

August, 2005

Interdisciplinary Trends in Evidence Scholarship

Roger C. Park* & Michael J. Saks**

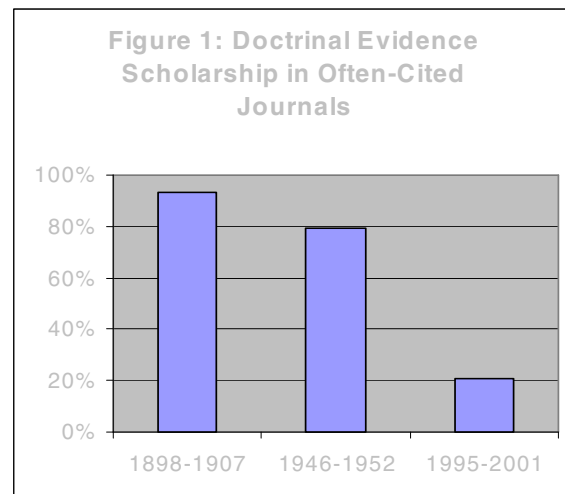
In recent decades, evidence law scholarship has taken on what appears to be a permanent interdisciplinary aspect. As Figure 1 suggests rather dramatically, doctrinal scholarship on evidence is in steep decline, having been replaced by inquiries of other kinds.¹ In pondering the wisdom and purpose of evidence doctrine, and in formulating rules, jurists and scholars have, no doubt, always thought about background facts and theories which today, certainly, can be recognized as falling within the domains of fields outside of law. In recent times the links to those other fields have become more explicit and more numerous. We discuss where these connections have come from and suggest where they might be headed.

* James Edgar Hervey Professor of Law, University of California, Hastings College of the Law; J.D., Harvard Law School. We wish to give special thanks to Craig Callen, David Kaye, Aviva Orenstein, and Joan Williams for their comments.

** Professor of Law and Professor of Psychology, Arizona State University; Ph.D., Ohio State University; M.S.L., Yale Law School.

¹ A “doctrinal” article is one that describes rules of law and synthesizes them. It may also suggest improvements or reforms. Its use of information from other disciplines is ancillary. The first author sampled evidence articles from three periods approximately 50 years apart to assess the relative frequency of doctrinal articles on evidence versus other kinds of evidence scholarship. To reduce the number of articles to be read, the inquiry was limited to often-cited American law reviews that were published under the same name in all three periods. Examination of citation studies conducted in 1930 (Scott Finet, *The Most Frequently Cited Law Reviews and Legal Periodicals*, 9 LEGAL REFERENCE SERVICES QUARTERLY 227, 229 Table 1 (1989), presenting material from Maggs, *Concerning the Extent to which the Law Review Contributes to the Development of the Law*, 3 S.CAL.L.REV. 181 (1930)) and 1996 (James Lindgren and Daniel Seltzer, *The Most Prolific Law Professors and Faculties*, 71 CHICAGO-KENT L. REV. 781, 789 Table 2(1996)) disclosed 10 law journals that appeared among the top twenty in both studies. Those were the Harvard Law Review, Yale Law Journal, Michigan Law Review, Columbia Law Review, University of Pennsylvania Law Review, Virginia Law Review, California Law Review, New York University Law Review, Cornell Law Quarterly/Law Review, and the Minnesota Law Review. Research assistants compiled a list of evidence articles from these journals at approximately 50 year intervals. The exact dates were chosen for index-searching convenience. A search for evidence articles was made in the INDEX TO LEGAL PERIODICALS and the tables of contents of the listed journals. The articles were read and categorized as doctrinal treatments of evidence law or other types of evidence scholarship. (Some articles were removed from consideration after they were read and found not to be on the subject of evidence law.)

In this article we review five of those interdisciplinary junctions. Three of them seem to be established, enduring intersections of evidence scholarship and other fields: psychology (Part I), forensic science (and perhaps science more generally) (Part II), and the “new evidence scholarship” (the application of formal understandings of probability and proof) (Part III).² And we look at two important newcomers to the world of interdisciplinary evidence scholarship: feminist studies (Part IV) and economics (Part V).



I. Psychology and Evidence

For obvious reasons, psychology is the most important of the

interdisciplinary threads that can be woven with evidence law scholarship. Evidence law is much concerned with the abilities of witnesses to perceive, to remember, and to report what they have observed; and with the abilities of jurors to comprehend, evaluate, and draw inferences from the evidence presented to them, including their ability to assess the sincerity of lay witnesses and to understand and not be overwhelmed by expert witnesses. All of these are psychological issues. By psychology we are referring to experimental psychology, cognitive psychology, and social psychology, rather than to clinical psychology. Experimental studies that deal with topics such as memory, perception, judgment, inference, and decisions under

² No doubt it seems ironic to characterize something that is called “new” as being one of the mainstays of interdisciplinary evidence scholarship. But that is the risk of any endeavor that gives itself the name “new” – after a few decades it no longer is.

conditions of uncertainty, and of jury behavior, are plainly relevant to evidence law.³

A. Historical Background

The history of experimental psychology and law has been one of bursts of enthusiasm followed by periods of disenchantment. At present the collaboration appears to be on a steady footing that will make it a permanent part of the landscape of evidence scholarship.

1. Hugo Münsterberg

The first major event in the conjunction of experimental psychology and law was the publication in 1908 of *On the Witness Stand* by Hugo Münsterberg.⁴

Münsterberg had earned his Ph.D. at the University of Leipzig in 1885 and an M.D. two years later at the University of Heidelberg. His mentor, Wilhelm Wundt, is widely recognized as the first true psychologist, the founder of experimental psychology and of the first psychological laboratory, at Leipzig in 1879. As a young scholar at the University of Freiburg, Münsterberg founded its first psychological research laboratory, attracted numerous students, and conducted a series of highly original experiments which attracted considerable attention, quite a bit of it hostile. But Harvard's William James⁵ was far more impressed with Münsterberg's research than with the critics' arguments. He corresponded with Münsterberg and eventually persuaded him to accept a visiting post at Harvard (1892-95) and later to become a permanent member of the faculty (1897-1916). Harvard's plan had been that Münsterberg would take charge of the University's psychology laboratory and become Harvard's leading exponent of the new experimental psychology. That worked for a time, during which he was a successful administrator, teacher and researcher.⁶ But Münsterberg's mind was so energetic and original that he could not remain focused on basic research in experimental psychology and ventured on into philosophy, the problems of theoretical psychology (in which he published well regarded work), and soon founded the meta-

³ These areas of psychology of primary interest to the scholarship of evidence law are the offspring of the marriage of philosophy and experimental physiology which took place in the latter decades of the Nineteenth Century and the early Twentieth. Clinical psychology has different intellectual ancestors. See generally, EDWIN G. BORING, *A HISTORY OF EXPERIMENTAL PSYCHOLOGY* (2nd ed., 1929).

⁴ HUGO MÜNSTERBERG, *ON THE WITNESS STAND* (1925) (hereinafter MÜNSTERBERG).

⁵ An M.D. physiologist whose primary interests were in philosophy, but who along the way became the senior American psychologist.

⁶ See, e.g., the four volumes of Hugo Münsterberg, ed., *HARVARD PSYCHOLOGICAL STUDIES* (1903).

field of applied psychology⁷ – exploring applications of psychology to industry,⁸ education, medicine, psychotherapy, business, and, of course, the law. The range of his interests, ideas, research and insights, is remarkable.

One indication of Münsterberg's professional stature might be seen in his becoming the seventh president of the American Psychological Association in 1898.⁹ Later, Münsterberg became well known to the larger society, writing for popular magazines and taking stands on issues of the day. He was an early twentieth century media figure, widely known and much quoted, arguably the best known psychologist in America.¹⁰

On the other hand, Münsterberg was arrogant and sometimes abrasive. Soon after meeting Münsterberg, William James wrote that he was “vain” and “loquacious,” “desiring to please and to shine,” with a certain “superficiality in his cleverness” but “a

⁷ Boring, *supra* note ___, at 428.

⁸ Münsterberg's work in industrial and organizational psychology was highly experimentally based. He looked at problems of monotony, attention and fatigue, physical and social influences on working ability, the effects of advertising, and helped set the stage for the future development of economic psychology.

⁹ William James had been its third president four years earlier. Ernest R. Hilgard, ed., *AMERICAN PSYCHOLOGY IN HISTORICAL PERSPECTIVE*, 21 (1978).

¹⁰ Ludy T. Benjamin, Jr., Hugo Munsterberg: Portrait of an Applied Psychologist, 4 Portraits of Pioneers in Psychology ch. 7 (Gregory A. Kimble et al. eds 2000), citing M Hale, Jr., *Human Science and the Social Order: Hugo Munsterberg and the Origins of Applied Psychology* (1980). His untimely death came in 1916. “Münsterberg died during the First World War, broken in spirit by the shattering of his dreams of rapprochement between Germany and America and by the hostility of Americans to him in that period of fear of Germany.” BORING, at 428-429. Also see, William Stern, *Hugo Munsterberg: In Memoriam*. 1 J. APPLIED PSYCHOL. 186 (1917):

For many years, Münsterberg considered it his special problem to strengthen the relations between his first and his second homelands. His books on America and the Americans, the founding of the American Institute in Berlin, his proposal to shape the Hamburg institution of higher learning as an embodiment of American University ideals.... as well as many other things, testify to this. He did not always find sympathy in these attempts either here or there.

But at the beginning of the world war, his attitude became completely unambiguous and unconditional. He realized that he was a German and regardless of consequences he began at once a spiritual war against the traditional English sentiment at Harvard. What this meant to him is shown in the following portion of a letter sent to me in February, 1916: “... Of course almost all of my old relations are severed, especially here in Boston. Most of my friends here no longer recognize me: I have been thrown out of clubs and academies. All their rage has concentrated upon me. ...”

man of big ideas in all directions, a real genius.”¹¹ Münsterberg’s arrogance and self-confidence, as well as his brilliance, came through in his book, *On the Witness Stand*, published when he was 45 years old. Münsterberg not only was the first to discover psychology’s relevance to the law, his insights foreshadowed much later work on various issues, including repressed memories, false confessions, and the sources of eyewitness error. At the same time, he scolded the legal profession for being “obdurate” in its disregard for psychological insights that would be valuable in the courtroom. He then proceeded to set forth a description of those insights.

The first chapters of Münsterberg’s book deal with problems of memory and perception. They are written somewhat casually, suggesting the book was intended for an audience of lawyers and other lay readers. For example, the results of experiments conducted in class, where students were asked to estimate quantities such as time and speed, are given as ranges (e.g., “the answers varied between half a second and sixty seconds, a good number judging 45 seconds as the right time”¹²) without giving any measure of central tendency or variance (no mean or median, no standard deviation). However, his basic point – that both perception and memory are trickier and more faulty than lay people think – is consistent with modern psychological research. In fact, one psychologist writing recently has noted:

The chapter on the fallibility of recall in witnesses is particularly well developed; it draws many of the same conclusions that are part of contemporary psychology’s admonitions about the inaccuracy of eyewitness testimony. In fact, these contemporary accounts often cite Münsterberg’s writings as anticipations of this modern work.¹³

Chapters 3 and 4 discuss crime detection, more specifically lie detection by monitoring matters such as associations and emotions. Always confident, Münsterberg believed he could detect bad acts by word association, presenting the suspected wrongdoer with words that are neutral and words that have something to do with the crime, and then measuring response times. Hesitation, unusual associations, and changes in associations on repetition of the experiment will expose the wrongdoer. Münsterberg noted that emotions can also be monitored by measuring respiration, blood pressure, perspiration and pulse, naming the instruments but not using the word

¹¹ Ludy T. Benjamin, *supra*, citing B. Kuklick, *The Rise of American Philosophy*: Cambridge, Massachusetts, 1860-1930, at 198 (1977) .

¹² MÜNSTERBERG at 22. The correct time was 22 seconds.

¹³ Ludy T. Benjamin, *supra* n. ____

polygraph, and says that this approach has promise.¹⁴ He says it is not yet ready for use in crime detection because “experiment gives us so far no sufficient hold for the discrimination of the guilty conscience and the emotional excitement of the innocent.”¹⁵ He suggests a form of guilty knowledge test, where hidden facts about the crime will cause excitement only in someone (notably, the perpetrator) who knows which details are connected to the crime.¹⁶

Chapter 5 contains speculation about untrue confessions, including case histories and references to historical events such as the Salem witch trials. Münsterberg offers reasons why confessions may be untrue, from melancholia to plea bargains to the third degree. Though casually written by contemporary standards, some of the points he makes are consistent with findings of modern research.¹⁷ Other parts seem very speculative, such as his theory that flashing lights cause suggestiveness and increase the risk of false confessions.

Chapter 6, “Suggestions in Court,” makes some points about human suggestiveness that are well-grounded. Münsterberg casually describes an experiment in which suggestive questions planted incorrect information in subjects,¹⁸ and another in which the size of boxes influenced ideas about their weight,¹⁹ reminiscent of later work by Elizabeth Loftus.²⁰ However, his idea that cross-examination using leading questions brings incorrect answers reveals naivete about the adversarial context.²¹

¹⁴ One of Münsterberg’s disciples, William Moulton Marston, was later to claim credit for invention of the polygraph. Marston, a colorful character himself, also created the cartoon character Wonder Woman. Ludy T. Benjamin, *supra* n. ____

¹⁵ MÜNSTERBERG at 132.

¹⁶ *Id.* at 132. Concerning the guilty knowledge test, see, DAVID L. FAIGMAN, DAVID H. KAYE, MICHAEL J. SAKS & JOSEPH SANDERS, MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY, *Polygraph Test: Guilty Knowledge Tests*, § 19-2.1.2[4] (2002).

¹⁷ See, Richard A. Leo & Richard J. Ofshe, *The Consequences of False Confessions: Deprivations of Liberty and Miscarriages of Justice in the Age of Psychological Interrogation*, 88 J. CRIM. L. & CRIMINOLOGY 429 (1998).

¹⁸ MÜNSTERBERG at 181.

