences but not to direct consequences of cultural stress); Mosteller, supra note 8, at note 57 and accompanying text; infra notes 128-58 and accompanying text.  
61. The validity of the "test" obviously depends upon choosing questions and answers upon which researchers agree. Another difficulty involves multiple-choice formats. As Penrod and Cutler, supra note 4, at 54, correctly point out, such tests may overestimate the knowledge that jurors have, because the questions may sensitize them to the importance of factors that might otherwise remain latent as they consider trial evidence.
identify groups of experts in the relevant area and then administer the test to them. The responses of the laypersons can then be compared to those of the experts. A particular advantage of this second strategy is that it permits estimation of the degree of consensus among the experts.62

A third strategy is to describe to laypersons the conditions of a psychological experiment or the conditions of some other setting in which data on human behavior have been collected. The laypersons, ignorant of the outcomes of the research, are asked to predict how the people in the experiment or study responded. The laypersons' predictions are then compared to the actual data collected in the study.63

A fourth strategy is to question actual jurors who served in trials involving the issues that could have been, but were not, addressed by social framework evidence, and to attempt to determine after the fact whether the jurors had adequate knowledge or held beliefs at variance with expert knowledge. This strategy could be more useful if the sample of jurors involved some who participated in trials where no expert testimony was tendered and some who heard expert testimony bearing on the issue.

A final strategy is to conduct simulated trials, varying the presence or absence of expert testimony and the form or type of testimony permitted. Individual jurors can then be questioned before and after trial about their knowledge, and these data can be compared as a function of experimental conditions. Additionally, the study could require that the jurors deliberate and that the deliberations be recorded. Subsequent content analysis of the deliberations can quantify the amount of information and misinformation expressed by the jurors, again comparing across the conditions.64

Strategies to assess the impact of social framework evidence are more complicated than strategies for examining the amount of knowledge jurors possess. Impact must be assessed in two comparative ways: comparison of the presence or absence of expert testimony, and comparison of how different parts of the trial evidence are utilized in relation to one another. Thus, while the ultimate legal concern is with the fairness of the outcome, the more immediate issue is the jury's decisionmaking processes, and specifically how the jurors use social framework evidence.

One research strategy is to measure changes in verdicts as a function of presence or absence of expert testimony or of variations in the form of expert testimony. This strategy, however, can provide only limited information about impact. Consider the introduction of evidence about battered woman syndrome by a defendant who claims self-defense to a murder charge. If the effect of the expert evidence, when compared to a trial in which no expert evidence is offered, is to reduce the verdict to a lesser included offense or to not guilty, the verdicts alone do not indicate why the different result occurred. The jury may have properly used the evidence in a limited way to assess the

---

62. See, e.g., infra notes 107-17 and accompanying text.
63. See, e.g., Human Intuition, supra note 60, at 262-65.
64. See infra notes 79, 101, 118 and accompanying text.
defendant's credibility, or may have improperly relied exclusively on the expert evidence as dispositive on the issue of her state of mind.  

A second strategy is to draw inferences about the process of decisionmaking by comparing expert testimony in conjunction with other variables. Consider expert testimony on battered woman syndrome in self-defense cases. If jurors are not being overwhelmed by expert testimony, then they should be sensitive to other evidence such as the degree of planning. This sensitivity should be reflected both in verdicts and in jurors' judgments on the merits of other evidence. Thus, if jurors are not being overwhelmed by expert evidence, we should expect them to give it less credence, and consequently for it to have less impact on verdicts, under conditions where battered woman syndrome is less relevant to the facts of the case. Studies can compare the presence or absence of social framework testimony (or the form of the testimony) in interaction with variables specific to the case. To the extent that verdicts (or judgments of evidence) vary from the pattern that proper legal usage would predict, we can infer that the jurors are using the evidence improperly. This strategy could be used with statistical data derived from actual juries or simulations. The weakness of the strategy, however, is that the data derived from the comparisons allow only an indirect inference about how the jurors used the evidence in arriving at a verdict.

Another research strategy for investigating jury decisionmaking processes is one that directly assesses how jurors use the expert evidence and the weight that it is given in relation to other evidence. This strategy involves interviewing actual or mock jurors after they have been exposed to variations in expert testimony conditions. Such studies can provide important insights into questions about how jurors used the evidence, and specifically about how much weight they accorded various pieces of evidence. However, such studies are subject to criticism because they are conducted after the fact and depend on the juror's memory and interpretation of how the evidence was used.

A final strategy is to examine the use of the evidence during the deliberation process itself. Such a strategy precludes the study of real juries. While simulation studies are subject to concerns about artificiality and generalizability to a real-world setting, they nevertheless allow examination of the basic social psychological processes that are the focus of legal concern. Thus, systematic analysis of jury deliberations can show how much time jurors

65. See Saunders, Vidmar & Hewitt, Eyewitness Testimony and the Discrediting Effect, in EVALUATING WITNESS EVIDENCE, supra note 60, at 71-76; Human Intuition, supra note 60, for similar critiques regarding the impact of expert testimony about eyewitness reliability.
66. See Penrod & Cutler, supra note 4, for an example of the strategy. See also H. Kalven & H. Zeisel, THE AMERICAN JURY 482-89 (1966); Zeisel, Reflections on Experimental Techniques in the Law, 2 J. LEGAL STUD. 107 (1973), for other examples and discussion of this inferential strategy.
devote to the various pieces of evidence, how they scrutinize it, and how they consider it in relation to other evidence.\textsuperscript{68}

C. The Psychological Process of Jury Decisionmaking

Legal analyses of jury decisionmaking, while recognizing that the task of integrating evidence and applying the law is a complex process fraught with potential hazards, seldom attempt to view the process in a holistic way. Social psychologists have developed models of the decisionmaking process, which are useful in addressing legal concerns about the impact of social framework evidence. Specifically, research and theoretical work indicate that the evidence and legal instructions provided to the jury during the trial are interdependent, such that changes in the presentation of one aspect of evidence is likely to create changes in the interpretation of other aspects. Moreover, the process of integrating the evidence and the law is a constructive or interpretive one in which the jurors bring their personal experiences to bear. A brief review of the social psychology of this process leads to the hypothesis that we should not expect other trial evidence to remain uninfluenced by social framework testimony.

There are a number of models of juror decisionmaking processes,\textsuperscript{69} but the most widely accepted model evolved from the work of Bennett, an anthropologist. He observed criminal trials in an attempt to determine the meanings that were given to the trial evidence.\textsuperscript{70} He proposed a model of the decisionmaking process fashioned around a metaphor of storytelling.\textsuperscript{71} His thesis is that jurors employ storytelling practices to organize and guide interpretation of trial information. The development of the story is guided by the implicit framework of judgment derived from the juror's life experience. This framework serves three interpretive functions that the juror must perform as she listens to the vast amount of trial information: She locates the central action of the story, she draws inferences regarding the relationship between key elements of the story, and she then evaluates the internal consistency and plausibility of the story.\textsuperscript{72} The story of the events at issue emerges out of this interpretive work and is used to fashion the verdict.

Building upon these essential insights as well as on a substantial body of other research on human decisionmaking processes, Pennington and Hastie
developed an elaborated model. It describes the jurors' task as consisting of three stages: a “story construction” stage, a “verdict category representation” stage, and a “story classification” stage. The first stage, which begins during the trial and extends into deliberations, involves the jurors' active organization of the evidence into a story form that describes what happened. The jurors use their own experience and beliefs about human behavior and its underlying causes to organize the events brought out at trial to produce one or more plausible accounts, or stories, of what occurred. During the second stage, the verdict alternatives or options outlined by the judge at the conclusion of the trial are developed into categories with defining features and decision rules that specify how they will be combined. These decision rules are built around the elements of the legal charge. Finally, during the third stage, the juror searches for the best possible match between the features of the stories developed in the first stage and the verdict categories elaborated in the second stage.

Story construction during the first stage of the process, therefore, is crucial to the two other stages. Moreover, jurors may develop a number of alternative stories about what occurred. The jurors draw upon two sources to construct an interpretation of the events: the evidence brought out at trial and their “factual” knowledge of the social and physical world. What is more, their pre-existing attitudes and beliefs influence the way that the trial evidence is perceived, recalled, and interpreted. There is substantial empirical support for this proposition.


74. Pennington & Hastie, Evidence Evaluation, supra note 73, at 242, 243-44.

75. Id. at 244.

76. Id. at 244-45.

77. To best illustrate this decision process, consider the following example provided by Pennington and Hastie:

Suppose that a person believes that the defendant . . . argued with the victim (initiating event), became very angry and decided to kill him (psychological state and intention), obtained a weapon, found the victim, got in a fight and stabbed the victim (actions), resulting in the victim’s death (consequence). The juror must decide which verdict category this story exemplifies by checking the mental state required by a verdict category (e.g., intent to kill for first degree murder) against the psychological states and the goals of the highest level episode of the believed story (intent to kill). There is a clear match in this example. However, in matching verdict-required circumstances (insufficient provocation) against the story-initiating events and states (argument), the juror has to judge whether an argument is an example of sufficient provocation. This matching process continues for all verdict attributes, for each verdict category. The best fitting verdict category is retained as a tentative decision.

Id. at 245.

deliberations frequently involves negotiating and reconciling the different stories developed by different jurors.