¹⁹ *Id.* at 186.

²⁰ E.g., Elizabeth Loftus et al., *Semantic Integration of Verbal Information into a Visual Memory*, 4 J. EXPERIMENTAL SOC. PSYCHOL. 19 (1978).

²¹ MÜNSTERBERG at 183.

Chapter 7, “Hypnotism and Crime,” makes broad claims for hypnotism. Münsterberg does not believe that strangers can hypnotize people, or that a hypnotist could get someone to murder someone through hypnotic suggestion. He does believe, however, that a hypnotist could get someone, through post-hypnotic suggestion, to lie on the stand or even to bequeath all their money and then commit suicide.²²

In the final chapter, “The Prevention of Crime,” Münsterberg says there is no single criminal type and sneers at phrenology. Criminals are generally less intelligent, he writes, but this can lead to many things. Environmental factors are important. We should try positive reinforcement, not punishment, a point he supports by describing an experiment in which subjects who had been asked to lift weights until their fingers became exhausted did not respond to verbal attacks aimed at getting them to lift more, but did get a second wind if the investigator modeled the finger lifting by lifting things with his own finger. Generalizing that to the correction of criminal conduct seems an inordinate stretch. He thought “penitentiaries” should be abolished, but also that there should be a way to take career criminals out of circulation.

2. Wigmore’s Response to Münsterberg

Münsterberg’s book soon attracted the attention of an opinionated and ferociously industrious 45 year old law professor, John Henry Wigmore. In 1909 Wigmore published an article entitled “Professor Muensterberg and the Psychology of Testimony,” containing a root and branch denunciation of Münsterberg’s book.²³

Wigmore chose a peculiar format for his response to Münsterberg. He created a fictional case, *Cokestone v. Muensterberg*, brought in Windyville, Illiana. It is hard to see how the article could have interested psychologists. It begins like a law school exam, silly names and all. It laboriously states an imaginary cause of action for libel against the legal profession and then, of all things, discusses the imaginary plaintiffs’ standing and the question whether the imaginary state had personal jurisdiction!²⁴ One might think that the purpose was humor, but the only attempt at it comes in the discussion of jurisdiction, and there it misses the mark.

Even when Wigmore approaches the point, he starts by laboriously demonstrating that the sources relied upon by Münsterberg were not generally available in English in publications accessible to American lawyers, and hence that it was unfair

²² *Id.* at 224-225.

²³ John H. Wigmore, *Professor Münsterberg and the Psychology of Testimony*, 3 Illinois L. Rev. 399 (1909) (hereinafter Wigmore’s Rebuttal).

²⁴ *Id.* at 399-403.

for Münsterberg to label lawyers as “obdurate.”²⁵ This is a matter of no lasting interest, since it does not address Münsterberg’s points about how psychology could be useful to lawyers and jurists, but only shows Münsterberg to have been unjustified in accusing the legal professions of being stubborn and ignorant.

When Wigmore finally does get to the issue of whether experimental psychology can help law, he causes Münsterberg to admit on the witness stand, mostly with yes or no answers, that he is in error or that his suggestions are impractical.²⁶ Where he is not admitting that he was wrong, Münsterberg admits that he was trite – judges already know about problems of memory and perception, as evidenced by judicial opinions from which Wigmore’s interrogator quotes.²⁷ At the end of the trial the jurors, after a few moments whispering, announce that they do not need to retire, and render a verdict for Wigmore’s side right then and there.²⁸ Perry Mason could not have done better.

Our evidence forefather Wigmore was quite a singular person. Who else would write an article attacking a Harvard²⁹ Professor of Psychology, a man whom William James had described as a genius,³⁰ and have the effrontery to have his adversary confess he was wrong in the attacker’s own article? Who else would purport, without apparent fear of contradiction, to have read every relevant article published in the academic psychology journals in English, German, French and Italian.³¹ Who else, for that matter, would think to have the psychologist admit on the stand that, unlike Wigmore, he was not familiar with the psychology articles published in any of the Italian journals³² or in the German legal journals.³³

²⁵ *Id.* at 407-12.

²⁶ *Id.* at 430-31.

²⁷ *Id.* at 418.

²⁸ *Id.* at 433.

²⁹ The same institution where Wigmore had earned his B.A., M.A., and in 1887 his LL.B.

³⁰ Ludy T. Benjamin, *supra* n. ____

³¹ According to his biographer, Wigmore was literate in a dozen languages, including Russian, Japanese and Arabic as well as all the major European languages. WILLIAM R. ROALFE, JOHN HENRY WIGMORE: SCHOLAR AND REFORMER 106 (1977). Münsterberg was fluent in at least three: German, English and French.

³² *See* Wigmore’s Rebuttal at 406.

³³ *Id.* at 405.

As for Münsterberg's points about the fallibility of memory and perception, Wigmore seeks to show that lay people are familiar with those points by quoting judges who have said the same thing in opinions. This is not a very convincing answer even for the judges, and certainly does nothing to show that jurors are familiar with these problems. Münsterberg's suspicions about lay ignorance seem consistent with later research showing that lay people overestimate the accuracy of eyewitness testimony.³⁴

When Wigmore gets to the merits of the association method, he effectively demolishes it, at least in its reported applications.³⁵ Examining one application, in exposing a fugitive who claimed to be a bookmaker as really being a lawyer, he showed the associations that would be appropriate for a bookkeeper were ignored by the observer, and those that seemed consistent with a lawyer were overinterpreted.³⁶ Wigmore also demonstrated that the reaction time betrayals were interpreted arbitrarily, with some seen as betrayals and others ignored.³⁷

Here are two of Wigmore's final questions on the subject:

Q. Now, do I exaggerate in saying that this whimsicality and arbitrariness of interpretation are constantly to be seen in the records of these experiments? For example, in Jung's own primal experiment, by which he is said to have detected an actual thief, did he not find a so-called "betrayal" in the reaction "stranger-look," by interpreting it thus: "The young thief thought that some one had looked when he was stealing, and had informed on him, so that I, a stranger, now knew of it?" A. Yes.³⁸

Q. Does that seem as amusing to you as it does to me? A. I do not know.³⁹

³⁴ ELIZABETH LOFTUS, EYEWITNESS TESTIMONY (1979); Gary Wells et al., *Accuracy, Confidence and Juror Perceptions in Eyewitness Identification*, 64 J. APPLIED PSYCHOL. 440 (1979).

³⁵ Wigmore's Rebuttal at 427-31.

³⁶ *Id.* at 428-29.

³⁷ *Id.* at 428-30.

³⁸ *Id.* at 429.

³⁹ *Id.* Almost every question by Wigmore's alter ego would have been objectionable – this one as argumentative, others as compound, assuming facts not in evidence, lawyer testimony, and blatantly summing up in the middle of cross-examination.

Wigmore quotes a psychologist who says that most eyewitness identification is accurate,⁴⁰ an assertion that is hard to either verify or disprove, but in any event is of little relevance: if a substantial minority of eyewitness identifications are in error, that creates a serious problem for any system which relies on such evidence. One thing that we know today is that erroneous eyewitness identifications are the leading cause of wrongful convictions,⁴¹ and that psychology has contributed much to an understanding of the conditions that lead to identification errors and to the development of procedures for reducing the risk of such errors.⁴²

Apparently Münsterberg never replied to Wigmore's attack. Our search for a response turned up nothing.⁴³ Most likely, Münsterberg either never learned of Wigmore's attack or was too busy with other projects and concerns.⁴⁴

For his part, Wigmore was obviously interested in psychology himself and he had some good things to say about it. He was receptive to experiments that used trial simulations.⁴⁵ He also upbraided the legal profession for not taking into account the literature about crime from psychology and other disciplines.⁴⁶ Here, he said, the legal

⁴⁰ *Id.* at 427 (quoting Guenther).

⁴¹ EDWARD CONNORS ET AL., U.S. DEP'T OF JUSTICE, CONVICTED BY JURIES, EXONERATED BY SCIENCE: CASE STUDIES IN THE USE OF DNA EVIDENCE TO ESTABLISH INNOCENCE AFTER TRIAL (1996); JIM DWYER, PETER NEUFELD & BARRY SCHECK, ACTUAL INNOCENCE: FIVE DAYS TO EXECUTION AND OTHER DISPATCHES FROM THE WRONGLY CONVICTED (2000).

⁴² The United States Department of Justice has distributed guidelines to all police agencies in the nation, specifying eyewitness identification procedures – based heavily on psychological research – designed to reduce the incidence of erroneous identifications. U.S. DEP'T OF JUSTICE, EYEWITNESS EVIDENCE: A GUIDE FOR LAW ENFORCEMENT (1999).

⁴³ A psychology scholar who has read Münsterberg's letters noted no reference to Wigmore, email communication from Ludy T. Benjamin to Roger C. Park, May 6, 2002, nor could we find a record of a reply with online research and emails asking psychology-law scholars whether they knew of any.

⁴⁴ The same year that Wigmore's article appeared, Münsterberg published two more books, one on values and one on the psychology of teaching. THE ETERNAL VALUES (1909) and PSYCHOLOGY AND THE TEACHER (1909). And the year after that he was appointed an exchange professor from Harvard to the University of Berlin, sent on a quasi-diplomatic mission to establish an American Institute. By then the stage was being set for the outbreak of the First World War, and Münsterberg was trying vainly to reverse the momentum by promoting cultural ties between his two homelands.

⁴⁵ Wigmore's Rebuttal at 426.

⁴⁶ *Id.* at 433.

profession deserved the censure that Münsterberg heaped on it in other contexts.⁴⁷ Later, he appointed a psychologist to be the editor of the *Journal of Crime and Criminology*.⁴⁸

Except for his attack on Münsterberg, Wigmore has almost nothing unkind to say about any scientific claims, however lacking they might be in systematic empirical grounding or however speculative and exaggerated. Perhaps the most remarkable contrast is provided by Wigmore's support of the admission of handwriting experts in court, a cause on behalf of which he waged a campaign for decades, despite the complete lack of any testing of the validity of their claims.⁴⁹ The difference cannot be that Wigmore's handwriting examiner friend, Albert S. Osborn, confined himself and his field to one small corner of the evidence universe. Osborn published a number of books in which he expounded on his speculations about inference, proof, trial practice, and

⁴⁷ *Id.*

⁴⁸ RAULFE, *supra* note ___, at 61.

⁴⁹ As observed by D. Michael Risinger et al., *Exorcism of Ignorance as a Proxy for Rational Knowledge: The Case of Handwriting Identification* "Expertise," 137 U. PA. L. REV. 731, 768 (1989): Though not incapable of a burst of energetic and telling skepticism, [Wigmore] seems to have been willing to accept many claims uncritically and at face value. This certainly seems to have been the case with Osborn's "scientific" approach to handwriting identification. [Wigmore] advocated its acceptance by all courts in his introduction to the first edition of Osborn's book in 1910 and continued to do so faithfully for the next thirty years. Together their advocacy brought "scientific" handwriting identification from a phenomenon barely tolerated by courts to a recognized source of useful and dependable information, despite the absence of a shred of empirical evidence of anyone's ability to do what Osborn claimed he and others of his trainees and followers could do.

Despite the absence of data and handwriting examiners' aversion to research testing their claims, in his treatise on evidence law Wigmore wrote of asserted handwriting identification expertise:

It ought to be now well understood that the identification and the decipherment of documents, including handwriting and all other features, are the subject of scientific study, made by the aid of instruments of precision, and that modern research has elevated the whole subject into the realm of an applied science. Under such conditions, tangible reasons can be given for every opinion, precisely as an engineer can explain the reasons for constructing a bridge of a certain type. A qualified expert's opinion may therefore now be tested and judged by the reasons on which it is based. Such an opinion is as much more valuable than the untrained layman's opinion as that of a scientific and experienced geologist surpasses that of the empiric oil-pro prospector.

JOHN WIGMORE, 7 A TREATISE ON THE ANGLO-AMERICAN SYSTEM OF EVIDENCE AT TRIALS AT COMMON LAW § 1998, at 190 (3d ed. 1938) (citations to two of Osborn's books omitted). Wigmore also roundly attacked judicial opinions which, as late as the 1930s, continued to be skeptical of and to disallow the opinions of asserted handwriting experts. *See*, Risinger et al., at 768-769 n. 174.

related matters, all of which were praised by Wigmore.⁵⁰ Whatever lack of thoroughness in testing and insufficiently cautious extrapolations Münsterberg indulged in, they pale next to Osborn's, which stand on nothing but anecdote and *ipse dixit*.

What, then, accounts for the special abuse Wigmore heaped on Münsterberg? The difference may be that while others showed respect for the intellectual structure of the law (for which Wigmore reciprocated by showing respect for whatever those others were peddling), Münsterberg had the impudence to criticize the legal factfinding process as suffering from fundamental flaws (to which Wigmore also reciprocated).⁵¹

In any event, the lengths to which Wigmore went in putting painstaking erudition to the service of vitriolic attack must, to the extent that it was known to them, have had a chilling effect on psychologists less bold than Münsterberg. Wigmore certainly did the discipline of psychology and evidence law no favors with his Illinois article. And when a bright young Yale law professor, Robert M. Hutchins, who had just begun to teach evidence for the first time, delivered a paper on evidence at a 1926 meeting of the Association of American Law Schools, at a round table on "Psychology, Deception Tests, and the Law of Evidence," Wigmore found the paper highly objectionable. Wigmore immediately wrote to Hutchins warning him away from experimental psychology and telling Hutchins to read the reply to Münsterberg. For good measure, Wigmore wrote a letter to the President of Yale University, James R. Angell, warning him of Hutchins' "extreme behaviorism" and his "irreverence towards judges and recorded experience." Angell sent the letter on to Hutchins, with a note saying, "I have the impression that you must have stepped on some of his most sensitive corns."⁵²

⁵⁰ ALBERT S. OSBORN, *THE PROBLEM OF PROOF* (1922); ALBERT S. OSBORN, *THE MIND OF THE JUROR* (1937).

⁵¹ That, at least, is the theory put forward by some commentators. See Risinger et al. *supra* note ___. Another possibility is political. Münsterberg favored rapprochement between the U.S. and Germany as tensions mounted in the period leading to the First World War, and he supported Germany's position against her European adversaries. His public support of Germany in speeches and in the popular press caused him problems with his Harvard colleagues as early as 1907. Wigmore's feelings toward Münsterberg could only have grown worse. Münsterberg continued to support Germany even after war broke out in 1914 and even after the sinking of the *Lusitania* in 1915. None of this could have been very pleasing to Col. Wigmore, the war hawk and patriot.

⁵² John Henry Schlegel, *American Legal Realism and Empirical Social Science: From the Yale Experience*, 28 *BUFF. L. REV.* 459, 474 n. 83 & 480 n. 101 (1979). Wigmore's attempt to undermine Hutchins not only was small-minded and cruel, it almost certainly fell on unreceptive ears. Angell was himself a psychologist, having done graduate work at Harvard under William James, graduating the same year that James first succeeded in luring Münsterberg there. After Harvard, Angell went to Germany for further study, returning to the U.S. to found the University of Minnesota's first psychology laboratory. Angell was the fifteenth president of the American Psychological Association (1906). Angell briefly was president of the University of Chicago (1919-1921), before moving to the presidency of Yale (1921-

Hutchins was not the least deterred. He took on a psychologist as a collaborator and together they wrote prolifically on psychology and evidence law. But others may well have been deterred.

One wonders what the body of psychology and evidence law scholarship might have developed into, and how much sooner, if the initial encounter had not been hobbled by the overreaching of its first important contributor and the overreaction of its first important critic.

3. Hutchins and Slesinger

The next landmark in psychology and evidence law is the Hutchins & Slesinger series in the 1920s.⁵³ Hutchins and Slesinger wrote articles in leading law journals about the pertinence of psychology studies to such evidence questions as the admissibility of spontaneous exclamations, refreshing memory, evidence of statements offered to prove state of mind and consciousness of guilt, marital privileges, and the competency of witnesses to testify.

Hutchins, a recent graduate of Yale Law School who joined its faculty in 1925 and became dean in 1927, while still in his late twenties, brought a psychologist, Donald Slesinger, to the Yale law school faculty to work with him on psychological studies of evidence law. Neil Duxbury, in his book, *Patterns of American Jurisprudence*, credits Hutchins with playing “the crucial” role in setting the scene for the emergence of realism at Yale during the 1930s, and Hutchins certainly shared the realist interest in studying law with the methods of social science.⁵⁴

Hutchins originally intended to do experimental work, but the articles he actually produced were examples of scholarship by consumers of social science, not scholarship by makers of social science. Professor Schlegel’s appraisal is that:

1937). Before becoming a professor at the law school, Hutchins had been secretary of the Yale Corporation, and therefore probably already had a good relationship with President Angell. The year after Wigmore wrote his letter, Hutchins was named dean of the law school at Yale (1927-1929), following which he became the president of the University of Chicago, where he served for two decades.

⁵³ *Some Observations on the Law of Evidence - Consciousness of Guilt*, 6 U. PA. L. REV. 725 (1929); *Some Observations on the Law of Evidence - State of Mind in Issue*, 29 COLUMBIA L. REV. 147 (1929); *Some Observations on the Law of Evidence - State of Mind to Prove an Act*, 38 YALE L. J. 283 (1929); *Some Observations on the Law of Evidence - The Competency of Witnesses*, 37 YALE L. J. 1017 (1928); *Some Observations on the Law of Evidence - Family Relations*, 13 MINN. L. REV. 675 (1929); *Some Observations on the Law of Evidence - Memory*, 41 HARV. L. REV. 860 (1928); *Some Observations on the Law of Evidence - Spontaneous Exclamation*, 28 COLUM. L. REV. 432, 437 (1928).

⁵⁴ DUXBURY, *supra* note ___, at 87-88.

The articles were of a generally high quality, although their effectiveness varied directly with respect to the quality and relevance of the underlying psychological literature: where good quantitative, behavioral studies were available, the articles were crisp and their criticisms effective, where an older, introspective psychology or new Freudian psychology provided the studies, the articles tended to be less well focused and their criticisms weak.⁵⁵

Several of the articles are of continuing interest. For example, in the spontaneous declarations article,⁵⁶ Hutchins and Slesinger describe a study – one that might not get past a present-day human subjects committee – in which the investigator staged a battle in the classroom. Consistent with other psychological research⁵⁷ on the effects of stress and arousal, students who were highly upset were least accurate. The most accurate students were the ones who were slightly stimulated; they wrote better descriptions than the ones left cold.⁵⁸

Interestingly, Hutchins and Slesinger did not recommend the abolition of the excited utterance exception. They said that the excited utterances then admitted by the courts, many of which required not only excitement but also personal involvement by the declarant, were less accurate than other spontaneous declarations excluded by the courts. In a conclusion that is consistent with the modern approach to present sense impressions under the Federal Rules of Evidence, they supported Professor Morgan's suggestion that declarations "closely connected in time" with the stimulus that brought them forth be admitted, whether excited or not and whether or not the declarant was a participant.⁵⁹

⁵⁵ Schlegel, *supra* note __, at 482.

⁵⁶ Hutchins & Slesinger, *Spontaneous Exclamation*, *supra* note __, at 437.

⁵⁷ The long established and well known Yerkes-Dodson law relating arousal to performance. Robert M. Yerkes & John D. Dodson, *The Relation of Strength of Stimulus to Rapidity of Habit-Formation*, 18 J. COMPARATIVE NEUROLOGY & PSYCHOL. 459 (1908).

⁵⁸ Hutchins & Slesinger, *Spontaneous Exclamation*, *supra* note __.

⁵⁹ Incidentally, the Morgan/Hutchins and Slesinger interaction is an example of cross-fertilization by doctrinal and interdisciplinary scholars that we could hope to emulate today. In 1922 Morgan had published an article in the Yale Law Journal making this suggestion. Edmund Morgan, *A Suggested Classification of Utterances Admissible as Res Gestae*, 31 YALE L.J. 229 (1922). It is a classic work of doctrinal scholarship. Morgan looked at the types of out-of-court statements that courts had been receiving as *res gestae* evidence, suggested a functional reclassification of those cases with different labels, and explained why as a matter of policy each of the classes he suggested should be admissible in evidence. Hutchins and Slesinger later found this work of classification, explanation and fireside policy analysis helpful in understanding how courts treat spontaneous exclamations and how the

Hutchins and Slesinger's memory article⁶⁰ offers psychological perspectives on legal doctrines about impeachment of witnesses, refreshing memory, and past recollection recorded. The authors concluded that impeaching witnesses by questioning them on cross-examination about otherwise irrelevant matters – e.g., cross-examining a witness to a contract transaction about the ages of his grandchildren – was psychologically invalid because human memory is highly situational. They argued that, where time and cost justify it, memory tests developed by psychologists would be better.⁶¹ Hutchins and Slesinger concluded that the past-recollection-recorded doctrine was basically sound, but that it should be refashioned to some extent to take psychological factors into account, such as the forgetting curve, which rises quickly and then flattens out. Among the things they recommend taking into account is whether the event was “definitely attended to” by an observer who was “determined to remember everything that happened.” In the absence of those factors, they argued that one should consider the length of time and number of times the event was “exposed,” and the length of time since the last “exposure,” in that way taking advantage of what is known about the forgetting curve. They approved generally of the concept of refreshing recollection, noting that it follows the psychological insight that recognition outlasts recall. But they worried about false memories caused by confusion with post-event information. Their recommendations on refreshing recollection are implicit and hence indeterminate, but they did identify factors that factfinders and perhaps policymakers should consider in determining accuracy.

As dean at Yale, Hutchins encouraged other realist projects in law and social science, undertaken by such faculty members as Charles Clark (on court administration and who later became the principal draftsman of the Federal Rules of Civil Procedure), William O. Douglas (on bankruptcy, later a justice of the U.S. Supreme Court), and Underhill Moore (studies of banking and compliance with traffic laws). Like other realist initiatives in law and social science, the Hutchins and Slesinger work just petered out. They never did experimental work themselves. The whole field of law and any social science was somewhat dormant for a time, suffering from lack of funding during the depression and a redirection of the energies of realist leaders during the depression and the Second World War.⁶²

doctrine might be reformed in light of findings in psychology. *Spontaneous Exclamation*, *supra* note _____. That is as it should be.

⁶⁰ Hutchins & Slesinger, *Memory*, *supra* note ____.

⁶¹ Hutchins & Slesinger, *Memory*, *supra* note ___, at 863, 870.

⁶² Schlegel, *supra* note ___, at 585-86.

Hutchins departed for the University of Chicago in 1929.⁶³ According to some sources, he became an “apostate realist”⁶⁴ who turned his back on empirical work as merely descriptive and who favored a metaphysical search for “first principles of law.”⁶⁵ In 1933, as the 34-year old President of the University of Chicago, he gave an address to the Association of American Law Schools that was later published in the *University of Chicago Law Review* under the title, “The Autobiography of an Ex-Law Student.”⁶⁶ In it, Hutchins said he had looked to psychology for behavioral insights relevant to the law of evidence. He discovered that psychology had dealt with very few of the issues raised by the law. It could tell you that spontaneous exclamations were not accurate, but it could not tell you whether they should be excluded. Psychologists, he said, “did not know enough about juries to tell you that; nor could they suggest any method of finding out enough about juries to give you an answer to that question.”⁶⁷ This opinion by the prominent university president and law professor, offered in an AALS address, must have been discouraging to scholars who were thinking of following in his path. Note, however, that it is not a wholesale condemnation of experimental psychology and law. Like parts of Wigmore’s earlier attack, it can be read as a call for ecological validity – for realistic simulations of trials and realistic studies of jury behavior.

4. Subsequent Developments

The late 1930s and 40s were a dry period, not only for psychology and evidence law but for all forms of behavioral science and law. But the idea that social science research would be helpful to legal scholars was too ingrained in the post-Langdell legal mind to die. Law professors had a positivist, functional approach to law, and if law was man-made and mutable, and it was important to consider its social consequences, then the social and behavioral sciences could not be ignored for long. It was against the temper of the times to avoid trying to use science in law, including social science. Leaders like Holmes and Pound had called for it before the realist era, and even Langdell invoked a scientific paradigm, though his concept of law as science hardly led to interdisciplinary work.

Hutchins himself did not completely turn his back on interdisciplinary work with law and social science. He participated in a curriculum reform at Chicago that resulted

⁶³ DUXBURY, *supra* note __ at 213.

⁶⁴ DUXBURY, *supra* note __, at 90 (citing EDWARD A. PURCELL, JR., *THE CRISIS OF DEMOCRATIC THEORY: SCIENTIFIC NATURALISM AND THE PROBLEM OF VALUE* 152 (1973)).

⁶⁵ DUXBURY, *supra* note __ at 213.

⁶⁶ Robert Hutchins, *The Autobiography of an Ex-Law Student*, 1 CHI. L. REV. 511 (1934).

⁶⁷ *Id.* at 513.

in the inclusion of a course in psychology in the law school curriculum.⁶⁸ In 1951, after Hutchins became an associate director of the Ford Foundation, Ford funded the landmark study of *The American Jury*, conducted at the University of Chicago Law School by Harry Kalven and Hans Zeisel.⁶⁹ The Kalven and Zeisel study, published in 1966, is recognized as one of the brightest lights in the law and social science firmament.⁷⁰

Another signal event was the use (or misuse) of social science data in *Brown v. Board of Education*.⁷¹ Regardless of what one thinks of the Supreme Court's famous footnote 11, it made the possibilities salient. But the revival of psychology and evidence law had to await the 1970s.

In the early 1960s, the "law and society" movement emerged, largely as a collaboration between legal scholars and social scientists (including sociologists, political scientists, anthropologists, psychologists, and historians). The legal scholars wanted to expand law and legal scholarship beyond its doctrinal habits, hoping to incorporate empirical social science methods and knowledge into the law school curriculum and to hitch their own scholarship to methodologies and substantive knowledge which they saw holding much promise for the law.⁷² The social scientists saw in their collaboration with the lawyers an expanded opportunity to contribute to progressive social change. Two charitable foundations were persuaded to make substantial grants to four universities to help advance these interdisciplinary efforts.⁷³ In 1964 the Law & Society Association (L&SA) was formed. As anyone involved with law schools can see, the integration of fields did not take hold, and spread little beyond the original schools, though the L&SA thrives.

The American Psychology-Law Society (AP-LS) was formed in the late 1960s, initially by psychologists interested in mental health law issues.⁷⁴ In a straw vote on

⁶⁸ DUXBURY, *supra* note __ at 90, n. 119 (quoting JOHN W. JOHNSON, *AMERICAN LEGAL CULTURE, 1908-1940* 138 (1981)).

⁶⁹ Schlegel, *supra* note __, at 238-40.

⁷⁰ KALVEN & ZEISEL, *THE AMERICAN JURY* (1966).

⁷¹ PAUL ROSEN, *THE SUPREME COURT AND SOCIAL SCIENCE* 182-188 (1972).

⁷² See Kermit L. Hall, ed., *OXFORD COMPANION TO AMERICAN LAW*, 489 (2002).

⁷³ The Russell Sage and the Walter E. Meyer foundations made grants to Wisconsin, Berkeley, Denver, and Northwestern. *Id.*

⁷⁴ E.g., insanity defense, diminished capacity, competence to stand trial, civil commitment. See, Tom Grisso, *A Developmental History of the American Psychology-Law Society*, 15 L. & HUM. BEHAV.

naming the organization, the name that got the most votes in 1968 was “Society for Forensic Psychology.”⁷⁵ But under the name AP-LS, the new organization's constitution identified its purposes as promoting a range of interests, including the psychological study of law. The organization's more influential founder and first president, Jay Ziskin, was a lawyer as well as a psychologist, who went on to write a book cataloguing the weaknesses of clinical expert testimony.⁷⁶ Though the book alienated many forensic psychologists from Ziskin, it along with AP-LS was a reflection of his desire to see psychology better serve law by promoting improved expert testimony through better scientific research.