Another implication of the story model is that because the various pieces of evidence are developed into a coherent scenario, they are interdependent: A change in the meaning or weight in one piece will likely effect changes in others. This process may be best illustrated by reference to a study by Saunders, Vidmar, and Hewitt that examined the effects of a discredited eyewitness.\textsuperscript{79} In a simulation experiment, the jurors heard evidence about a breaking and entering case in which some circumstantial evidence linked the defendant to the crime. In one condition, there was only the circumstantial evidence, but in another there was also an eyewitness. Two other conditions in the experiment also contained an eyewitness, whose identification was subsequently discredited. The jurors' deliberations in reaching a verdict were recorded and subsequently analyzed. The analysis quantified the time and weight given to both the circumstantial evidence and the eyewitness. One important finding for our present purposes is that when an eyewitness identified the defendant, the jury gave greater weight to the circumstantial evidence than when there was only circumstantial evidence. Another is that although the weight given to the eyewitness' evidence decreased significantly when her identification was discredited, the weight accorded the circumstantial evidence as a result of the eyewitness remained high. This interdependence of evidence has also been demonstrated in other studies.\textsuperscript{80}

This view of the jury decisionmaking process has several important implications for projections of the potential impact of social framework evidence on the jury. First, the jurors' beliefs and attitudes will influence the way other trial evidence is evaluated and integrated into the stories they construct. Second, expert testimony about social frameworks, to the extent that it adds new information or corrects misinformation, will change the jurors' stories. Third, the expert evidence is likely to have spillover effects on other evidence. Thus, it should not be too surprising if social framework evidence pertaining to one trial witness has some effects on judgments about that witness' character or on evaluations of other witnesses or evidence at trial.

IV

RESEARCH ON THE EFFECT OF SOCIAL FRAMEWORK EVIDENCE

A. Battered Woman Syndrome

The battered woman syndrome is a descriptive term that refers to a cluster of typical behaviors and emotional reactions that may develop in a woman

\textsuperscript{79} Saunders, Vidmar & Hewitt, \textit{supra} note 65.

\textsuperscript{80} Casper, Benedict & Kelly, \textit{supra} note 78; Casper, Benedict & Perry, \textit{supra} note 78; Hans & Doob, \textit{supra} note 68.
repeatedly subjected to mental and physical abuse by a male with whom she is intimately involved.  

Described initially by Dr. Lenore Walker, the syndrome has two components. The key component involves what Walker refers to as a "cycle of violence." She posits that abusive relationships develop a patterned, repetitive cycle. There is an escalation of tension involving "minor" abuses—name-calling and other forms of verbal aggression—typically accompanied by the woman's attempts to placate the batterer. In a second phase, the tension erupts, the male becomes increasingly more aggressive, and an acute battering incident occurs. In a third phase, the batterer shows remorse, attempts to assist his victim, shows kindness, and promises change. The abuse, however, recurs; and the cycle may repeat itself many times in the relationship.

A second component of the syndrome involves the theory of "learned helplessness." As a consequence of the abuse, the woman develops feelings and perceptions that she has no control over her environment and no ability to escape from the relationship. The feeling persists as a general state even when the abuser may not be immediately threatening her. At some point, the woman may break out of this pattern of behavior and attack the batterer, perhaps seriously injuring or even killing him. In some instances, the attack may occur during the battering incident, but alternatively it may occur under circumstances in which the abuser is posing no immediate physical threat to her—for example, while the batterer is sleeping.

If the woman faces criminal charges arising from her actions, she may plead self-defense. This defense can be problematic because experts on battered women contend that the general public believes certain myths about battered women: If a woman is suffering she can simply leave her abuser; battered women derive pleasure from their beatings; battered women are at least partially deserving of the beatings; violent relationships are largely confined to persons who are on the low end of the socioeconomic continuum; and a woman waives her rights when she remains with the batterer.

At trial, the defendant may call an expert on the battered woman syndrome in an attempt to corroborate her claim that she believed herself to be in imminent danger even though the batterer may not have been actually physically threatening her at the time of the incident. There are several

82. L. Walker, supra note 6.
83. L. Walker, supra note 81, at 55.
84. Id. at 43.
reasons for calling the expert testimony. It will provide the jurors with information that "battered women" form an identifiable class of persons. It will serve to dispel misconceptions jurors may have about the effects of spousal abuse—for example, that a battered woman is free to leave at any time, or that she is a masochist and enjoyed the battering—thereby making more credible the woman's own testimony about the effects of past abuse. Finally, expert testimony will provide the jurors with a framework for understanding why the woman would have perceived herself to be threatened with imminent harm at the time of her actions, thereby helping to illuminate her mental state.

Courts have differed in their approach to the admissibility of testimony about battered woman syndrome. Some have ruled it admissible on the grounds that it could supply the jurors with an interpretation of the facts that might differ from those of the ordinary layperson. Others, however, have ruled it inadmissible because of concerns about its scientific status, the extent to which the evidence is beyond the knowledge of the jurors, or its potential prejudicial impact. There is limited empirical evidence bearing on each of these issues.

Although the scientific status of the battered woman syndrome is controversial, recent research indicates that some consensus exists among experts about the causes and effects of spousal abuse. Dodge and Greene constructed an eighteen-item questionnaire and surveyed forty-five professionals who had published research on spouse abuse. The data showed substantial agreement among the experts with respect to the majority of the items on the questionnaire. Thus, by this empirical measure the

---

86. See, e.g., Ibn-Tamas v. United States, 407 A.2d 626, 634 (D.C. App. 1979); see also generally citations, supra note 85.
89. The conceptual bases underlying the syndrome have been challenged as significantly flawed. *See particularly Note, supra note 6; see also Ibn-Tamas*, 455 A.2d at 894-95.

Items in the questionnaire tapped aspects of spousal abuse for which there appeared to be general consensus among experts in the field (for example, the woman's sense of blame, anxiety, low self-esteem, belief that her partner might kill her, sense of helplessness to stop the beatings, etc.). In addition, one item was designed to assess a concern typically raised by the courts in these cases, namely that the woman could not have been suffering severely or else she would have left the relationship.

To select their sample, Dodge and Greene identified 73 researchers who had published in the area of spouse abuse; 45 responded to the survey, for a response rate of 62%. *Id.* at 12.

91. In the survey, respondents read a brief description of a relationship involving domestic abuse and were asked to indicate the extent to which statements applied to the setting. The 14 items on which there was agreement included whether a woman in the described situation would believe her husband might kill her; would feel dependent on the husband; would feel a constant potential threat; would be persuaded to stay by promises never to be hurt again; would show anxiety and depression; would blame herself; would believe deadly force is the only alternative; would feel helpless; would deliberately provoke violence; would believe leaving would result in further harm; and could predict violence. *Id.* at 15.
battered woman syndrome appears to have gained some degree of scientific acceptance.

The next issue is whether this body of knowledge is generally familiar to laypersons or whether they in fact hold beliefs that are contrary to the findings. Dodge and Greene administered the same questionnaire used in their survey of experts to prospective jurors. This second survey permitted a direct comparison between lay and expert knowledge. Statistically significant differences between the laypersons and experts were found on twelve of the eighteen items. Greene, Raitz, and Lindblad administered a twelve-question survey to approximately 300 prospective jurors. Those jurors were relatively knowledgeable on aspects of spousal abuse, with less than 20 percent of the prospective jurors scoring in the opposite direction of the experts' responses on most items. There were also some gender and age effects among the laypersons: The beliefs of males and of older persons diverged more sharply from those of the experts than did the beliefs of females and of younger persons. The researchers also found that jurors' beliefs about the psychology of spousal violence differed as a function of the socioeconomic status of the couple described in the question. Jurors agreed more with the research conclusions if they were made in reference to a woman of low socioeconomic status. Greene and her colleagues concluded that jurors may evaluate the actions of a battered woman in light of their stereotyped attitudes about spousal violence.

For a recent doctoral dissertation, the second author of this article also obtained data concerning beliefs about battered women. Like Greene, Raitz, and Linblad, she devised a set of scales to measure the extent to which people endorse various beliefs about battered women. At present, the scale has been tested only with samples of college students; but the results are

92. Id. at 20.
93. Id. at 13.
96. Greene and her colleagues found that if a statement referred to a woman of low socioeconomic status, in contrast to high, jurors were more likely to believe that the woman “would believe that she is helpless to stop the beatings,” “would show signs of extreme anxiety and depression,” “would stay with her husband because she feels dependent on him,” and that “once violence is used in a relationship, it is always there as a potential threat.” Greene, Raitz & Lindblad, supra note 94, at 118.
strikingly similar to those of Greene and her colleagues. Moreover, males were more likely than females to subscribe to beliefs that are contrary to the literature on the dynamics involved in physically abusive relationships.

Thus, several studies on beliefs about the social and psychological contexts in which battered women exist suggest that the average juror may have an understanding on some issues that varies from conclusions of experts who have studied the phenomenon. However, it appears that jurors may be better informed than critics have suggested. Thus, while there are grounds for concluding that jurors might be helped by expert testimony on the battered woman syndrome, the data are not overwhelming.