Partly due to that conflict, the young organization struggled to gain its footing. But as it approached the end of its first decade it came under the influence of the field's most influential leader, Bruce Sales, who served as its eight president. Sales had doctoral training in psycholinguistics as well as a law degree, and almost singlehandedly created the infrastructure of modern law and psychology. He personified the field's varied embryonic interests, imbued AP-LS with his own boundless energy, founded the field's first journal,⁷⁷ first book series, and first joint JD/Ph.D degree program. He “provided AP-LS with a structure and purpose that was dramatically more coherent and dynamic than anything it had in its earlier stage. He conceptualized a field of law and psychology that incorporated all of its facets....”⁷⁸ He and his students also produced a disproportionate amount of the research that initially filled the journals and that was presented at the conferences.

One of the first topics of major and continuing research has been eyewitness identification accuracy, an initiative of psychologists, not law professors. One of the early modern landmarks on this subject was Elizabeth Loftus's book, *Eyewitness Testimony*, published by the Harvard University Press.⁷⁹

213 (1991).

⁷⁵ *Id.* at 218.

⁷⁶ JAY ZISKIN, *COPING WITH PSYCHIATRIC AND PSYCHOLOGICAL TESTIMONY* (1970).

⁷⁷ *Law and Human Behavior*, whose citation count now rivals that of the top law journals. See Fred R. Shapiro, *The Most-Cited Law Reviews*, 29 J. LEGAL STUD. 389 (2000). *Law and Human Behavior* ranked first in the “specialized” category with 2,364 cites for the period 1987-97. *Id.* at Table 3. In the “general” category, the *Yale Law Journal* was first with 7,121 citations. The count of *Law and Human Behavior* placed it just ahead of the tenth-ranked “general” law review (the *California Law Review*, which had 2,359 cites). *Id.* at Table 2. This is all the more remarkable considering that LHB has, up to the present, not allowed itself to be included in Westlaw or Lexis.

⁷⁸ Grisso, *supra* note ___, at 224.

⁷⁹ ELIZABETH LOFTUS, *EYEWITNESS TESTIMONY* (1979).

B. Three Contemporary Stories of Success and Its Alternatives

1. Eyewitness Identification Research

There are literally hundreds of studies on the subject of eyewitness testimony; one bibliography of eyewitness research lists 2,000 entries, most of them scientific studies.⁸⁰ The body of research, both field studies and laboratory studies, has been growing at an increasing rate.⁸¹ On many questions, the findings show a high degree of convergence. Moreover, the research, unlike that of Münsterberg or that reviewed by Hutchins & Slesinger, shows a high degree of sensitivity to the legal context. The researchers do not stop at identifying factors, such as weapon focus, that create poor witnessing conditions. They also study witnessing within the legal system.

The weakness of eyewitness identification would not be such a concern if jurors gave it proper weight. But are jurors sensitive to witnessing conditions and problems? This topic has had the benefit of a substantial amount of sophisticated research. Do eyewitness experts help jurors understand? That has also been the subject of controlled experiments. Do judges and lawyers understand problems with witnessing conditions? Can judicial instructions help? Is cross-examination and adversarial testing sufficient to alert jurors to the dangers? All of these context issues have been the subject of empirical research by scholars in the field – just what Wigmore and Hutchins were asking for and bemoaning the absence of.

The topic had become so popular that one of us, when editor of *Law & Human Behavior*, wrote an editorial warning psycholegal scholars that the usefulness of their work can be limited by the range of topics that they have chosen to address.⁸² It even received notice on the pages of the *New Yorker*, in an article by Atul Gawande, who struck a tone similar to Münsterberg:

[The] legal profession has conducted no further experiments on the reliability of eyewitness evidence, or on much else, for that matter. Science finds its way to the courthouse in the form of ‘expert testimony’ –

⁸⁰ BRIAN L. CUTLER & STEVEN PENROD, *MISTAKEN IDENTIFICATION: THE EYEWITNESS, PSYCHOLOGY, AND THE LAW* 68 (1995).

⁸¹ *Id.*

⁸² Michael J. Saks, *The Law Does Not Live by Eyewitness Testimony Alone*, 10 *LAW & HUM. BEHAV.* 279 (1986) (“I have received many papers devoted to the study of eyewitness phenomena; more, in fact, than any other category. Yet the subject of eyewitnesses will occupy at the most only a few hours of a law student's academic life; only a fraction of the thousands of annual pages in law reviews; can be only one of a large number of issues a judiciary committee will address itself to in a year's policy making.”)

forensic analysis, ballistics, and so forth. But the law has balked at submitting its methods to scientific inquiry. Meanwhile, researchers working outside the legal establishment have discovered that surprisingly simple changes in legal procedures could substantially reduce misidentification. They suggest how scientific experimentation, which transformed medicine in the last century, could transform the justice system in the next.⁸³

In contrast to a century ago, today research on eyewitness testimony is far more complete, much of it is more carefully related to the legal context, and more legally sophisticated. And this time around, the legal academy has itself been much more receptive to and sophisticated about the research.⁸⁴

Psychologists have studied the effect of eyewitness age, eyewitness sex, sex of the target person, training of eyewitnesses (such as bank tellers) in how to identify, the dubious value of consistency of description and eyewitness confidence, the effects of disguise and weapon focus, the rate of decay of memory, the effects of post-event information (such as seeing mugshots before making a lineup identification), exposure time, distinctiveness of the target, length of retention interval, encoding instructions, biases in lineup structure and composition, the effect of context reinstatement (having the eyewitness do the identification in the same surroundings as the crime), live v. video v. still pictures at exposure and at identification, and the difficulties of cross-racial identification and whether people with friends of another race are better at it than others. Psychologists also have done studies of jurors, finding that juror subjects often overemphasize factors that have only a weak relationship to accuracy, such as witness confidence, and underestimate other factors, such as the witness's age, the effect of disguise, or the distinctiveness of the target person.⁸⁵

What have been the legal consequences of this research? One important area, dealt with in the evidence casebooks, is the question whether expert testimony is admissible to alert jurors to the dangers of eyewitness identifications (pointing out the factors which increase or decrease eyewitness accuracy).⁸⁶ Most courts continue to hold that it is within the discretion of trial judges to exclude this evidence, but there are

⁸³ Atul Gawande, *Under Suspicion*, THE NEW YORKER, Jan. 8, 2001, at 50.

⁸⁴ Thus, Gawande's criticism is unfair, at least as it pertains to eyewitness identification research.

⁸⁵ See CUTLER AND PENROD, *supra* note ___, at 171-212.

⁸⁶ See Jon R. Waltz & Roger C. Park, CASES AND MATERIALS ON EVIDENCE 935-45 (10th Ed. 2004); Christopher B. Mueller and Laird C. Kirkpatrick, EVIDENCE UNDER THE RULES: TEXT, CASES, AND PROBLEMS 599-600(4th Ed. 2000).

exceptions,⁸⁷ and of course there are trial judges who exercise their discretion in favor of admitting it, especially when it has a good fit to the facts of the case and the eyewitness identification is crucial.

More recently, the research has had an even more important effect on the administration of justice. It has suggested policy for the conduct of eyewitness identification procedures in order that the risk of erroneous convictions be minimized, without increasing the risk of erroneous failures to identify. The AP-LS produced a white paper on eyewitness identification by a number of leading eyewitness researchers.⁸⁸ Their recommendations included:

- preventing witnesses from discussing the incident with each other prior to attempting their identifications (by instruction and, when possible, by sequestration)
- selecting lineup foils (or fillers) to resemble the witness's description (rather than to resemble the suspect)
- blind administration of lineups (police officers interacting with the witness should not know who the suspect is)
- instructing the witness that the culprit might or might not be in the lineup (in order to reduce relative judgment⁸⁹)
- sequential lineups or photospreads (whereby witnesses see one lineup member at a time and must declare whether that person is or is not the perpetrator, also designed to reduce relative judgment).
- recording the confidence of the witness immediately following any identification (in view of the high persuasiveness of this relatively unimportant element, and the malleability of confidence after the witness receives confirming or disconfirming post-identification information)

⁸⁷ For leading cases holding exclusion of expert testimony about eyewitness identification to be an abuse of discretion under the circumstances, see *State v. Chapple*, 660 P.2d 1208 (Ariz 1983) and *People v. McDonald*, 690 P.2d 709 (Calif. 1984).

⁸⁸ Gary Wells, Mark Small, Steven Penrod, Roy Malpass, Solomon Fulero, and C.A.E. Brimacombe, *Eyewitness Identification Procedures: Recommendations for Lineups and Photospreads*, 23 L. & HUM. BEHAV. 603 (1998).

⁸⁹ "Relative judgment" is the phenomenon whereby eyewitnesses tend to select the person in the lineup who comes closest to looking like the perpetrator rather than the person they believe is the perpetrator.

The United States Department of Justice has adopted most of these recommendations and advised all police agencies throughout the nation to follow them.⁹⁰ The one that the Justice Department task force did not adopt was blind testing.⁹¹ Others have gone further. The Illinois Governor's Commission on Capital Punishment recommended double-blind lineups and the General Assembly funded a limited program of implementation and study of the procedure.⁹² New Jersey has adopted double-blind lineups on a statewide basis.⁹³

2. Research Relevant to the Character Evidence Ban

The legal literature on character evidence has varied in its use of insights from academic literature on psychology. Professor Uviller produced an influential article on character evidence without expressly relying upon insights from the psychology literature.⁹⁴ Other articles do cite and discuss the literature on psychology, but are relatively cautious in their use of it.⁹⁵ Still others have been more daring, using personality theory to argue that character evidence is worthless or virtually worthless.⁹⁶

⁹⁰ *Supra* note ____.

⁹¹ Police representatives on the DOJ task force felt that a requirement of blind administration reflected a lack of "trust" in officers conducting lineups. They agreed that officers should not give cues to witnesses as to who the officer knew the suspect to be, but that they could accomplish this by willing themselves to behave properly. Interestingly, blind protocols are common in scientific research, where those scientists do not feel at all demeaned to impose blind testing regimes on themselves and see good methodology as the surest path to accurate results. Gary Wells et al., *From the Lab to the Police Station: A Successful Application of Eyewitness Research*, 55 AM. PSYCHOLOGIST 581 (2000).

⁹² Thomas P. Sullivan, Preventing Wrongful Convictions – A Current Report from Illinois, 52 Drake L. Rev. 605, 607 (2004).

⁹³ Gary L. Wells, 7 Psych. Pub. Pol. and L. 791, 791 (2001).

⁹⁴ H. Richard Uviller, *Evidence of Character to Prove Conduct: Illusion, Illogic, and Injustice in the Courtroom*, 130 U.P.A.L. REV. 845 (1982).

⁹⁵ See, e.g., Richard D. Friedman, *Character Impeachment Evidence: Psycho-Bayesian [!?!] Analysis and a Proposed Overhaul*, 38 U.C.L.A. L. REV. 637 (1991); Miguel A. Mendez, *The Law of Evidence and the Search for a Stable Personality*, 45 EMORY L. J. 221 (1996); Miguel A. Mendez, *California's New Law on Character Evidence: Evidence Code Section 352 and the Impact of Recent Psychological Studies*, 31 U.C.L.A. L. REV. 1003, 1041-60 (1984); Roger C. Park, *Character at the Crossroads*, 49 HASTINGS L. J. 717 (1998).

⁹⁶ For legal scholars who have used the psychology literature to support an especially strong view of the uselessness of character evidence, see Teree E. Foster, *Rule 609(a) in the Civil Context: A Recommendation for Reform*, 57 FORDHAM L. REV. 1, 5, 32 (1988); Robert G. Spector, *Rule 609: A Last Plea for Its Withdrawal*, 32 OKLA. L. REV. 334, 351 (1979); David P. Leonard, *The Use of Character to*

The theory that character evidence lacks probative value finds support in a view of personality that sees situational pressures as being more important as a cause of human behavior than are general traits of character. Thus, whether a person is in a hurry to keep an important appointment is likely to be a more powerful determinant of whether he will stop to help someone in distress than what we can find out about that person's general disposition toward self-sacrifice.⁹⁷ Some legal scholars have used "situationist" personality theory to argue that character evidence ought to be broadly excluded.⁹⁸ They have found further support in studies of "fundamental attribution error," studies showing that people tend to attribute too much power to dispositions and too little to situations. For example, when experimental subjects are asked to predict how people will react in certain situations where behavior has been tested in prior experiments, they tend to err on the side of underestimating the power of situations.⁹⁹

Two dangers are present in applying this body of literature to legal issues: First, there is the danger of using psychology that is too little, too selective, and too old.¹⁰⁰ Second, there is the danger of using research that does not generalize to the legal situation. The first flaw is somewhat hard to avoid. We can't expect law professors to be on the cutting edge of research in another field. Still, one can take precautions against the danger of old and selective research. It is tempting to take one's psychology

Prove Conduct: Rationality and Catharsis in the Law of Evidence, 58 U. COLO. L. REV. 1, 26 (1986-87). But see Susan M. Davies, *Evidence of Character to Prove Conduct: A Reassessment of Relevancy*, 27 CRIM. L. BULL. 504, 535-36 (1991) (using personality theory to argue that opinion and specific-acts evidence on character should be admissible during the prosecution's case in chief.)

⁹⁷ See John M. Darley and C. Daniel Batson, "From Jerusalem to Jericho": A Study of Situational and Dispositional Variables in Helping Behavior, 27 J. PERSONALITY AND SOCIAL PSYCH. 100 (1973) (study of effect of being in a hurry on seminarians' behavior in helping strangers).

⁹⁸ See Robert G. Spector, *Rule 609: A Last Plea for its Withdrawal*, 32 OKLA. L. REV. 334, 351 (1979) ("Character evidence used as a basis for predicting human behavior is useless . . . the legal conclusion therefore ought to be that evidence of character has no probative value"); David P. Leonard, *The Use of Character to Prove Conduct: Rationality and Catharsis in the Law of Evidence*, 58 U. COLO. L. REV. 1, 26 (1986-87) ("[A]s currently conceptualized by the law, character evidence often fails the test of logical relevance"); Teree E. Foster, *Rule 609(a) in the Civil Context: a Recommendation for Reform*, 57 FORDHAM L. REV. 1, 32 (1988) ("Social psychology data reflect the conclusion that prior convictions have virtually no probative value as a predictor for determining a witness' in-court veracity"). Cf. Robert G. Lawson, *Credibility and Character: A Different Look at an Interminable Problem*, 50 NOTRE DAME LAWYER 758, 783 (1975).

⁹⁹ For a description of these experiments, see Ross & Nisbett, *supra* note —, at 119-44 (1991).

¹⁰⁰ Sometimes old research is every bit as good, or better, than the most recent studies. Here, we use the elderliness of research as a shorthand for research which has been supplanted by better and more complete research, leading to different overall conclusions.

from law review articles and empirical studies cited in them, but something more needs to be done, even if it is only to check one's conclusions against recent editions of standard college texts on the subject.¹⁰¹ Or, better yet, law professors should follow Hutchins's example and take a psychologist as a co-author.

Having a psychologist co-author does not solve the second challenge, that of making sure that the psychological literature "fits," that is, that it generalizes to the relevant legal context. That requires examination of the underlying literature and use of the basic skill exercised by doctrinal scholars, that of drawing distinctions. Sensing the lack of generalizability does not require going to somebody else's library and digging through poorly indexed psychology journals. One only needs to look at whatever psychology research was cited and think about its applicability to the trial context. For example, the two works most heavily relied upon in law review articles¹⁰² were Hartshorne and May's massive 1928 work¹⁰³ and Walter Mischel's 1968 book, *Personality and Assessment*.¹⁰⁴ These were seminal works, deservedly influential in psychology. In assessing their applicability to issues of evidence law, however, legal scholars need to consider the behavior examined in those works and compare it to the behavior at issue in those situations where the law admits or excludes evidence of character.

Hartshorne and May studied deceptive behavior in thousands of schoolchildren. Under observation, children were given the opportunity to steal coins, to lie, or to cheat on a test.¹⁰⁵ The authors compared children's propensity to engage in different kinds of dishonest or deceptive behavior – for example, cheating on self-graded classroom tests compared to stealing coins or cheating on tests of athletic ability. They found that

¹⁰¹ For an example, see David Crump, *How Should We Treat Character Evidence Offered to Prove Conduct?*, 58 U. COLO. L. REV. 279, 283 n. 10 (1987). Professor Crump reviewed two college textbooks to check the current status of trait theory, noting that, "Both of these books were used as texts in courses offered at the University of California (Davis) in 1987 and were purchased at the campus bookstore by the author of this Essay." Admittedly, following Professor Crump's example would put legal scholars in the unwonted role of paying money for books, but perhaps interlibrary loans or faculty expense accounts could soften the blow.

¹⁰² See articles cited in *supra* notes ____.

¹⁰³ HUGH HARTSHORNE & MARK A. MAY, *STUDIES IN DECEIT* (1928).

¹⁰⁴ WALTER MISCHEL, *PERSONALITY AND ASSESSMENT* (1968).

¹⁰⁵ See Hartshorne & May, *supra* note ____, at 90-103.

correlations between different kinds of dishonest or deceptive behaviors were very modest.¹⁰⁶

Mischel evaluated the accuracy and utility of psychological assessments of character traits and of psychodynamic states (ego strength, defenses, repression, dependency, etc.). Much of his path-breaking book deals with the reliability and validity of personality tests and clinical assessments. Clinical assessments did not stand up well either in terms of interrater reliability or in terms of external measures, such as progress of the patient. Pencil-and-paper personality tests and tests such as the Rorschach test, sentence completion tests, and the Thematic Apperception Test (TAT) did not correlate very highly with each other or with measures of behavior. His views about the lack of utility of trait-state theories is tied to his preference for “social behavioral” therapy, e.g., counter-conditioning and extinction of phobias by small steps during which the patient becomes desensitized to the object of fear, as opposed to psychodynamic analysis based upon the belief that a person’s basic personality can be inferred from cues such as performance on projective tests, Freudian slips, or interpretation of dreams. He found that hypotheses about generalized traits often did not stand up. For example, he noted that some psychodynamic theorists believe that reactions toward authority stemming from attitudes towards parental figures would generalize to attitudes towards superiors in the workplace. However, experimental testing indicates that there was no substantial correlation between attitudes toward fathers and attitudes toward bosses. Similar results were obtained in examining attitudes towards peers and comparing them to attitudes toward siblings. Thus, Mischel concluded that “there was little evidence for generality of attitudes either toward authority or toward peers.”¹⁰⁷ He also expressed doubt about the generality of the psychological construct of an “authoritarian personality.” Psychologists had hypothesized that persons with authoritarian personalities had little tolerance for ambiguity. But tests that measure different types of intolerance of ambiguity had little correlation with each other.¹⁰⁸ He noted similar results with the posited personality trait of “rigidity.” Similarly, pencil and paper anxiety scales correlated substantially with each

¹⁰⁶ See *id.* at 382-83, Tables LXXI and LXXII (“Intercorrelations among deception tests of the same type” ranged from .462 to .836, while “Average correlations between single tests of different techniques” ranged from -.003 to .312, with only one average correlation above .3) The authors stated that, “as we progressively change the situation we progressively lower the correlations between the tests. . . . [W]e interpret these facts to mean that the consistency of the individual is a function of the situation.” *Id.* at 384. Hartshorne and May concluded that, “[t]he results of these studies show that neither deceit nor its opposite, ‘honesty,’ are unified character traits, but rather specific functions of life situations.” *Id.* at 411.

¹⁰⁷ *Id.* at 22.

¹⁰⁸ *Id.* at 28.

other, but not with psychological measures of anxiety.¹⁰⁹ He noted generally that when behaviors are measured by tests with similar formats and content there are substantial correlations but the correlation often does not hold when the same behaviors are measured by different means.¹¹⁰ Though much of his book deals with correlations between psychological tests or between tests and personality ratings by subjects or their acquaintances, sometimes he also considered studies that involved observed behavior. For example, he referred to the Hartshorne and May study as one in which cheating on one test did not have a high correlation with cheating on a different type of test, and lying in classroom situations showed almost no correlation with deception in out-of-classroom situations. He noted that experiments in which childrens' propensity to delay gratification was observed indicated that this behavior was also malleable, and could be influenced by such situational factors as observation by the children of the behavior of adult confederates whose behavior modeled patience.¹¹¹

The question for the law of evidence is whether the findings of the studies reviewed and reported by Hartshorne & May and Mischel – and cited by proponents as the basis for a strict ban on character evidence – generalize to the situation of greatest interest to the law of evidence. One can question the relevance of the studies involving children. To what extent is the moral behavior of children stable enough to think that what is learned from those studies tells us about the behavior of adults?¹¹² Another way in which the populations may differ is that most if not all of the psychological research reviewed involves normal persons, while the accused in a serious criminal case is disproportionately likely to be suffering from some kind of abnormality.¹¹³ Similarly, valid research about the predictive value of ordinary behavior does not necessarily generalize to the prediction of extreme behavior.¹¹⁴

¹⁰⁹ *Id.* at 81.

¹¹⁰ *Id.* at 101.

¹¹¹ *Id.* at 154.

¹¹² This has been itself a research problem of psychology: what factors enable the prediction of a person's adult behavior from his or her behavior as a child? One of the best answers has been temperament (activity level, emotionality, sociability, impulsivity). See ARNOLD H. BUSS & R. PLOMIN, A TEMPERAMENT THEORY OF PERSONALITY DEVELOPMENT (1975); Jerome Kagan, *Temperamental Contributions to Social Behavior*, 44 AMERICAN PSYCHOLOGIST 668 (1989).

¹¹³ For further discussion, see *infra* notes ___-___ and accompanying text.

¹¹⁴ See Roger Park, *Character at the Crossroads*, *supra* note ___ at 737 & n. 57 (1998); LEE ROSS AND RICHARD NISBETT, THE PERSON AND THE SITUATION: PERSPECTIVES OF SOCIAL PSYCHOLOGY 116-17 (1991).

When the legal issue is whether evidence showing propensity for criminal violence should be received, there is an additional problem of “fit”: none of the underlying research in the studies referred to dealt with criminal or violent behavior. And though decades of research have carried the Hartshorne & May findings beyond the lying, cheating, and stealing of schoolchildren to a far wider array of behavior and situations, and though the studies reviewed by Mischel, and their progeny, similarly examined a wide array of behavior, it is not clear that they generalize to violent behavior. Conclusions about the lack of cross-situational consistency of behavior have been based on studies that generally examine nonviolent, noncriminal behavior of normal research participants.¹¹⁵ This is not a criticism of that body of research, but simply the observation that they do not include the behavior of greatest relevance to the law, and a caution about generalizing from those findings to the legal question of the predictive power of character and other-crimes evidence. Some of the studies do deal with aggression, a trait construct that seems somewhat analogous to violent criminal conduct, but the reported research on that trait gives little comfort to those who believe that behavior is highly unstable and situation-dependent.¹¹⁶ At the end of the day, external validity (generalizability) cannot be determined by logic, but only empirically.

Even for the behavior studied, Mischel did not offer extreme claims about the absence of individual differences or the irrelevance of prior behavior. He even looked favorably upon the use of other act evidence for prediction of behavior in certain situations, considering it superior to clinical psychodynamic assessments based on constructs such as “infantile dependency needs” or “passive-aggressive character make-up.” Indeed, one of the important implications that psychologists took from Mischel’s findings was that personality measures are weaker predictors than past behavior is, leading to the aphorism that “the best predictor of future behavior is past

¹¹⁵ For a useful review of these studies, see ROSS & NISBETT, *supra* note ____ (1991).

¹¹⁶ See WALTER MISCHEL, INTRODUCTION TO PERSONALITY 480-81 (5th ed. 1993) (“Aggression is a dimension of behavior on which stable individual differences have been identified beginning in grade school. These differences remain stable even over long periods of time.”) See also LEONARD BERKOWITZ, AGGRESSION: ITS CAUSES, CONSEQUENCES, AND CONTROL 131 (1993); MICHAEL R. GOTTFREDSON & TRAVIS HIRSCHI, A GENERAL THEORY OF CRIME 50 (1990) (“[W]hat is not arguable is that aggressive behavior, however engendered, once established, remains remarkably stable across time, situation, and even generations within a family”); Dan Olweus, *Stability of Aggressive Reaction Patterns in Males: A Review*, 86 PSYCHOLOGICAL BULLETIN 852-75 (1979). The famous Milgram experiments, in which subjects were induced to give seemingly dangerous (but fake) electric shocks to confederates when told to do so by an authority figure, could be adduced as evidence that criminal aggression is highly situational, but these experiments are better regarded as testing obedience to authority than as a test of antisocial aggressiveness. See Michael R. Gottfredson and Travis Hirschi, *A Control Theory Interpretation of Psychological Research on Aggression*, in AGGRESSION AND VIOLENCE: SOCIAL INTERACTIONIST PERSPECTIVES 47, 53 (Richard B. Felson and James T. Tedeschi eds., 1993). (Milgram’s studies “measure compliance, obedience, or...acquiescence,” rather than aggression); ROSS & NISBETT, *supra* note — at 52-58.

behavior.” By focusing on the behavior, as a matter of sheer prediction, it does not matter whether the causes are internal biological or personality factors that we have not yet discovered and measured or whether the causes are external situational ones (and the behavior recurs because similar situations are repeatedly encountered). In this respect, one might find in Michel a trace of support for admission of bad-act evidence that evidence judges would see as character evidence, and for exclusion of expert clinical assessments that judges might see as falling outside the character ban.¹¹⁷

What about the findings of research that *has* focused on individual differences in and predictors of violence and aggression? Some of that research suggests more useful explanatory and predictive factors. For example, the effects of age and gender are well recognized, with a greatly disproportionate amount of violent crime committed by males between 18-21.¹¹⁸ Among prisoners who had been convicted of unprovoked violent acts (as compared to prisoners convicted of nonviolent crimes) the former had higher levels of testosterone,¹¹⁹ and among non-prison populations, boys and men with higher testosterone levels were more likely to respond aggressively to provocation.¹²⁰ When aggression has been rewarded, especially intermittently, it tends to persist.¹²¹ Some men have long records of criminal violence, which has been found to serve one of two major purposes: violence used to command respect among associates and

¹¹⁷ Evidence judges are sometimes more willing to let in bad-act evidence to impeach a witness when it supports an expert psychiatric assessment by a clinician than when it is offered for the naked inference that because the subject previously engaged in bad-act behavior, he is likely to do so again. The expert testimony can be viewed as being about an illness or medical condition, whereas naked bad-act evidence is likely to be viewed as forbidden evidence of character. *See, e.g., United States v. Lindstrom*, 698 F.2d 1154, ___ n. 6 (11th Cir. 1983).

¹¹⁸ *See* JAMES Q. WILSON & RICHARD HERRNSTEIN, *CRIME AND HUMAN NATURE* (1985).

¹¹⁹ James Dabbs, *Testosterone Measurements in Social and Clinical Psychology*, 11 J. SOC. & CLIN. PSYCHOL. 302 (1992); J.M. Dabbs, T.S. Carr, R.L. Frady, J.K. Riad, *Testosterone, Crime, and Misbehavior Among 692 Male Prison Inmates*, 18 PERSONALITY AND INDIVIDUAL DIFFERENCES 627 (1995).

¹²⁰ J. Archer, *The Influence of Testosterone on Human Aggression*, 82 BRIT. J. PSYCHOL. 1 (1991); James Dabbs & R. Morris, *Testosterone, Social Class, and Antisocial Behavior in a Sample of 4,462 Men*, 1 PSYCHOL. SCI. 209 (1990).