Several studies have addressed the question of how jurors use battered woman syndrome testimony in their decisionmaking. Ewing reviewed forty-four cases in which a woman charged with homicide attempted to present expert testimony on the battered woman syndrome. In eighteen trials the testimony was excluded, and in each instance the accused was convicted. In contrast, in the remaining twenty-six cases in which the testimony was admitted, only one-third of the defendants were acquitted. This finding suggests that admitting battered woman syndrome testimony did not automatically result in acquittals. In a portion of the cases in which the expert testimony was allowed, however, the defendants were convicted of lesser included offenses. From verdicts alone, however, we cannot be sure whether the expert evidence caused the jurors to fail to convict on the main charge or whether other considerations shaped the decisions. After all, the verdicts were reached in different cases with different fact patterns.

Schuller conducted two experiments to explore the effects of battered woman syndrome evidence. In the first study, mock jurors, who did not deliberate before rendering verdicts, read a detailed transcript in which a battered woman who killed her husband claimed self-defense. Two presentational forms of the testimony were compared. In one condition, the jurors read expert testimony in which only the general research pertaining to the battered woman syndrome was presented (general condition). In the second condition the expert testimony also included an opinion indicating that the defendant fit the description of the syndrome (specific condition). A third group of jurors was presented with the identical trial material, but was not exposed to any expert testimony (control condition). The jurors were then required to render a verdict from among three verdict alternatives: guilty of second-degree murder, guilty of manslaughter, or not guilty by reason of self-defense. They also answered a number of questions about various aspects of the trial testimony.

The study found that the testimony affected verdicts, but only in the specific condition where the connection between the expert’s testimony

98. Compare id. at Table 2.5 with Table 3.9.
99. M. Dodge & E. Greene, supra note 90, draw a similar conclusion.
100. C. Ewing, supra note 6.
101. R. Schuller, supra note 97.
regarding the description of battered woman syndrome and the defendant was explicitly drawn. The effect was to shift verdicts away from second-degree murder. Insight into the way in which the specific expert testimony influenced decisions was gleaned from the jurors' ratings of various aspects of the trial. For example, they were asked to describe the pieces of evidence that they felt most influenced their verdicts. There were clear differences in the type of testimony cited. Jurors who were not exposed to expert testimony cited considerably more evidence unfavorable to the defendant's case than jurors who were provided expert testimony. Moreover, those presented with expert testimony accompanied by an opinion linking the battered woman syndrome to the defendant cited far fewer pieces of evidence that were unfavorable to her (for example, the conflicting testimony of another witness). Findings on other measures indicated that the specific form of expert testimony also led to differential perceptions of the defendant's ability to leave and escape the situation, the extent to which the defendant may have feared that her life was in danger, and the extent to which the defendant intended to kill her husband. In short, these jurors tended to believe that the defendant's "story" of self-defense was more plausible.

Schuller's second experiment also had general, specific, and no-expert conditions, but the study additionally examined group deliberations. The trial transcript used in the first study was presented to the jurors in the form of an audiotape with different actors playing the parts of witnesses, lawyers, and the judge. The jurors deliberated in groups of four or five persons to reach a verdict, and the individual jurors were then asked questions similar to those in the first experiment. The deliberations of the juries were tape-recorded and systematically analyzed to determine how the jurors used the battered woman syndrome testimony.

Unlike the first experiment, the second experiment found few differences between the general and specific expert conditions; but both were only moderately different from the control (no-expert) condition. As in the first experiment, the effect of expert testimony was to shift verdicts away from second-degree murder; however, the shift was to a verdict of guilty of the lesser included offense of manslaughter rather than to not guilty. Examination of the deliberations indicated that the effect of the expert testimony was to cause the jurors to attend more carefully to testimony favorable to the defendant, suggesting that the defendant was abused, that she feared for her life, and that she had experienced a loss of control over her actions. In addition, the jury tended to treat less favorably the testimony of a witness whose testimony contradicted part of the defendant's version of the events. The jury devoted very little deliberation time to explicit consideration of the expert testimony; to the extent that such deliberation did occur, it came when points of contention arose and was used to back up interpretations favorable to the defendant's position. The jury spent little time discussing matters pertaining to the defendant's character.
In neither of the experiments was there evidence that the expert testimony affected jurors' interpretation and understanding of the law. However, individual juror assessments in the second experiment, which also obtained some general measures of the defendant's character and credibility, showed that the expert testimony did cause them to evaluate her character more favorably. The effect was very modest, however.

Finally, in Schuller's first study, the more the individual jurors endorsed what some experts consider to be "myths" about abused women, the less likely they were to render a not guilty verdict. Moreover, in both studies males not only were more likely to endorse these myths than females, but they were also more likely to favor guilty verdicts and to offer unfavorable interpretations of the defendant's state of mind and of her ability to leave the situation in which she found herself.

We uncovered one additional simulation study that attempted to examine the effects of battered woman syndrome testimony on juror decisions. Unfortunately, the study is plagued by a number of methodological problems; some of the results, however, seem consistent with those reported by Schuller. For jurors who reported that they believed the expert, there was an associated leniency in verdicts.

In summary, the research data indicate that potential jurors may hold beliefs and attitudes about abused women at variance with the views of experts who have studied or had experience with abused women. In particular, males are likely to be skeptical about the fear the woman feels in an abusive relationship and about her inability to leave a setting in which abuse is threatened. These beliefs influence the way jurors interpret evidence about the defendant's state of mind. The introduction of expert testimony regarding the battered woman syndrome appears to have a salutary effect on the more prejudicial beliefs, but the primary effect seems to be to shift jurors toward the lesser included offense of manslaughter rather than toward a verdict of not guilty, although for some circumstances the effect could be a choice of not guilty by reason of self-defense.

While confirmatory research is needed, the data also suggest that expert testimony causes the jurors to give more attention to the social and psychological context within which the defendant claimed to be afraid and helpless, rather than causing them to reinterpret her character or


103. This finding is consistent with recent concerns that legal scholars have raised about the testimony. Schneider, supra note 85, for instance, has argued that battered woman syndrome testimony fails to provide an adequate explanation for the reasonableness of the defendant's actions. She suggests that although the term "battered woman syndrome" is used only descriptively, the content of the testimony and the import of the term carry a different meaning . . . .

The notion of battered woman syndrome contains the seeds of old stereotypes of women in new form—the victimized and the passive battered woman, too paralyzed to act because of her own incapacity . . . . It is in tension with the notion of reasonableness necessary to self-defense since it emphasizes the woman's defects and incapacity.
misinterpret the law. The two studies conducted by Schuller raise questions about the impact of the form of the expert testimony. In her first experiment, which examined only individual juror responses, only specific expert testimony that provided general information about battered woman syndrome and specifically identified the defendant as probably belonging to that class seemed to have an impact on the jurors. However, in her second experiment, where jurors deliberated as a group, the general testimony condition was as efficacious as the specific condition in affecting interpretation of the evidence.

B. Rape Trauma Syndrome Evidence

Rape trauma syndrome is a form of post-traumatic stress disorder used to describe an identifiable cluster of physical, behavioral, and psychological symptoms that may result when a woman is sexually assaulted. Expert testimony regarding rape trauma syndrome has been tendered for a number of purposes, but its most common use has been in rape trials where the defendant concedes that sexual intercourse occurred but asserts that the complainant consented. The prosecution tenders the evidence to demonstrate that the complainant’s emotional state and her behavior following the incident were consistent with the behavior of typical rape victims. Thus, the jury is provided with information that may help in the evaluations of the complainant’s testimony.

Courts in various jurisdictions have reached conflicting conclusions about the admissibility of rape trauma syndrome evidence. Some of the disagreement concerns whether jurors are adequately knowledgeable about the circumstances of rape and the responses of rape victims. Other disagreement centers on the potential prejudicial impact of the testimony. When the testimony has been held admissible, some courts have ruled that the expert may only testify about the general characteristics of rape victims without specific reference to the complainant because such references would speak too closely to the ultimate issue. In contrast, other courts have allowed testimony that the complainant exhibited symptoms consistent with a diagnosis of rape trauma. Research studies have addressed many of the issues raised by the courts.

Frazier and Borgida conducted an empirical study directed specifically to the question of the degree of lay knowledge about rape. After reviewing

---

Id. at 207, 216.
104. See citations, supra note 7.
106. See generally citations, supra note 105.
research on rape, Frazier and Borgida developed a Sexual Assault Questionnaire (SAQ) that contained eighteen questions pertaining to the circumstances of rape and its effects on the victims (for example, victim hesitancy to report the assault, or the degree of the trauma). The questionnaire was administered to two groups of experts on rape; the first were authors of empirical studies on rape and the second a group including social workers, psychologists, psychiatrists, nurses, members of the clergy, and counsellors who had attended a meeting of the Society for Traumatic Stress Studies. An additional questionnaire assessed the experts’ opinions about the scientific status of the data base related to rape trauma and its potential helpfulness to jurors.

Both expert groups showed high agreement with the statements that rape trauma syndrome is a form of post-traumatic stress disorder, that responses to rape are distinguishable from pre-existing psychological problems of rape victims, and that the adequacy of the existing scientific data base on rape would not dissuade them from testifying as experts in a rape trial. However, while the experts who had published studies on rape showed modest agreement with the proposition that “a coherent and internally consistent body of knowledge on the aftereffects of rape exists,” the experts from the Society for Traumatic Stress Studies showed modest disagreement. On the SAQ, the experts on rape averaged 84 percent correct answers regarding empirical findings about rape, and the post-traumatic stress disorder experts averaged 74 percent accuracy; but these differences were not statistically significant. More importantly, on many of the eighteen items both groups showed very high accuracy rates. Thus, while there were some reservations and disagreements regarding the scientific coherence of the body of knowledge on rape and its aftereffects, the expressed willingness to testify and the generally high rates of knowledge and agreement about specific scientific findings suggests that rape trauma syndrome would likely pass muster as having gained general acceptance in the field.

The important question, then, is how knowledgeable nonexperts were in comparison to experts. To address this issue, the researchers also administered the SAQ to two groups of laypersons, college students enrolled in an introductory psychology course and nonacademic employees of a state university. The student sample scored 58 percent correct answers on the SAQ, and the university employees scored 57 percent. The differences between the nonexpert and expert groups were statistically significant. When we turn to individual items on the SAQ, the nonexperts, in comparison to the experts, appeared informed on matters such as the frequency of acquaintance rape, the fact that victims often do not report rapes, the amount of time it takes to recover from an assault, and the effects of prior stress upon recovery.

108. Id. at 112, Table 2.
109. Id. at 112.
110. Id. at 113, Table 3.
111. Id. at 112.
The nonexperts, however, appeared to be less well informed regarding such matters as who is likely to be raped, the frequency of multiple victimization experiences, the behavioral changes following a rape, and whether anger is the most common initial reaction to being raped. The nonexperts also seemed to endorse views that delays in reporting a rape raise suspicions about the veracity of the claim and that there are more false reports of rapes than of other crimes.

On the whole, then, the Frazier and Borgida study suggests that there may be a number of areas where jurors could be helped by expert evidence, though on some matters they appear to have adequate knowledge. While Frazier and Borgida's study directly compared groups of nonexperts to experts, other research also tends to support the view that the average juror may have inadequate information about rape or may hold attitudes that would predispose him or her to be skeptical of complainant testimony in a trial involving a consent defense. In a random sample of 598 adults, Burt found high levels of endorsement of rape myths. Moreover, the more that people endorsed the myths, the less likely they were to consider a sexual assault committed on an acquaintance to be rape. Similarly, Feild studied 896 adult citizens and found that their pre-existing attitudes affected how they viewed the circumstances surrounding various allegations of rape. Another study by Feild showed attitudinal and knowledge differences between rape counsellors and laypersons. Taken as a whole, these studies and others support the view that jurors may not be adequately informed about certain aspects of the social and psychological context in which acquaintance rape occurs and, further, that the absence of this information may cause jurors to be skeptical of a complainant's version of the events.

Brekke and Borgida conducted a series of simulation studies that examined the utilization of rape trauma syndrome testimony under several different conditions of presentation. In all of the studies, the mock jurors heard an audiotaped version of a trial in which the defendant was charged with third-degree sexual misconduct for allegedly forcing a female

complainant to have sexual intercourse. Both parties had been acquainted prior to the incident, and both agreed that sexual intercourse had occurred. The issue at trial was consent.

In the first study, four conditions were compared to a control condition involving no expert testimony. In a "standard expert" condition, a psychiatrist testified about sexual assault, attempting to focus particularly on misconceptions about rape. In a "specific hypothetical" condition the expert provided the same testimony, but the prosecutor also posed a legally permissible hypothetical example to the expert, thus allowing the expert to link the general testimony to the evidence regarding the complainant’s specific behavior. In both the standard and the hypothetical conditions, the expert’s testimony came either at the beginning of the prosecution’s evidence or at the end. Although the jurors deliberated on the trial evidence, the focus of this first study was individual juror decisions.

With respect to verdicts, the specific hypothetical expert condition produced a significantly greater number of convictions than the control group condition. The standard expert condition produced a greater proportion of guilty verdicts than the control condition; but the differences were not statistically significant, suggesting that its effect on the jurors was weaker than the specific hypothetical condition. Moreover, other data suggested that the effects of the specific hypothetical expert condition were most pronounced when the testimony was given at the beginning of the prosecution’s case than at the end.

The researchers also assessed the perceptions of the jurors regarding various aspects of the evidence and the expert. In the specific hypothetical condition, in comparison to the control condition, the jurors were less likely to believe that the complainant consented or that she was responsible for the incident. The jurors in this condition also rated her as having greater credibility and viewed her moral character more favorably. Another finding was that males were less favorably disposed toward the complainant than females, assigning more responsibility to her and believing it more likely that she consented. On the other hand, there were no gender differences regarding the perception of the complainant’s moral character.117

Brekke and Borgida conducted additional analyses to attempt to determine how the jurors actually used the expert testimony.118 The taped deliberations of the jurors were examined to answer two sets of questions. The first set concerned discussion of the expert testimony itself, while the second set involved jurors’ discussion of other issues in the trial. Not surprisingly, the data showed that the expert testimony was used to support arguments in favor of conviction. Juries exposed to the more specific

117. In their second study, Brekke and Borgida created a "concrete" expert condition in which the expert testimony was interspersed with very specific examples intended to link the scientific evidence even more directly with the complainant. The concrete expert evidence had a greater impact on verdicts and perceptions of the complainant than did the more abstract evidence. Id.
118. Id.
hypothetical expert evidence, particularly when it came first in the prosecution’s case, were more likely to discuss its helpfulness for understanding the other evidence. Overall, however, direct discussion of the expert testimony was brief and limited. With respect to the discussion of other evidence, the findings showed that in comparison to the no-expert control condition, discussion centering around the complainant’s credibility was more frequent in all of the expert conditions. Furthermore, in the control condition discussion of the issue of the complainant’s resistance was tinged with more prodefendant interpretations.

Another study conducted by Brekke examined the impact of expert testimony on rape trauma syndrome evidence that was challenged by an opposing expert, thus addressing the “battle of experts” issue. The study consisted of three conditions: a single-expert condition in which the expert described general research on rape trauma syndrome and provided an opinion that the complainant exhibited characteristics consistent with rape trauma syndrome; a two-expert condition in which an opposing expert for the defendant testified about problems with the methodology and premises underlying rape trauma syndrome research (the opposing expert did not directly address the specific opinion offered by the prosecution’s expert); and a control condition in which no expert testimony was presented.

In comparison to the no-expert control condition, the expert conditions resulted in moderately small but significant effects on juror’s predeliberation verdicts and judgments about evidence. Following deliberations, jurors in the expert testimony conditions reported that the complainant was more credible and had a better moral character, and that she was less responsible for the assault than did jurors in the no-expert control condition. They also indicated that it was more likely that the defendant had committed the rape; as a result, they rendered significantly more verdicts of guilty and recommended harsher sentences than jurors in the no-expert control condition.

Jurors who heard the expert testimony also reported that they gave greater consideration to the issue of the defendant’s credibility and to the issue of force. These expert testimony effects, however, occurred primarily when the prosecution expert’s testimony was not contested by an opposing defense expert. On all measures except evaluation of the complainant’s moral character, the opposing expert condition did not differ from the no-expert


120. The study also included two additional conditions to examine the impact of another form of expert testimony, namely polygraph evidence. Paralleling the conditions for rape trauma syndrome evidence, in one condition the expert described polygraph techniques and research as well as providing an opinion that the complainant passed a polygraph test, while in the other condition an additional expert was called by the defense to rebut the prosecution’s expert. The opposing expert’s testimony focused on the problems associated with polygraph tests and not on the specific results of the test administered to the complainant. Few differences for the two types of testimony were found; that is, the results tended to parallel those found with respect to the rape trauma syndrome evidence conditions. Id. at 32, 39 at Table 1.
control condition, thus suggesting that the influence of the prosecution’s expert witness was negated by the testimony of the opposing expert.

In summary, the research findings suggest that potential jurors may harbor conceptions about rape and its aftereffects that differ from those held by persons considered to be expert in the field of sexual assault. Laypersons’ beliefs influence how they view a complainant’s version of events surrounding an alleged rape. Males tend to view complainants less favorably than do females. The experimental studies of the effect of expert testimony on juror verdicts indicated that it produced a greater proportion of guilty verdicts than did no-expert control conditions. Expert testimony called jurors’ attention to situational evidence, increased their discussion of the complainant’s credibility, and subsequently decreased their beliefs that the complainant consented or that she was responsible for the incident. The testimony also produced some spillover effects, causing jurors to rate the complainant as having greater credibility and better moral character. The research also suggested that the effects of expert testimony were stronger when the expert made direct links from general findings about rape victims to the complainant and when the testimony came early in the trial. One experiment indicated that challenge by an opposing expert reduced the influence of the first expert.

C. Eyewitness Reliability

Expert testimony bearing on potential eyewitness unreliability constitutes the third area of social framework evidence under research investigation. Psychologists have devoted more scholarly effort to the examination of factors influencing eyewitness identification than to any other area of law. For the past two decades this has been the high-profile area of social framework evidence, generating controversy within the field of psychology itself and evoking concerns in the legal community.121 While this subject receives detailed treatment in the article by Penrod and Cutler,122 we will summarize some of the findings here in our attempt to discover parallels with juror responses to battered woman syndrome and rape trauma syndrome evidence.

The concerns about expert evidence on eyewitness reliability are identical to those raised by the two other subjects. First, is there a body of knowledge about the factors influencing eyewitness identification, and can it be generalized to the legal setting? Second, to the extent that a body of knowledge does exist, can it help the jurors? Third, is the impact of expert testimony likely to be more prejudicial than probative?