¹²¹ *See, e.g.,* G.R. Patterson, R.A. Littman & W. Bricker, *Assertive Behavior in Children: A Step Toward a Theory of Aggression*, 32 MONOGRAPHS OF THE SOCIETY OF RESEARCH IN CHILD DEVELOPMENT 5 (1967); J.F. McCarthy & B.R. Kelly, *Aggression, Performance Variables, and Anger Self-report in Ice Hockey Players*, 99 J. PSYCHOL. 97 (1978); J.F. McCarthy & B.R. Kelly, *Aggressive Behavior and its Effect on Performance over Time in Ice Hockey Athletes: An Archival Study*, 9 INT’L J. SPORT PSYCHOL. 90 (1978).

violence used for more instrumental purposes.¹²² A review of a large body of research on people who participate in violent and other criminal activity found that such conduct clusters among genetically closely related individuals.¹²³

The most voluminous work, which has focused squarely on the question of the predictability of violence, has been in the area of mental illness and violence. Consistent with the findings of Mischel and others, clinicians using psychiatric theories and diagnoses long have had difficulty predicting who among the mentally ill will commit acts of violence and who will not.¹²⁴ That persistent finding has been helpful in leading researchers away from reliance on clinical assessment to study specific risk factors¹²⁵ and to apply those risk factors actuarially. This has increased the power of violence prediction. Mental disorder itself is “a risk factor of modest magnitude for the occurrence of violence.”¹²⁶ The factors most useful for predicting the violence proneness of the mentally disordered are the same as those useful for predicting in the general offender population, namely: criminal history, antisocial personality, substance abuse, and family dysfunction.¹²⁷ Prediction tools using the risk factor approach have greatly improved predictive power, though it remains less than a precise art/science.¹²⁸

¹²² HANS TOCH, *VIOLENT MEN* (1980).

¹²³ WILSON & HERNNSTEIN, *supra* note _____. For researchers, the debate over these findings focuses on whether they reflect biological differences (inherited or otherwise) or whether they are differences in how these closely related individuals were raised in their families of origin. Leon Kamin, *Is Crime in the Genes?*, *SCI. AM.* 22 (January, 1986); Heathcote W. Wales, *Tilting at Crime: The Perils of Eclecticism*, 74 *GEO. L. J.* 481 (1985). The day may be drawing near when extensions of the human genome project will allow far more precise predictions of at least some kinds and causes of violent behavior. *See, e.g.*, Dan Rather, *Criminal Genes*, 60 *Minutes*, 1999 WL 16209084 (program concerning Jeff Landrigan, an Arizona death row prisoner with an extensive criminal record matching that of his biological father, Darrell Hill, on death row in Arkansas, though the son was adopted out of his biological family in infancy). For the issue of the admissibility of past acts, the cause of the behavior matters little; the persistence of the behavior is the relevant finding.

¹²⁴ John Monahan, *Clinical and Actuarial Predictions of Violence*, § 9-2.3, in *MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY* (DAVID L. FAIGMAN, DAVID H. KAYE, MICHAEL J. SAKS & JOSEPH SANDERS, eds., 2002) (“[T]he sober conclusion that clinicians are ‘modestly better than chance’ at predicting violence appears to be becoming the consensus view.”).

¹²⁵ JOHN MONAHAN ET AL., *RETHINKING RISK ASSESSMENT: THE MCARTHUR STUDY OF MENTAL DISORDER AND VIOLENCE* (2001).

¹²⁶ Monahan, *supra* note ___, § 9-2.1.1.

¹²⁷ *Id.*, § 9-2.2.2.

¹²⁸ *Id.*, § 9-2.1.1.

The occurrence of past violent acts are predictive; the more of them and the more serious they are, the higher the predictive accuracy.¹²⁹

Ultimately, what the law needs to know is whether reform will increase the accuracy of factfinder conclusions. The possible reforms run from exclusionary proposals (barring character and other acts evidence) or the more inclusive proposals (making character evidence in some form, as well as other acts evidence, more available). In order to reach such conclusions, the law must learn more about the actual utility of “character” and prior acts in predicting behavior (else it must guess), as well as about how factfinders will process those kinds of information. The research referred to is clearly relevant to those tasks, but also insufficient.

What we have seen so far is that clinical opinions about personality traits may be less useful than practitioners believe as a predictor of future behavior, but that past behavior might be more useful, especially if the violent behavior has been recurrent, if the situation settings for the behavior at issue were similar, and if factfinders can be given realistic, informative, data-based cautions about the predictive power of the evidence.¹³⁰ It is by no means out of the question that a high comparative propensity to engage in violent behavior could have great value, when used in conjunction with incident-specific evidence, in post-dicting whether a person committed a single act of that type of behavior on a particular occasion.

The larger point we have tried to make is not about whether the body of psychological research points toward an expansion or a contraction of the prohibition on character evidence. The larger point is about generalizing and distinguishing studies, about making fair and informed use of studies, and about trying to reach a correct answer to evidence policy questions. In using social science literature, legal scholars should not lose the standards that they would apply in creating first-rate doctrinal scholarship. First-rate doctrinal scholarship is not advocacy scholarship; the authors do not snatch quotations from cases on their side and ignore cases on the other side. They are expected to squarely confront authority on the other side and to argue why it is inapposite or wrongly decided, if they conclude that it is. Secondly, doctrinal scholars are expected to read the authorities that they rely upon with care and determine whether they are applicable or distinguishable. They should do no less when using the psychology authority.

3. Empirical Research on the Hearsay Rule

¹²⁹ See MONAHAN ET AL., RETHINKING RISK ASSESSMENT, *supra* note ____.

¹³⁰ Though we say this with some trepidation, a Bayesian format for presenting the information might be most appropriate.

The first empirical study of the juror use of hearsay published in 1991.¹³¹ Since then there have been at least 26 other studies, most of them original empirical work as opposed to reviews or commentary on empirical work by others.¹³² Unlike the eyewitness studies, the hearsay studies often involve collaboration between law professors and psychologists. Fourteen of these studies were co-authored by law professors, one by a lawyer, and three by William C. Thompson, a psychology professor who has a law degree and who probably spends more time in court than most law professors who teach evidence.¹³³ These close collaborations are entirely understandable. For hearsay research, lawyers and psychologists need each other. The legal setting of eyewitness identification problems is easy for someone with no formal law training to understand. But the legal context in which hearsay issues arise, the rationales of the hearsay rule, and the situations in which out-of-court statements would be admissible in a real trial are things that can be understood only by someone who has carefully studied the concepts.

In essence, what these studies do is to present some mock jurors with the “live” testimony of a witness with personal knowledge and other mock jurors with the same information from a hearsay witness. The research question asked by most of these studies is whether the jurors discount the evidence they acquired through hearsay, reflecting their recognition of its weaknesses compared with the firsthand witness. A few studies have been clever enough to ask whether the jurors are able to make proper use of the hearsay evidence in order to move closer to an accurate conclusion about the underlying events. For example, if the out of court declarant’s statements were made under the influence of suggestive questioning, does that troublesome fact come through to the jurors as well by way of a hearsay witness as it does by examination of the firsthand witness? Some studies created mock cases, the “truth” of which is not only not known but does not exist; others created a set of true underlying facts, and the jurors’ conclusions about the event could be compared to its reality.¹³⁴ Some of the studies

¹³¹ Stephan Landsman & Richard Rakos, *A Preliminary Empirical Enquiry Concerning the Prohibition of Hearsay Evidence in American Courts*, 15 L. & PSYCHOL. REV. 65 (1991).

¹³² For a bibliography, see See Roger C. Park, *Visions of Applying the Scientific Method to the Hearsay Rule*, 2003 Mich. St. DCL L. Rev. 1149, 1171-74 (2003).

¹³³ In addition to arguing appellate cases involving high tech forensic science, he was co-counsel for O.J. Simpson in *State v. Simpson*.

¹³⁴ This reveals a major advantage to simulation studies in this context. In actual trials, rarely can one know what the true underlying facts were. Some of the studies were ingenious enough to begin by having the “hearsay witnesses” observe a videotaped event (the underlying reality that would be at issue) and then be videotaped being interviewed about what they had seen. Those interviews became the “in court” hearsay testimony heard by the mock jurors. The results of such studies can provide insights into the ability of mock jurors to reach accurate conclusions about the underlying events which are otherwise completely unavailable.

presented the “case” by way of brief written transcripts or even briefer summaries, others by audio or videotaped presentations of a mock trials.¹³⁵ In some studies the hearsay witness was a layperson, in others an expert witness. In some the firsthand witness was cross-examined and in others not. And so on.

Put succinctly, some of these studies appear to offer greater verisimilitude than others.¹³⁶ And, perhaps more important, the studies asked different research questions, thereby having more (or less) relevance to the legal policy questions about hearsay that need to be asked and answered. The results, however, cannot be put succinctly. The circumstances of some studies revealed jurors to be quite capable of heavily discounting hearsay testimony as compared to firsthand witness testimony. In other studies the jurors credited the hearsay as much as they did the firsthand testimony. It is not always clear what the right or ideal response should be to the hearsay testimony, in contrast to the testimony of the firsthand witness.

Whatever their methods and findings, these studies are an important start, but certainly no more than a start, in a complex area of evidence law and policy. Reflecting on how they relate to evidence policy, and especially on where they cast light or fall short of casting light, should guide more and better studies in the future. Our reflections, below, are organized in terms of three important issues of hearsay policy.

¹³⁵ There is one criticism of the hearsay experiments that is entirely without merit. Professor Michael Seigel argues that all hearsay experiments in which the trial stimuli were videotaped fail to test any difference between hearsay and non-hearsay conditions. Professor Seigel writes that, “[A]lthough each of the studies purports to examine the ability of human subjects to differentiate and discount hearsay in comparison with witness testimony, they actually measure individuals’ abilities to differentiate among different types of hearsay and hearsay within hearsay. Remarkably, the articles are silent on this critical issue concerning their internal design.” Michael L. Seigel, *A Pragmatic Critique of Modern Evidence Scholarship*, 88 NW. U. L. REV. 995, 1042-43 (1994). As applied to videotaped simulations, this criticism is conceptually barren. The hearsay rule is actually two rules: the important and consequential rule that a witness cannot report another person’s statement if it is used for its truth, and the relatively trivial rule that witnesses testifying to firsthand observations cannot testify by videotape. Because these two rules both share the label “hearsay” instead of being called the “secondhand information rule” and the “videotape rule,” Professor Seigel believes that a trial presented by videotape cannot tell us anything about the processing of secondhand information. He gives no reason for this other than the labels. To be sure, videotaped trials are not real trials, but this challenge to validity is no more serious in the hearsay context than in any other context using videotaped stimuli, such as studies of the effects of jury instructions.

¹³⁶ It is far from clear how necessary a realistic simulation is to getting correct answers that are generalizable to actual trial settings. A comparison of simulations with more and with less verisimilitude found that the results reached did not differ. Brian Bornstein, *The Ecological Validity of Jury Simulations: Is the Jury Still Out?* 23 L. & HUM. BEHAV. 75 (1999). But it almost certainly is the case that lawyers and judges are more persuaded by studies which share many superficial similarities with trial procedures.

a. Whether Hearsay Should be Received When There Is a Choice Between Hearsay and Live Testimony

This first issue of hearsay policy arises when a live firsthand witness is available, but a party would like to offer a hearsay witness instead. The policy-maker must ask whether harm is done by allowing adversaries to substitute hearsay for available live testimony. Most jurists would agree that the principal harm would be the loss of cross-examination.¹³⁷ It is risky to allow adversaries to substitute hearsay evidence for cross-examined evidence. They will substitute hearsay for live testimony when the substitution helps their case. An advocate would choose to forgo the benefits of a vivid live presentation when the advocate does not want to have her witness questioned by the opponent, which is exactly the situation in which cross-examination might turn up information helpful in finding the truth.

The existing empirical studies of hearsay simply do not address this issue. They do not try to study whether adversaries would choose to present inferior evidence in the absence of a hearsay rule, and they do not attempt to examine the value of cross-examination in uncovering new evidence. The experiments focus on the modality of presentation of the evidence, not on whether the trial procedure of cross-examination uncovers useful new information. What the studies test is the impact of the medium, just as in an experiment comparing the effects of live testimony with the effects of closed-circuit TV testimony. There is no attempt to determine whether cross-examination might bring out new facts or reveal collateral facts bearing on the credibility of the witness.

That is not so much a flaw in the studies as it is a limitation on the inferences that can be drawn from them. The hearsay studies that seek to say something about the accuracy of hearsay, as opposed to its impact, do not seek to simulate any kind of cross-examination, much less typical cross-examination.¹³⁸ Thus it is obvious that they are of limited usefulness in assessing the value of hearsay testimony as a substitute for cross-examined testimony.

¹³⁷ The substitution has other consequences, such as precluding physical confrontation between the witness and the accused in criminal cases, allowing evidence to be received from out-of-court declarants who were not under oath or subject to observation for demeanor cues.

¹³⁸ See Margaret Bull Kovera, Roger C. Park, & Steven D. Penrod, *Jurors' Perceptions of Eyewitness and Hearsay Evidence*, 76 MINN L. REV. 703 (1992) and Maithilee K. Pathak & William C. Thompson, *From Child to Witness to Jury: Effects of Suggestion on the Transmission and Evaluation of Hearsay*, 5 PSYCHOLOGY, PUBLIC POLICY, AND LAW 1 (1999).

b. Whether Hearsay Should be Received When There is No Choice

A different issue of hearsay policy is presented when the declarant is not available, so that hearsay testimony is the only way of learning her account of the facts. Here, there is a strong common-sense argument that hearsay is better than nothing at all, that one need not go in the dark because the light is not perfect. But the hearsay rule sometimes bars such testimony.¹³⁹ The empirical question is whether hearsay evidence is so misleading that it ought to be excluded even when live testimony is not an alternative. Here, empirical studies can be helpful even if they do not simulate adversarial incentives to substitute hearsay for live testimony. Nonetheless, it is difficult to use the existing experiments to reach conclusions about this issue.

The legal policy issue is whether hearsay would be helpful to the jury in reaching the ground truth. In most of the experiments, the ground truth was not known to the experimenter, as we noted above. In two of the experiments,¹⁴⁰ the investigators did have accurate knowledge of the ground truth. They controlled and recorded the events witnessed by the hearsay declarant and by the person reporting the content of the hearsay statements, and examined the question whether the juror subjects were able to use the secondhand information to accurately reconstruct what actually happened.

Neither of those experiments simulated cross-examination. Here it is necessary to distinguish between cross-examination of the declarant and cross-examination of the in-court hearsay witness. If the declarant is not available for cross-examination, the legal policy issue does not turn on whether it would be better to hear from a cross-examined declarant or from a hearsay witness, because cross-examining the declarant is not an option. But an experiment would cast clearer light on the legal situation if it involved cross-examination of the hearsay witness. The cross-examiner could ask questions that might undermine the believability of the hearsay witness. At a minimum, these questions make it more apparent to the jurors that the hearsay witness did not see the event, does not know from firsthand knowledge what happened, is relying solely on the declarant, and does not know the witnessing conditions of the declarant (such as whether something was in the way, whether the declarant was paying attention, etc.).

¹³⁹ See *United States v. Day*, 491 F.2d 861 (D.C. Cir. 1978); Roger C. Park, *Trial Objections Handbook* § 4:6 (2d Ed. 2001). The increasing popularity of the concept that a defendant forfeits his hearsay/confrontation objection by silencing the victim could, however, someday lead to free admission of victim statements, at least in cases in which there is enough evidence of homicide to allow the trial judge to find that the prosecution has laid the foundation for forfeiture by showing that the defendant killed the victim. See *People v. Giles*, 19 Cal.Rptr.3d 843 (Ca. App. 2004); *State v. Meeks*, 277 Kan. 609, 88 P.3d 789 (2004) (holding that defendant forfeited the right to confrontation by murdering victim, citing *Richard D. Friedman, Confrontation and the Definition of Chutzpah*, 31 *Israel L.Rev.* 506 (1997)).

¹⁴⁰ *Kovera et al.*, *supra* note __; *Pathak and Thompson*, *supra* note __.

c. Whether Erroneous Admission of Hearsay Should be Deemed Prejudicial Error

Appellate courts often confront the question whether erroneous admission or exclusion of hearsay is prejudicial error. When an evidence error is harmless in the sense that it was quite unlikely to have affected the result, then appellate courts refuse to reverse.¹⁴¹ The hearsay studies have relevance to this policy issue even when they do not address the question whether hearsay aids in reaching accurate verdicts. If, as some of the studies suggest, juries heavily discount hearsay, then the erroneous admission of hearsay (at least where it is substituted for live testimony) should less often be considered to be prejudicial error.

The hearsay experiments reach divergent results, and without further research we cannot be sure why.¹⁴² Based on the features of the extant studies, the reason might be a difference in the medium by which the trial evidence was presented (written versus audio or videotaped). Or it might have been a difference in whether or not the hearsay witnesses were especially credible individuals (e.g., expert witnesses). In some studies, the mode of presentation may have made it particularly hard for jurors to realize they were dealing with hearsay, a problem that is much less likely to occur in a trial. Another possibility has to do with who the out-of-court declarant was. Substituting hearsay adults for child declarants may be more acceptable to the mock jurors than other hearsay because they could readily sympathize with the possible need for it. Because there are many different situations in which hearsay can be offered, and different witnesses through whom it can be offered – parties, police, informants, children, experts, to name a few – it is hard to generalize. Or it might have been due to something else.

Because of divergent results, the wide range of circumstances in which hearsay issues can arise, and differences between the simulations and the actual courtroom environment, caution must be used in deriving broad policy recommendations from the experiments on hearsay impact.

Despite the reservations noted above concerning those three basic hearsay policy issues, the experiments are a good start in answering more narrow questions.

¹⁴¹ See, e.g., ROGER C. PARK, TRIAL OBJECTIONS HANDBOOK (2d ed. 2002) at § 12:2.

¹⁴² It might have been due to a difference in the medium by which the trial evidence was presented (written versus audio or videotaped); it might have been a difference in whether the hearsay witnesses were especially credible individuals (e.g., expert witnesses); it might have been due to something else. In some studies, the mode of presentation might have made it particularly hard for jurors to realize they were dealing with hearsay, a problem that is much less likely to occur in a trial.

For example, the experiment by Warren & Woodhall¹⁴³ suggests a danger of hearsay in the child witness context – that persons interrogating child witnesses do not accurately remember the mode of interrogation, and how suggestive they were or not – while the Pathak and Thomson experiment¹⁴⁴ indicates that when the suggestive interrogation causes a complete change of story, the trier of fact may be able to detect the truth even without realizing the full danger of suggestiveness. Other experiments may be useful to lawyers making strategic decisions, such as whether to substitute a hearsay witness for a child witness. And, as has been the case with studies of eyewitness testimony, the studies are more likely to be useful as they multiply and converge. Studies examining particular hearsay dangers – for example, the danger of suggestive questioning of child witnesses – in particular situations are likely to be more useful than studies that attempt to answer the global question of whether hearsay is good evidence.

But the studies that are available now are just a start. The hearsay problem is complex because there are so many different situations. A study that tests whether jurors (or judges, for that matter) are successful in using hearsay evidence of eyewitness identification will not necessarily tell you anything about whether they would be successful in using hearsay accounts of medical diagnosis or hearsay accounts of declarations by child witnesses. And the purposes of cross-examination can be different in different situations, hence the use of hearsay will differ as an adequate substitute for cross. Hearsay might be an adequate substitute in situations in which the purpose of cross is to show defects in perception, but not where the purpose is to show deception. A hearsay experiment that shows jurors successfully using hearsay from a neutral declarant would not show that the jurors could detect a lying declarant through the medium of a hearsay witness.¹⁴⁵

Consequently, it makes sense to approach the vast problem of hearsay and its exceptions bit by bit, studying a specific situation such as child witnesses and specific dangers such as insensitivity to suggestiveness. To get a good grip on it, there probably will need to be hundreds of studies, as in the eyewitness area. Until then, policy-makers seeking answers to questions such as whether the hearsay rule should be abolished or modified in a particular class of situations will have to continue to rely upon their traditional tools of history, experience, and fireside induction, with occasional help from the empirical research that is available.

¹⁴³ Amye R. Warren and Cara E. Woodall, *The Reliability of Hearsay Testimony: How Well Do Interviewers Recall Their Interviews With Children?*, 5 PSYCH. PUB. POL. AND L. 355 (1999).

¹⁴⁴ Pathak and Thompson, *supra* note ____.

¹⁴⁵ Suppose, for example, that the lawyer plans to cross-examine a lying police officer by getting the police officer committed to the story that the officer had not used the N word in the previous ten years, a story that the cross-examiner knows can be definitively disproved. A study that shows you don't lose anything when you use hearsay instead of a neutral witness would not mean that you do not lose anything when you use hearsay in lieu of cross-examining Detective Fuhrman.

II. Law and Forensic Science

Law and forensic science has long been a topic of evidence scholarship – there have always been articles about new ways of gathering evidence to prove the identity of suspected perpetrators. But the flourishing of this field has been remarkable since the Supreme Court's *Daubert* decision in 1993.¹⁴⁶ *Daubert*'s command to judges to do their own screening for scientific validity rather than merely deferring to self declared experts has encouraged scholarship about whether expertise has been scientifically tested, what the results of testing have been, and how well judges have been doing with the law and the science.

Daubert has already led to an astonishing amount of scholarship¹⁴⁷ and has created a legal environment that will lead to even more scholarly (as well as judicial) discussion of issues concerning the scientific validity of scientific claims and forensic techniques. It quickly gave birth to new treatises, seeking to look at the world of scientific evidence through the framework of *Daubert*.¹⁴⁸ These discussions focus not only on issues of the legal admissibility of scientific evidence, or other limitations that might be placed on it, but *Daubert* and *Kumho Tire*¹⁴⁹ require inquiry into the scientific basis for the evidence. In the post-*Daubert* world, judges, lawyers, and legal scholars cannot do their work properly without becoming scientifically literate.

What lawyers, scholars, and the courts are finding is that some kinds of evidence, most notably some of the forensic sciences, which had been all but unquestioned under older admissibility tests, appeared to have startling weaknesses when viewed through the lens of the new test.¹⁵⁰ What courts should do with this new

¹⁴⁶ *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993).

¹⁴⁷ A search of the Westlaw JLR database (journals and law reviews) on July 26, 2005, found 1,253 articles citing *Kumho Tire v. Carmichael* and 4,918 articles citing *Daubert v. Merrell Dow Pharmaceuticals*.

¹⁴⁸ See, DAVID L. FAIGMAN, DAVID H. KAYE, MICHAEL J. SAKS & JOSEPH SANDERS, MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY (2nd ed. 2002; Supp. 2003); FEDERAL JUDICIAL CENTER, REFERENCE MANUAL ON SCIENTIFIC EVIDENCE (2nd ed. 2000).

¹⁴⁹ *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999).

¹⁵⁰ Michael J. Saks, *Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science*, 49 HASTINGS L. J. 1069 (1998).

realization, what the exact requirements of *Daubert* are, and what the contours of *Daubert*'s exclusionary zone are or should be are the subjects of countless articles.¹⁵¹

Another important development in the area has been the rise and success of DNA evidence. At first DNA evidence was accepted blindly, like many other forensic sciences; then came the *Castro* case – real scientists got involved and criticized the methodology of the DNA labs, successfully.¹⁵² Then the issues of methodology and fit were addressed and cleaned up. It now appears that DNA may, by example, provide benefits to the rest of forensic science, because its more careful and explicit scientific approach puts to shame so much of what had previously been accepted as forensic expertise.¹⁵³ And exoneration by DNA may also help throw light on the flaws of other forensic science, just as it has already reinforced, in the eyes of policy makers, the scientific knowledge about the fallibility of eyewitness identification.¹⁵⁴

There already have been successful challenges to a form of expertise that, for most of the 20th Century, had been regarded as being of settled admissibility – handwriting identification.¹⁵⁵ Even the holy grail of forensic science, fingerprinting, has been challenged. We are not referring to rolled prints used to prove that a person once arrested in Phoenix is the same person who is now on trial in San Francisco, but rather the methods used in comparing latent prints, lifted from crime scenes, with rolled prints. The scientific basis for the claim that identification by latent prints is infallible if established procedures are followed just does not exist.¹⁵⁶ Most federal courts which have confronted this realization responded to it by manipulating the law to permit admission.¹⁵⁷ One federal judge initially responded by limiting admission (excluding the

¹⁵¹ For discussion of specific scientific topics, see the various chapters of Faigman et al., *supra* note ____.

¹⁵² See, e.g., Jennifer L. Mnookin, *Fingerprint Evidence in an Age of DNA Profiling*, 67 Brooklyn L. Rev. 13 (2001).

¹⁵³ Michael J. Saks & Jonathan J. Koehler, *What DNA "Fingerprinting" Can Teach the Law About the Rest of Forensic Science*, 13 CARDOZO L. REV. 361 (1991).

¹⁵⁴ Michael J. Saks & Jonathan J. Koehler, *The Coming Paradigm Shift in Forensic Identification Science*, ____ SCIENCE ____ (forthcoming).

¹⁵⁵ See *supra* notes ____ - ____ and accompanying text.

¹⁵⁶ See Faigman et al., *supra* note ___, *Fingerprint Identification*, § 27 (2002; Supp 2003); SIMON COLE, *SUSPECT IDENTITIES* (2001); Robert Epstein, *Fingerprints Meet Daubert: The Myth of Fingerprint "Science" Is Revealed*, 75 S. CAL. L. REV. 605 (2002).

¹⁵⁷ See FAIGMAN et al., *supra* note ___, *Fingerprint Identification: Post-Daubert Challenges in Federal Courts*, § 27-1.2 (Supp., 2003).

fingerprint expert's conclusion of identity, but allowing testimony concerning the similarities and differences between the latent and known prints),¹⁵⁸ but later reversed himself and went the way of other courts.¹⁵⁹ The issue has even begun to interest the general public.¹⁶⁰

Of course, classifications are always disputable, and some might argue that the field of "law and forensic science" is really too small if it embraces only the traditional forensic sciences, and that a better title would be "law and the scientific method," because people who write in this field also write about other *Daubert* topics, such as the admissibility of epidemiological evidence, psychological syndrome evidence, and so on. Certainly the scientization of law, and of society in general, with or without *Daubert*, will inevitably lead to more evidence scholarship about law and science of all kinds.

III. The "New Evidence Scholarship" – Probability and Proof

The term "New Evidence Scholarship," coined by Richard Lempert,¹⁶¹ is broad enough to could cover all interdisciplinary scholarship or even all innovative scholarship. But the term has most often been applied to scholarship on probability and proof, including evidence scholarship that applies formal tools of probability theory, such as Bayes' theorem.

Formal analysis of legal evidence is not a wholly new idea. Wigmore's *Science of Judicial Proof* used symbols and charts as part of a system of evidence analysis.¹⁶² And there is a long, if intermittent, history of interest in legal uses of statistics and basic probability theory: in the Howland Will case in the 1860s, for example, an eminent

¹⁵⁸ United States v. Llera-Plaza, 2002 WL 27305 (E.D.Pa.); see also, United States v. Crisp, 324 F.3d 261 (4th Cir. 2003) (Michael, J., dissenting) (explicating the shortcomings of fingerprint and handwriting expert evidence).

¹⁵⁹ United States v. Llera-Plaza, 188 F.Supp.2d 549 (2002) (finding, nevertheless, that the "science" of fingerprint identification suffered from major shortcomings).

¹⁶⁰ See, e.g., Michael Specter, *Do Fingerprints Lie?*, NEW YORKER (May 27, 2002), at 96. That article predicts, accurately we think, that the issue is not going to go away.

¹⁶¹ Richard Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof*, 66 B.U.L.REV. 439 (1986)) ("Evidence is being transformed from a field concerned with the articulation of rules to a field concerned with the process of proof. Wigmore's other great work [THE SCIENCE OF JUDICIAL PROOF, cited *infra* n. ___] is being rediscovered, and disciplines outside the law, like mathematics, psychology and philosophy are being plumbed for the guidance they can give.")