A recent survey by Kassin, Ellsworth, and Smith asked sixty-three experts located in the United States, Canada, and Europe about the reliability of scientific data bearing on twenty-one factors associated with eyewitness identification.123 At least 80 percent of the experts agreed that research

121. For a review of the issues and controversy, see generally 10 LAW & HUM. BEHAV. 1 (1986).
122. Penrod & Cutler, supra note 4.
showed that eyewitnesses can be influenced by the following factors: the wording of questions, line-up instructions, misleading post-event information, attitudes and expectations, exposure time, unconscious transference, show-ups, and forgetting curves. They also agreed that the confidence of an eyewitness is often spuriously associated with accuracy. Over 70 percent of the experts also agreed that there was a cross-race identification bias in white witnesses, a tendency for witnesses to overestimate the duration of events, and problems with the construction of line-ups. Over 50 percent agreed about five other findings.\(^{124}\) If one accepts the sample of experts as representative,\(^{125}\) there are grounds for concluding that a consensus exists in the relevant scientific community about a number of findings pertaining to eyewitness reliability.

Other studies have attempted to determine the extent to which laypersons hold beliefs that are inconsistent with the body of expert knowledge. Deffenbacher and Loftus\(^ {126}\) and Yarmey and Jones \(^ {127}\) used a questionnaire approach, asking groups of laypersons questions similar to those posed by Kassin, Ellsworth, and Smith in their survey of experts. These studies indicated that although laypersons held beliefs consistent with some of the research findings, their beliefs varied substantially from the researchers’ conclusions on other issues.

Brigham and Bothwell took another approach.\(^ {128}\) They described the conditions of experiments involving eyewitness reliability to laypersons and asked them to predict the results. The predictions were then compared to the results actually obtained in the experiments. On the whole, laypersons grossly overestimated eyewitness reliability, causing Brigham and Bothwell to infer that jurors do not intuitively give proper weight to eyewitness identifications. Other studies have used trial simulation experiments to discover how laypersons evaluate eyewitness evidence. In a series of experiments, Wells and Lindsay and their colleagues \(^ {129}\) have consistently found that jurors place a great deal of weight on the confidence with which the witness identified the defendant, even though confidence has been shown

\(^{124}\) These findings were as follows: Hypnosis does not facilitate the retrieval of an eyewitness’ memory; judgments of color made under monochromatic light are not reliable; police officers are no more accurate as eyewitnesses than the average person; the presence of a weapon impairs the ability to identify the perpetrator’s face; hypnosis does not facilitate the retrieval of an eyewitness’s memory. Id. at 1090.

\(^{125}\) The list of experts for the survey was compiled from a search of the eyewitness literature in scientific journals between 1980 and mid-1986 and was supplemented by other sources. This selection strategy resulted in a pool of 113 prospective experts, of whom 56% responded to the survey.


\(^{127}\) Yarmey & Jones, supra note 60.


\(^{129}\) See Human Intuition, supra note 60. For a review of the earlier work and for a recent replication and elaboration, see Lindsay, Wells & O’Connor, Mock-Juror Belief of Accurate and Inaccurate Eyewitnesses: A Replication and Extension, 13 LAW & HUM. BEHAV. 333 (1989).
not to be correlated with accuracy. Some of the research reported in Penrod and Cutler's article also indicates that witness confidence is a significant component in juror judgments of credibility.

A final set of data bearing on juror knowledge comes from the Kassin, Ellsworth, and Smith survey of experts. In addition to asking the experts about the reliability of various findings, Kassin, Ellsworth, and Smith also sought their opinions about whether jurors would subscribe to the findings as a matter of common sense. The answers indicated that the experts believed jurors would be reasonably well informed on some issues, but that they would be unaware of others.

Taken as a whole, these sources indicate that laypersons' beliefs about the factors that make eyewitnesses credible often do not comport with the body of expert knowledge. It should be added, however, that on some matters lay intuition and experience seem consistent with the research findings. Thus, whether expert evidence might assist the jurors in a specific case will depend upon which factors are at issue in that case.

Loftus conducted two of the first experiments intended to examine the impact of expert testimony about potential eyewitness unreliability. In the first experiment, simulating jurors read a criminal trial transcript in which the primary evidence against the defendant involved eyewitness identification. In half of the trial transcripts, an expert testified about factors that might lead to eyewitness unreliability; in the other half, there was no expert evidence. The expert condition reduced the number of convictions. The second experiment was similar to the first, but the jurors deliberated. In the expert condition, the jurors spent significantly more time discussing the eyewitness testimony. A similar study conducted by Hosch, Beck, and McIntyre also found that an expert witness increased the amount of deliberation time devoted to eyewitness testimony and to other relevant evidence. While all three of these experiments indicated that the effect of expert evidence was to direct attention to the eyewitness, the experiments were not analyzed in a way that permits a conclusion about whether the effect was a desirable one. The expert evidence may have made the simulating jurors too critical of the eyewitness. Rather than cause the jurors to give closer scrutiny to the eyewitness

130. For a review of the literature assessing the relationship between eyewitness confidence and accuracy, see Bothwell, Deffenbacher & Brigham, Correlation of Eyewitness Accuracy and Confidence: Optimality Hypotheses Revisited, 72 J. APPLIED PSYCHOLOGY 691 (1987).
131. Penrod & Cutler, supra note 4, at Part V.C.4. Penrod and Cutler, it should be noted, did not find other factors having a significant influence on juror responses. Id.
132. Kassin, Ellsworth & Smith, supra note 123, at 1094.
133. The experts believed that jurors would be reasonably well informed on issues such as length of exposure time affecting reliability, cross-racial identification, and line-up fairness; they would not be adequately informed regarding the findings about the confidence-accuracy relationship, overestimation of the duration of events, postevent suggestions, or unconscious inference. See id.
testimony, the expert evidence may have caused them to discount it altogether.\textsuperscript{136}

Wells and his colleagues have conducted experiments examining the effect of expert evidence in a different way.\textsuperscript{137} Persons who had witnessed a staged theft attempted to identify the thief from a photograph line-up. The researchers then selected both witnesses who accurately identified the thief and those who did not. These witnesses were then examined and cross-examined in a manner similar to ordinary courtroom procedure. Videotapes of the examinations were played for simulating jurors, who were asked to rate the extent to which they believed the witness. Before undertaking this task, some of the jurors saw an additional videotape in which an expert offered cautions about witnessing conditions, such as findings about the relationships between witness confidence and accuracy, and between the amount of detail relayed by a witness and accuracy. In one study, the expert testimony did not increase jurors' ability to distinguish between accurate and inaccurate witnesses, but did reduce the extent to which the jurors believed the eyewitness. In the second study, however, jurors in the expert testimony condition believed inaccurate witnesses significantly less often than accurate ones. Additionally, the expert testimony induced the jurors to be sensitive to the effects of different viewing conditions. These effects did not occur among jurors in a no-expert control condition.\textsuperscript{138}

Maass, Brigham, and West conducted a simulation that varied the form in which expert testimony was presented to the jurors.\textsuperscript{139} In comparison to control group conditions the expert evidence conditions resulted in more lenient judgments of the defendant; however, the data suggested that the eyewitness was not entirely discounted. The study also found that when the expert testimony was linked to the eyewitness and accompanied by causal explanations, it had greater impact on the jurors.

Fox and Walters also studied the effects of different types of expert testimony in a simulation trial involving eyewitness evidence.\textsuperscript{140} In a general testimony condition, the expert discussed general findings about perception and memory processes, whereas in a specific testimony condition the expert

\textsuperscript{136} This observation has been made previously by Wells, \textit{supra} note 60, at 88-91, and by Penrod & Cutler, \textit{supra} note 4. Penrod & Cutler refer to the legally desirable impact of the expert as a "sensitivity" effect and to inappropriate discounting as a "skepticism" effect.

\textsuperscript{137} Wells, \textit{supra} note 60; Wells, Lindsay & Tousignant, \textit{Effects of Expert Psychological Advice on Juror Judgments in Eyewitness Testimony}, \textit{4 Law & Hum. Behav.} 275 (1980).

\textsuperscript{138} Sanders, \textit{Expert Witnesses in Eyewitness Facial Identification Cases}, \textit{Tex. Tech L. Rev.} 1409, 1439-64 (1986). Sanders has criticized the Wells, Lindsay, and Tousignant findings by suggesting that the effect of the expert may have been to cause the jurors to discount the eyewitness too much. While his criticism may be valid, it is blunted by a second study, performed by Wells and Wright and described in Wells, \textit{supra} note 60, which found differential effects according to witnesses and witnessing conditions. However, Wells has pointed out some limitations of his own research, including the artificiality of the setting, the fact that the expert was not cross-examined, and the fact that the jurors did not deliberate. \textit{Id.} at 88.


\textsuperscript{140} Fox & Walters, \textit{supra} note 42.
discussed effects of twelve specific factors relevant to the case that are known to influence eyewitness reliability. Though jurors in both expert testimony conditions were less likely to believe the eyewitness than those in a no-expert control condition, the specific expert testimony had a greater effect than the general testimony.

The research by Penrod, Cutler, and Dexter that is described in considerable detail in the Penrod and Cutler article also examined the effects of expert testimony. In a complex simulation trial, those authors varied a number of factors, including the absence or presence of expert testimony. The expert was rigorously cross-examined by the prosecutor. Additionally, the form of the expert testimony was varied. In some cases, the expert discussed only research bearing on memory processes, whereas in another condition the expert added quantitative data involving percentages of false and correct identifications obtained in eyewitness experiments. The simulating jurors who saw the videotapes in the various conditions of the experiment were asked to render verdicts and make ratings about the defendant's guilt as well as to respond to other questions. One important finding of the experiment was that expert testimony increased juror sensitivity to factors relevant to the eyewitness' ability to identify the defendant; there was little indication that the testimony caused the jurors to discount excessively the eyewitness testimony, or that it had significant effects on other evidence. Another important finding was that the elaborated expert testimony involving quantitative estimates had a greater impact than the more general expert testimony.

Judicial commentary or instructions merit attention, as they are sometimes offered as an alternative to expert testimony. Three studies have examined the effect of judicial instructions or commentary regarding eyewitness reliability. These studies varied the form of instructions given to the jurors. Katzev and Wishart presented simulating jurors with a videotaped trial in which an eyewitness played a crucial role in identifying a defendant accused of burglarizing a warehouse. Before they began deliberations, the juries heard one of three sets of instructions. In a control condition they heard standard instructions regarding jury responsibility, types of evidence, definition of the charge, and the burden of proof. In a second condition they heard the same instructions accompanied by a recapitulation by the judge of arguments that had been set forth by both prosecution and defense counsel. The third condition contained everything in the second condition plus judicial commentary that cautioned the jurors about eyewitness identification and reviewed some of the psychological issues involved in such identification. The third condition produced the greatest number of not guilty verdicts among

141. See Penrod & Cutler, supra note 4, at Parts VII, VIII.
142. The experiment obtained only the responses of individual jurors who did not have the opportunity to deliberate.
143. See Walker & Monahan, supra note 1.
individual jurors. It also reduced the amount of deliberation time in
comparison to the other two conditions.

Hoffheimer also conducted a simulation study involving eyewitness
identification in criminal trials. In addition to a control condition with general
instructions, there were two other instruction conditions. In the first, there
was an instruction that the government had to prove the defendant's identity
beyond a reasonable doubt. In the second, there was a particularized
identification instruction modeled after the instruction used in some federal
courts. Contrary to the intent of the particularized instruction, and to what
was predicted, the particularized instruction increased conviction rates.
Hoffheimer argues that the particularized instruction drew attention to factors
that would bear on eyewitness accuracy. This added attention encouraged
jurors to speculate and make inferences based on their own understanding,
which studies have demonstrated to be inaccurate.\textsuperscript{147}

Greene compared the absence or presence of judicial instructions about
eyewitnesses in two simulation experiments.\textsuperscript{148} The first judicial instruction
had a minimal impact on jury verdicts, but the instruction did cause the jurors
to place more weight on the eyewitness testimony. In a second experiment, a
more particularized judicial instruction reduced conviction rates.

Finally, a study described by Penrod and Cutler attempted to examine
many of the issues regarding expert testimony within a single experiment.\textsuperscript{149}
Using the basic trial stimulus materials employed in their previous studies,
Penrod, Cutler, and Dexter created six experimental conditions: a no-expert
condition; a defense-hired expert condition; an opposing expert condition; a
court-appointed expert condition; a Telfaire\textsuperscript{150}-type judicial instruction
condition; and a defense expert plus judicial instruction condition. In
comparison to the control group, the opposing expert condition induced the
jurors to be sensitive to the conditions under which the eyewitness made the
identification. Interestingly, however, both the opposing expert condition
and the court-appointed expert condition caused the jurors to devalue aspects
of the eyewitness testimony inappropriately. The defense expert condition
appeared to make the jurors sensitive to eyewitness factors without undue
devaluation of the testimony. As in the previous studies by Greene and by
Hoffheimer, the Telfaire-style instruction had minimal effect.

In summary, the accumulated findings suggest that laypersons' beliefs
about eyewitness identification are often inconsistent with those held by


\textsuperscript{146} United States v. Telfaire, 469 F.2d 552 (D.C. Cir. 1972).

\textsuperscript{147} Hoffheimer, supra note 145, at 60. Unfortunately, like so many other studies, Hoffheimer's experiment did not require the jurors to deliberate. An analysis of those deliberations could have provided data bearing on his intriguing speculations.


\textsuperscript{149} Penrod & Cutler, supra note 4, at Part VIII.

\textsuperscript{150} 469 F.2d 552.
experts in the area. On the other hand, as with the other two topics covered in this article, jurors are fairly well informed on some matters about eyewitness identification. Hence, the informativeness of the testimony for any particular case depends on the unique factors surrounding the identification. The experiments investigating the effects of the testimony on how jurors decide cases also suggest a number of conclusions. The inclusion of expert testimony does seem to alter aspects of the decision process. The testimony causes jurors to discredit eyewitness evidence and to render verdicts more favorable to the defendant. In those studies that have also included group deliberations, some have found that the introduction of the testimony increased deliberation time. Unfortunately, the lack of attention to the specific content of the deliberations makes interpretation of these effects problematic. Another finding is that the impact of expert testimony corresponds to the degree that it is linked to the specific issues in the case. Finally, some studies suggest that social framework evidence is likely to have more impact on jurors when it is presented through expert testimony rather than through judicial instructions.

D. The "Aura of Science" and the Expert: Corroborative Evidence

There is a persistent concern expressed in judicial opinions and legal commentary about jurors being unduly influenced by expert testimony because of the "aura of science," or the prestige of the expert. The findings from the studies we have reviewed indicate that while jurors do use social framework evidence they are certainly not overwhelmed by it. In light of the serious legal concerns, however, it is informative to consider any other findings consistent or inconsistent with these results. There are in fact a number of field studies and simulation experiments about reactions to social science evidence that address jurors' capacity for skepticism and critical evaluation.

Bridgeman and Marlow conducted lengthy interviews with sixty-five jurors who had participated in felony trials in California during 1974 and 1975. While the jurors stated that they attached considerable importance to the testimony of other witnesses, they claimed that they were not greatly influenced by the testimony of doctors or psychiatrists. Saks and Wissler interviewed ninety-seven jurors and found that psychiatrists and psychologists were deemed about as competent as eyewitnesses but less competent than other types of medical doctors, chemical/drug experts, firearms experts, and accountants. Rosenthal interviewed eight members of a jury in an arson case involving voice spectrogram evidence. The validity of the evidence was challenged, with the result that the prosecution called three additional experts

151. See supra notes 46-50 and accompanying text.
and the defense called seven. From interviews with eight jurors, Rosenthal concluded that when the jurors did not comprehend much of the highly technical discussion of spectrography, they shifted their attention to the personal characteristics of the experts as a basis for judging the credibility of their testimony.

In *The American Jury*, the classic study of criminal juries, Kalven and Zeisel compared rates of agreement between the judge and the jury. Agreement rates did not differ as a function of the presence or absence of experts in the trial. Myers also attempted to assess the impact of expert witnesses in her study of Indiana felony juries. She concluded that the presence or absence of expert witnesses was not related to conviction rates. Neither of these two studies assessed the impact of expert testimony with much rigor, but they do suggest that the jury is not overwhelmed by the mere presence of experts. This suggestion is consistent with simulation studies that examine the impact of specific types of expert evidence.

To explore jurors' responses to the insanity defense, Simon conducted a series of elaborate trial simulations with jurors selected from jury rolls in Chicago and St. Louis. Analysis of the tape-recorded deliberations indicated that jurors did not accept psychiatric evidence uncritically. Considerable skepticism was voiced toward the testimony. Far from rubber-stamping the psychiatric opinions, the jurors relied heavily on the trial record, especially details of the circumstances in which the defendant's behavior occurred. Simon also found that jurors made distinctions between the psychiatrist's clinical diagnosis and the moral and legal decision that they were required to render. Ellsworth, Bukaty, Cowan, and Thompson conducted another simulation experiment to determine whether jurors' responses to the insanity defense would differ as a function of the asserted etiology of the defendant's mental problems. The research did not assess the impact of expert testimony itself, but it did indicate that the substance of psychiatric testimony may affect the jury's receptiveness to that testimony. While jurors readily accepted evidence regarding mental retardation and psychomotor epilepsy, many were likely to reject the diagnosis of schizophrenia. The data from these studies seem consistent with nationwide court statistics indicating that the success rates of defendants who plead not guilty by reason of insanity is low; between 75 and 99 percent of defendants who enter such pleas are found guilty.

A number of studies have attempted to assess the impact of polygraph testimony on jurors. Polygraph evidence, like social framework evidence,
draws upon statistical data about normative responses of groups of people as part of its underlying scientific rationale. Its reliance on sophisticated physiological measurement devices, electronic recording equipment, and assertions about statistical accuracy rates would seem to enhance its "aura of science."

A 1938 study involved interviews with jurors who had heard a case in which polygraph evidence was used, and indicated that the testimony was accorded substantial weight. But nearly four decades later, a set of interviews with jurors who heard another case suggested that the jurors had used the evidence only in a limited fashion, principally as corroboration of the other evidence.

Other relevant data come from simulation experiments. Markwart and Lynch assessed the impact of polygraph evidence in a simulated felony-murder trial. A control condition containing no polygraph evidence was contrasted both with a condition in which polygraph evidence unfavorable to the defendant was introduced and with a condition in which the polygraph evidence was favorable to the defendant. The trial evidence was followed by a forty-minute videotaped instruction by a judge, who characterized polygraphs as having a 90 percent accuracy rate but emphasized that they tended to misidentify innocent persons as not telling the truth. The unfavorable polygraph results produced more verdicts against the defendant than the control condition, and the favorable polygraph results produced more verdicts in favor of the defendant. Tape recordings of the deliberations of the jurors, however, suggested that little time was devoted to discussion of the polygraph evidence and that some jurors dismissed it out of hand. Post-deliberation questioning of the jurors suggested that the polygraph testimony was rated as having less significance than the other evidence. These last two analyses, however, are not consistent with the fact that the verdicts were affected. One explanation is that, as found in the Saunders, Vidmar, and Hewitt experiment described earlier, the polygraph evidence caused the jurors to give more weight to the other evidence consistent with the polygraph testimony about the truthfulness of the defendant.

Cavoukian and Heslegrave conducted two other experiments involving polygraph evidence that tended to exonerate the defendant in a felony-murder trial. In both experiments the polygraph testimony produced higher acquittal rates than the control conditions in which polygraph testimony was absent. However, the study also found that a caution by a judge or by an expert witness about the accuracy rate of polygraphs reduced the impact of the testimony. These authors concluded that there is no evidence to support

the hypothesis that jurors will blindly accept polygraph evidence, at least not when they are cautioned about the potential unreliability of the instrument.

Loftus conducted a simulation experiment to determine the relative weight jurors would give to scientific evidence in comparison to testimony by an eyewitness. The stimulus case involved charges of passing a bad check. In the control condition, an eyewitness identified the defendant as the person who had passed the bad check. In a second condition, expert polygraph evidence incriminated the defendant. A third condition involved expert evidence that the fingerprints on the check matched those of the defendant; and a fourth condition involved expert evidence that the handwriting on the check matched that of the defendant. The rates of conviction in the four conditions were as follows: eyewitness, 78 percent; polygraph, 53 percent; fingerprints, 70 percent, and handwriting, 34 percent. Thus, while the impact of the expert on verdicts varied between expert conditions, none was as influential as the layperson witness.

Greene, Wilson, and Loftus studied the impact of hypnotically refreshed testimony on jury decisionmaking. Some courts exclude such testimony, others admit it under limited conditions, while still others have ruled it admissible, reasoning that the jury can decide how much weight it should be accorded. Some of the reservations about admissibility stem from the belief that, since the jurors have no experience in assessing the potential unreliability of the body of science on hypnosis, they will view it as having mystic infallibility, and that they will therefore give hypnotically refreshed testimony undue weight. To examine these assertions, Greene, Wilson, and Loftus conducted simulation studies with prospective jurors and with college students. The jurors watched a ninety-minute videotape of a trial involving a stabbing incident, answered a predeliberation questionnaire, and finally deliberated in groups to reach a verdict. In one of three conditions, an important prosecution witness provided a police officer with crucial details immediately after the incident; in another condition, the details were not recalled until a week later; and in a third condition, the details were recalled only after hypnotic refreshment one week after the incident. The major finding, derived from both the predeliberation questionnaires and from systematic analyses of the tape-recorded deliberations, was that the jurors viewed the hypnotically refreshed testimony with a reasonable degree of skepticism. Additionally, the results showed that jurors upon learning that one prosecution witness' testimony had been refreshed hypnotically, tended to treat the testimony of all the prosecution witnesses as less credible.

Goodman conducted two simulation experiments intended to assess how jurors respond to probabilistic evidence. The first involved college

students assuming the role of jurors in a murder trial in which expert testimony about the matching of blood samples was critical to the prosecution’s case. The jurors rated the blood test evidence favorably but accorded it lower levels of reliability than that asserted by the experts. Only 23 percent of the students found the probabilistic evidence adequate to convict the defendant.

In Goodman’s second experiment, the simulating jurors were 223 persons eligible for jury duty (31 percent of whom had previously served as jurors). The case involved a claim of arson, either in the context of criminal charges or of a civil lawsuit in which the owner of a factory sued his insurance carrier. Each side of the case called both lay and expert witnesses, supporting either a finding of arson or a finding that the fire was accidental. The expert evidence involved chromatographic tests on gasoline samples. A number of findings from this experiment bear on the impact of expert witnesses.

First, jurors gave more weight to the testimony of one of the lay witnesses than to either of the experts. Second, they appeared to give careful attention to relevant factors in deciding which testimony of the conflicting experts should be given more weight. For example, one expert had conducted an on-site inspection while the other had not. Third, the jurors assigned relatively more weight to the probabilistic evidence when it had more probative value. Fourth, there was no evidence that the jurors confused the burden of proof with the probabilistic evidence. Regardless of whether the burden of proof was beyond a reasonable doubt or balance of probability, the proportion of findings of arson did not vary. Fifth, there was an indication that while the expert evidence was used, it was given less weight than it should have been accorded. The last finding is consistent with more general findings about human decisionmaking, but is inconsistent with the hypothesis that jurors will accord such evidence too much weight.

It is important to note that Goodman’s jurors did have difficulty understanding probabilistic evidence, particularly in converting probabilities to percentages and vice versa. Moreover, the impact of the evidence varied according to the form in which it was described to the jurors. Jurors were less likely to use evidence presented in percentages than evidence presented in probabilities. These findings are generally consistent with earlier research by

---

168. The purpose of these two conditions was to develop two burdens of proof, “beyond a reasonable doubt” in the criminal trial and “balance of probabilities” in the civil trial, in order to test the hypothesis that jurors might confuse the probabilistic evidence provided by the expert with burdens of proof.

169. The probabilities associated with the evidence were also varied in the experiment to test jurors’ sensitivity to its incriminating value.

Thompson and Schumann and by Faigman and Baglioni. The results bear on the ability of the jury to evaluate the validity of scientific techniques as opposed to foundational evidence.

Case studies of jurors involved in civil trials in which experts played an important role provide some additional information. Austin interviewed jurors from two trials arising out of an antitrust case. His interviews suggested that the jurors were quite skeptical of an economist witness, probably as a consequence of a rigorous cross-examination. Selvin and Picus interviewed jurors involved in a complex asbestos case. One of their conclusions was that the "jurors were . . . generally skeptical if not negatively disposed toward many of the medical experts who testified." Finally, as part of an extensive project on medical malpractice litigation, the first author, his colleagues, and their students have interviewed jurors who were the factfinders in a number of medical malpractice trials. These trials involved not only medical experts, but also economists and actuaries. One of the most striking findings from these interviews is the skepticism that jurors displayed toward experts. They were sensitive to the fact that the trial is an adversary process; in many cases, they spontaneously indicated to interviewers that experts are chosen to testify because their opinions favor the party that calls them.

V

Conclusion

We can now return to the five areas of legal concern about social framework evidence discussed in Part III.

A. Providing Knowledge to the Jury

With respect to the three areas reviewed in this article—battered women, rape victims, and eyewitnesses—empirical studies show a degree of consensus existing among experts about each of the respective bodies of knowledge. The studies also indicate that laypersons frequently hold beliefs or impressions at variance with these bodies of expert knowledge. This last


173. See Imwinkelried, supra note 46, at 600-04.

174. A. Austin, Complex Litigation Confronts the Jury System: A Case Study (1985).


176. Id. at 27. Selvin and Picus also observed that when the testimony was complex and confusing the jurors reported that they evaluated the witness' evidence based on their personal impressions of that expert.

177. The Medical Malpractice Research Project was conducted with my colleagues Thomas Metzloff and David Warren under the auspices of the Private Adjudication Center at the Duke University School of Law. The data have not yet been reported in scholarly articles.
conclusion must be qualified in light of the fact that within each of the areas the data indicate that laypersons tend to be relatively better informed on some matters than others. Additionally, on the topics of battered women and rape victims there appear to be some gender differences, with women being better informed than men. 178

The potential assistance of social framework evidence for jurors depends, of course, on the specific subject matter and the context to which it is applied. Thus, the potential contribution of other types of social framework evidence to jurors' knowledge will have to be assessed independently.

B. Acceptance, Integration, and Utilization

The studies covered in our review indicate, on the whole, that laypersons will accept social framework evidence and incorporate it into their decisionmaking. The data do indicate that the likelihood of the evidence being integrated into the decisionmaking process is greater when the information conveyed by the expert is directly linked to the witness to whom it relates and the information is conveyed in concrete terms. For example, testimony about the responses of rape victims is more likely to be used by jurors when it is specifically tied to the complainant in the case than when it only describes typical rape victim behavior. This finding is not surprising in light of a large body of recent work on human decisionmaking showing that scientists and lawyers as well as laypersons are prone to value concrete examples more than abstract examples or statistical evidence about base rates. 179

178. The implications of these differences among jurors for the jury deliberation process have not been explored in any of the existing literature, but they raise some interesting questions. An important rationale for the jury's existence is that the perspectives and experience of the various individual jurors will be shared during deliberations, making the collective wisdom of the jury greater than the sum of its individual deliberators. See Forston, Sense & Non-Sense: Jury Trial Communication, 1975 B.Y.U. L. REV. 601; V. HANS & N. VIDMAR, supra note 159, at 47-61, 97-112. Empirical studies have shown that this process does indeed occur, see, e.g., Dashiel, Experimental Studies of the Influence of Social Situations in the Behavior of Individual Human Adults, in HANDBOOK OF SOCIAL PSYCHOLOGY 1097 (C. Murchison ed. 1935); Forston, supra note 178; Forston, Judge's Instructions: A Quantitative Analysis of Jurors' Listening Comprehension, 18 TODAY'S SPEECH 34 (1970), and is quite consistent with other literature and theory about social influence in groups, see generally I. STEINER, GROUP PROCESS AND PRODUCTIVITY (1972). Thus, one could hypothesize that during the deliberation process the better-informed jurors might convey their insights to the less-informed jurors. If this hypothesis was confirmed, it might be argued that there is no need for expert testimony. Critics of this argument, however, could contend that the less-informed juror would resist the arguments of the better-informed jurors and that the expert testimony will be used to bolster the arguments of the better-informed jurors during deliberations. Additionally, those in favor of introducing expert testimony might point out that the data tend to show that even the best-informed jurors have knowledge levels below those of experts, and might argue that expert testimony is, therefore, needed in any event. We point out these possible conflicting positions as a matter to be investigated in subsequent empirical research.

The research data also suggest that when social framework evidence precedes the relevant lay witness evidence, it tends to have a greater impact. This result is consistent with other research on jury decisionmaking and with additional bodies of social psychological literature.\textsuperscript{180}

Research comparing the delivery of social framework information by judicial instructions to delivery by means of expert testimony has been limited to studies of information bearing on eyewitness reliability. These studies indicate that delivery of social framework evidence through judicial instructions will not be as effective as delivery by means of expert testimony. More research is needed, however, to determine whether this finding can be generalized and why judicial instructions prove less effective. The topic is of particular importance given Walker and Monahan's proposal that judicial instruction is the preferable vehicle for delivering social framework evidence.\textsuperscript{181}

C. The Impact of Science

The data in the studies reviewed in this article speak forcefully to the assertion that jurors will be seduced by social framework evidence and accord it an unwarranted aura of trustworthiness and reliability. The data include interviews with actual jurors and simulations involving a number of kinds of social science testimony. Jurors do not appear to suspend their own judgment in deference to the expert. The expert's testimony is evaluated in light of the juror's own experience, common sense, and recognition of the adversarial nature of the trial process.

A limited number of simulations studying rape victim testimony and eyewitness expert testimony have included conditions involving conflicting expert testimony. These studies led the researchers to conclude that the jurors were not confused by a "battle of experts.”

The existing studies have not directly addressed the question of whether jurors are capable of understanding the scientific rationale that underlies the various types of social framework evidence. The work of Thompson and others\textsuperscript{182} indicates that jurors may misunderstand probabilistic evidence; yet probabilistic estimates lie at the heart of much social framework research. To the extent that a scientific critique of social framework evidence becomes an issue that jurors are called upon to judge at trial, there could be a problem. However, the methods for deriving social framework evidence may be easier for jurors to grasp intuitively than the methods for fingerprint or hair-sample matching. Research is needed on this issue.

\textsuperscript{180.} Elwork, Sales & Alfini, \textit{Juridic Decisions: In Ignorance of the Law or in Light of It?}, 1 LAW & HUM. BEHAV. 163 (1977).

\textsuperscript{181.} See Walker & Monahan, \textit{supra} note 1.

\textsuperscript{182.} See Faigman & Baglioni, \textit{supra} note 172; J. Goodman, \textit{supra} note 167; Thompson, \textit{supra} note 171; see generally Saks & Kidd, \textit{supra} note 170.
D. Impact on Other Evidence and Decision Standards

This heading involves issues such as whether social framework testimony might affect the weight that jurors give to other trial evidence, cause them to draw legally inappropriate conclusions about a witness' character, or cause them to misunderstand the law. In Part III.C we hypothesized that because jurors proactively develop story schemas in arriving at their verdicts, we should expect some spillover effects on other evidence. Some of the studies address the magnitude and extent of that spillover.

As predicted, the various sets of data did show some spillover effects, but the magnitude of the effects was relatively small. For example, Brekke's study of rape trauma syndrome testimony found that the evidence affected assessments of the complainant's credibility and character. However, in comparison to no-expert control conditions, the testimony had only minor effects on the jurors' ratings of the defendant's character. Similarly, Schuller found that while testimony about the battered woman syndrome increased jurors' beliefs in the defendant's credibility and honesty, it had no statistically significant effect on the perception or evaluation of the deceased husband. The research described by Penrod and Cutler involving eyewitness reliability indicated that while expert evidence caused jurors to be more sensitive to witness viewing conditions, it did not cause them to give extra weight to factors irrelevant to the witness in question. Finally, Schuller's study indicated that jurors' understanding of the legal verdict categories did not differ as a consequence of the expert testimony.

The total number of studies that have directly addressed the issues raised under this rubric is limited. Nevertheless, our tentative conclusion is that from a statistical perspective, the spillover effects of social framework evidence are not large. We leave to legal decisionmakers the question of what constitutes undue impact.

E. Excessive Impact on the Determination of Credibility

The accumulated data from the various studies reviewed in this article lend no support to the argument that jurors allow the expert evidence about social frameworks to substitute for their own judgments about the credibility of a lay witness. The central purpose of social framework evidence is to help jurors assess credibility; and to the extent that the evidence bolsters or detracts from the relevant witness's evidence, we would expect it to influence credibility judgments. Thus, the key issue is whether it affects those judgments directly or indirectly. Given that, under the law, the jury is granted ultimate responsibility for assessing credibility, it would be preferable that the effects be indirect.

With respect to evidence about battered woman syndrome, Schuller's study suggests that the effects of social framework testimony are in fact indirect. Schuller's analysis of jury deliberations showed that while jurors discussed matters such as the woman's ability to leave the setting, her perceptions of fear and danger, her intentions, and the extent to which she
was in control of her actions, the issue of her general credibility accounted for less than 3 percent of deliberation time. Other studies we reviewed, which used different methodologies, are consistent with these results, as are the findings that jurors treat social science experts with appropriate caution.

Before concluding, we need to address the methodological limitations and weaknesses of the research studies. Such a discussion is important for two reasons. First, any weaknesses in methodology constrain our ability to extrapolate from the findings. Second, because the empirical investigation of jury behavior is an ongoing process, it is to be hoped that methodological criticisms will spur improvements in future studies.

The most serious limitations of the body of literature reviewed in this article is that the vast bulk of the studies involve simulations. The results, therefore, might have limited applicability to the conditions under which real juries operate.\textsuperscript{183} For example, simulating juries typically are faced with an abbreviated trial situation that is less rich than that experienced by real jurors. Instead of watching live witnesses, lawyers, and judges, the jurors read a transcript, listen to an audiotape, or view a videotape. Further, their verdicts are hypothetical; the decisions have no real consequences.

Lawyers frequently criticize jury research on these grounds, and to some extent the criticism is valid. Yet, in this instance a spirited defense can be made of the research literature. First, we are not dealing with just one or two simulation studies but a large number, varied in substantive topics and involving diverse subject populations and methodologies. Despite the heterogeneity of the studies, consistent trends emerged. Second, the results of the studies that involved actual jurors were generally consistent with the results of the simulated jury studies. Third, the findings are consistent with more general bodies of social psychological theory and research about human decisionmaking processes. This last point is particularly germane, since the assertions that judges and other legal scholars have made about jury behavior are, after all, hypotheses about psychological mechanisms underlying juror decisionmaking. Many of the simulation studies address these posited psychological mechanisms. In the absence of the ability to observe actual jury deliberations, simulation studies provide the next best approximation.

While we would subscribe to the defense of the literature and the conclusions that we have drawn from it, some important weaknesses must be noted. There have been no attempts to interview real jurors who have served in trials involving matters where social framework evidence might be relevant. Particularly when expert evidence has been tendered, the reactions and observations of jurors could provide much insight about the impact of such testimony. In most of the simulation studies the jurors have been presented

with minimal trial evidence; more realistic materials could have been developed. There has been an unwarranted emphasis on the responses of individual jurors when the actual focus of concern is how twelve (or six) men and women pool their insights to reach a decision. Both legal and psychological hypotheses ultimately focus on the deliberation process. Yet most empirical researchers have only made inferences about the deliberations from the variables manipulated in the experiments and from the resulting verdicts. Finally, the empirical research has often not directly addressed the various legal concerns that we attempted to tease out and articulate in Part III of this article. As further efforts are made to understand how juries respond to social framework evidence, it is to be hoped that these deficiencies will be corrected.

In this review, our task has been confined to the question of how juries are likely to respond to expert evidence of the sort that Walker and Monahan have labelled social framework evidence. We have largely avoided the issue of whether the scientific foundation of the evidence is solid. We have also avoided policy considerations, other than jury competence, that judges may consider in weighing its admissibility. Although we have drawn on research literature assessing jury reactions to other types of expert evidence, our conclusions should be seen as pertaining only to social framework testimony. Also, we need to be cautious about generalizing beyond expert testimony involving battered women, rape victims, and eyewitness reliability. Nevertheless, our overall judgment from the body of literature that currently exists is that expert testimony about social frameworks can be helpful to juries. The data also indicate that jurors' use of the evidence is largely confined to the legally appropriate functions it is intended to serve.

184. See R. Hastie, S. Penrod & N. Pennington, supra note 73, at 53; Ellsworth, Are Twelve Heads Better than One?, LAW & CONTEMP. PROBS., Autumn 1989, at 205; Holstein, Jurors' Interpretations and Jury Decision Making, 9 LAW & HUM. BEHAV. 83, 84 (1985); Wells, supra note 60.