¹⁶² John H. Wigmore, THE SCIENCE OF JUDICIAL PROOF AS GIVEN BY LOGIC, PSYCHOLOGY AND GENERAL EXPERIENCE AND ILLUSTRATED IN JUDICIAL TRIALS, (3d Ed. 1937).

scholar used the product rule in an attempt to calculate whether commonalities in two signatures could have been coincidental.¹⁶³

But the word “new” is a fair one, at least in terms of the volume of scholarship. Among the seminal works we can count John Kaplan’s 1969 article on decision theory,¹⁶⁴ Richard Eggleston’s 1978 book entitled, *Evidence, Proof and Probability*, and Jonathan Cohen’s 1977 challenge to the applicability of standard probability reasoning in legal contexts.¹⁶⁵ The 1968 case of *People v. Collins*,¹⁶⁶ in which the prosecution misused evidence of probabilities in an ingenious fashion, stirred up interest in how probabilities might properly be used in trials. The celebrated debate in the *Harvard Law Review*,¹⁶⁷ with Finkelstein and Fairley on one side and Tribe on the other, no doubt inspired other scholars to pursue the topic, even though Tribe’s skepticism toward the practicality of using Bayesian approaches at trial may have also had a dampening effect. William Twining’s calls to action no doubt also helped,¹⁶⁸ as did conferences on new evidence scholarship organized by Peter Tillers.¹⁶⁹ David Schum, a nonlawyer,

¹⁶³ See *The Howland Will Case*, 4 AM. L. REV. 625 (1870) (testimony of Charles Pierce). Other early work included the rigorous and largely ignored efforts of Michael & Adler to trace the logic of proof, see Jerome Michael & Mortimer J. Adler, *The Trial of an Issue of Fact: I*, 34 COLUM. L. REV. 1224 (1934); Jerome Michael & Mortimer J. Adler, *The Trial of an Issue of Fact: II*, 34 COLUM. L. REV. 1462 (1934); Jerome Michael & Mortimer J. Adler, *Real Proof: I*, 5 VAND. L. REV. 344 (1952). William Twining detected a “continuous intellectual tradition from Bentham, through Wills, Best, Stephen, Thayer, Gulson, Wigmore, Michael and Adler to Leo Leven, Irvin Rutter, and contemporary teachers of law” See William Twining, *Taking Facts Seriously*, in William Twining, RETHINKING EVIDENCE: EXPLORATORY ESSAYS 15 (1990; reprint of 1982 essay). Nonetheless, Twining recognized that by the 1970s the field of rigorous analysis of facts seemed “rather neglected” and that there was almost no mention of topics such as epistemology and probability theory in the standard texts on the law of evidence. See William Twining, *Rethinking Evidence*, in RETHINKING EVIDENCE: EXPLORATORY ESSAYS 341-42 (1990).

¹⁶⁴ John Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065 (1968).

¹⁶⁵ L. JONATHAN COHEN, *THE PROBABLE AND THE PROVABLE* 6 (1977)

¹⁶⁶ 438 P.2d 33 (Cal. 1968).

¹⁶⁷ Michael Finkelstein & William Fairley, *A Bayesian Approach to Identification Evidence*, 83 HARV. L. REV. 489 (1970); Laurence Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971).

¹⁶⁸ See William Twining, RETHINKING EVIDENCE: EXPLORATORY ESSAYS *passim* and 12-31, 341-72 (1990).

¹⁶⁹ Professor Tillers was the chair (with Eric Green), moderator and a panelist at the Boston University School of Law Symposium on Probability and Inference in the Law of Evidence, April 4-6, 1986 and the organizer, chair and a panelist at the Cardozo Conference on Decision and Inference in

also deserves much credit for interesting law professors in formal analysis of evidence.¹⁷⁰

One “new evidence” topic revolves around the question whether standard probability logic is or ought to be consistent with judicial fact-finding. Sometimes naked statistical evidence seems intuitively insufficient to justify a judgment. If the only proof that the plaintiff was injured by the defendant’s bus instead of another company’s bus was mere evidence that a majority of the blue busses in town belonged to the defendant, many of us would hesitate to find that identification sufficient. Or if the defendant was chosen at random from a crowd in a rodeo, and was sued on a claim that he had not paid for his ticket, we would hesitate to issue judgment based merely on proof that only 499 tickets had been sold and there were a thousand spectators, so that defendant, chosen at random, had just over a 50% chance of being a gatecrasher. Does the fact that our intuition makes us cringe from issuing judgment on these facts alone mean that the standard logic of probability does not or should not apply to trials?¹⁷¹ Scholars who are comfortable with using standard probability logic have provided many answers to the paradox, including the argument that if that is all the plaintiff chooses to present then we cannot be sure that the probability is actually over 50%, because a negative inference can be drawn from lack of other proof.¹⁷²

Litigation, Cardozo School of Law, Yeshiva University, March 24-26, 1991. Email I from Peter Tillers to Roger Park, July 8, 2005. When contacted, Professor Tillers also noted the influence of activities such as William Twining’s 1982 conference on “Facts in Law” at Durham University and the honors seminars organized by Adrian Zuckerman at Oxford that started in 1984. Email II from Peter Tillers to Roger Park, July 8, 2005. For a European perspective on the development of the field, see William Twining, *RETHINKING EVIDENCE: EXPLORATORY ESSAYS passim* and 349-50 (1990).

¹⁷⁰ See Peter Tillers, *Web of Things in the Mind: A New Science of Evidence*, 87 MICH. L. REV. 1225 (1989) (review of David Schum’s *EVIDENCE AND INFERENCE FOR THE INTELLIGENCE ANALYST* (1987), with comments on his contributions to the science of evidence).

¹⁷¹ See JONATHAN COHEN, *THE PROBABLE AND THE PROVABLE* (1977) (describing anomalies that arise if Pascalian probability theory is applied in certain legal situations, and arguing that a “Baconian” approach would be more just).

¹⁷² See, e.g., Richard A. Posner, *An Economic Approach to the Law of Evidence*, 51 STAN. L. REV. 1477, 1508-10 (1999). Judge Posner also argues that, even if other evidence cannot be obtained, the cost of trying such marginal cases outweighs the benefits. *Id.* at 1510. For further discussion of the general topic of “naked statistical evidence,” see Daniel Shaviro, *Statistical-Probability Evidence and the Appearance of Justice*, 103 HARV.L.REV. 530, 533-36 (1989); Ronald J. Allen, *On the Significance of Batting Averages and Strikeout Totals: A Clarification of the “Naked Statistical Evidence” Debate, The Meaning of “Evidence,” and the Requirement of Proof Beyond a Reasonable Doubt*, 65 TUL.L. REV. 1093 (1991); Craig R. Callen, *Adjudication and the Appearance of Statistical Evidence*, 65 TUL. L. REV. 457 (1991); Daniel Shaviro, *A Response to Professor Allen*, 65 TUL. L. REV. 1111 (1991); Daniel Shaviro, *A Response to Professor Callen*, 65 TUL. L. REV. 499 (1991).

The conjunction problem raises similar questions. Standard instructions tell jurors that if they find each element to be true by a preponderance of the evidence, they should find for the plaintiff. The problem is that proving each element to be more probable than not does not prove that the conjunction of all elements is more probable than not. If we make the assumption that the elements are independent of each other, this instruction is technically inaccurate because if there are two elements and the probability of each of them being true is 60%, then the probability of both being true is $.6 \times .6$, or 36%. Does that mean that the standard logic of probability is inapposite, and that we therefore must think of legal reasoning as proceeding from some other basis? The conjunction problem is one reason why Professor Allen has proposed reconceptualization of civil trials in a way that tells juries to decide whose story is most believable, as opposed to telling them to decide whether plaintiff has established each element by a preponderance of the evidence.¹⁷³ Again, there are answers, including the answer that jury instructions are often ambiguous on the question whether it is sufficient merely to prove each element by a preponderance, and that instructions sometimes suggest that the joint occurrence of all elements must be proven.¹⁷⁴ Moreover, the approach of telling juries to judge relative plausibility of the parties' stories would lead to its own anomalies. Suppose, for example, the defendant tells one story, the plaintiff tells a slightly more plausible story, and the jury believes that a third story not offered by either party (but favoring the defendant) is the true one.¹⁷⁵

Another topic, the use of Bayes' Theorem to evaluate evidence and evidence law, has become one of the centerpieces of the new evidence scholarship.¹⁷⁶ Bayes' Theorem is a basic tenet of probability theory that can be used to adjust a probability assessment upon receiving new evidence. For example, imagine a case in which the issue is whether the defendant is the source of a hair found at a crime scene. After hearing testimony of lay witnesses, the fact-finder forms an opinion that the odds are 2 to 1 that the defendant is the source of the hair. In addition, there is expert testimony that mitochondrial DNA sequencing shows that the hair has a genetic profile that has a population frequency of 1%. The defendant's hair matches that profile. Bayes' theorem would provide a way of updating the prior estimate of 2 to 1 odds with the new information of the match.

¹⁷³ Ronald J. Allen, *A Reconceptualization of Civil Trials*, 66 B.U. L. REV. 401 (1986)

¹⁷⁴ See Dale A. Nance, *A Comment on the Supposed Paradoxes of a Mathematical Interpretation of the Logic of Trials*, 66 B.U. L. Rev. 947 (1986).

¹⁷⁵ See Dale A. Nance, *Naturalized Epistemology and the Critique of Evidence Theory* 87 VA. L. REV. 1551, 1575-80 (2001). For another discussion, see Richard A. Posner, *An Economic Approach to the Law of Evidence*, 51 STAN. L. REV. 1477, 1512-15 (1999).

¹⁷⁶ For a useful introduction, see David McCord, *A Primer for the Nonmathematically Inclined on Mathematical Evidence in Criminal Cases: People v. Collins and Beyond*, 47 WASH. & LEE L. REV. 741 (1990).

In the above example, the probability of this test result, given a defendant who is not the source of the hair, is 1 in 100. (The random match probability in the population.) That does not mean that there is a 1 in 100 probability that the defendant is the source, given the mt-DNA test results. The error of confusing these two probabilities is known as “transposing the conditional.”

The following example illustrates the error of transposing the conditional. The probability that a person has committed a crime, given that he is in prison $\text{Pr}(\text{CIP})$ is obviously not the same as the probability that a person is in prison, given that he has committed a crime $\text{Pr}(\text{PIC})$. The latter probability is much lower, considering that not all criminals are caught and that not all crimes are punished with incarceration. To believe that $\text{Pr}(\text{CIP})$ must be equal to $\text{Pr}(\text{PIC})$ is the error of transposing the conditional.

Another way to understand the error of transposing the conditional is to suppose that there is absolutely no evidence against the defendant except the mt-DNA test on the hair. Suppose there are 100,000 other people living in the vicinity and no reason except the mt-DNA finding to point a finger at the defendant instead of one of the others. Suppose also that the defendant’s genetic profile matches and the probability of a random match is 1%. The probability of a match, given that the defendant is not the source, is 1%. That does not mean that the probability the defendant is not the source, given that there is a match, is only 1%. One would expect, out of the 100,000 other possible suspects, that 1,000 of them would also match. If so, the probability that the defendant is not the source, given a match, is closer to 99.9% than to 1%.¹⁷⁷ Of course, the probability that defendant is the source would increase dramatically once other inculpatory evidence is offered, such as testimony that the defendant had a motive and was seen near the crime scene.

A fact-finder who used Bayes’ theorem to take account of the mt-DNA test could arrive at a probability that the defendant is the source without falling into transposition error. Suppose again that the fact-finder estimates the prior odds that the defendant was the source (before the DNA test) to be 2 to 1. That fact-finder could multiply the prior odds by a statistic called the “likelihood ratio” to obtain the posterior odds (the revised estimate of the odds the defendant was the source, after taking the mt-DNA test into account). The likelihood ratio is derived by dividing the probability of finding the evidence of a match, given that the defendant is the source, by the probability of finding it, given that the defendant is not the source. Assume that the probability of finding a match, given that the defendant is the source, is 100%.¹⁷⁸ If the defendant is not the

¹⁷⁷ For purposes of the example, we are ignoring the possibility that someone not living in the vicinity is the source.

¹⁷⁸ This assumption is not logically compelled. It would not be valid, for example, if laboratory error were taken into account, or if hairs from the same person could have different mt-DNA profiles. But here we assume that we would find a match every time if the hair came from the defendant.

source, the probability of finding a match is 1/100 or 1% (the probability of a random match in the general population).¹⁷⁹ The likelihood ratio therefore is 100% divided by 1%, or 100. To calculate the posterior odds, one would multiply the prior odds of 2 to 1 by the likelihood ratio of 100. Thus, the fact-finder using this method would conclude that the updated odds that defendant was the source are 200 to 1.

In an article published early in the Bayesian debate, Finkelstein and Fairly¹⁸⁰ suggested using Bayes' theorem to aid jurors in cases in which the issue is the identity of the perpetrator and the perpetrator has left trace evidence at the scene of the crime. They constructed a hypothetical case in which the accused is charged with murdering his girlfriend, and the perpetrator of the crime left behind a hand print. They illustrated how Bayes' theorem could be used to get from the random match probability (the frequency of handprint features in the general population) to an estimate of how likely it was that the defendant left the print at the scene.

In a celebrated reply, Professor Laurence Tribe, then an evidence teacher, raised several objections to this use of Bayes' theorem in the trial process. Jurors who are not proposition bettors might have mistaken or inconsistent understandings of the meaning of prior probability.¹⁸¹ There is a danger of "dwarfing soft variables," that is, the danger that the impressiveness of statistics would obscure other issues (such as whether there might be an innocent explanation for the presence of the print). Moreover, uncertainty about facts upon which the Bayesian calculations would be based could require additional quantification decisions about so many issues that use

¹⁷⁹ These assumptions require that we ignore the danger of laboratory error. Putting the danger of lab error aside is controversial among scholars, even in the usual case in which frequency evidence is presented without any Bayesian interpretation. The fact-finder may be prejudiced when an extremely small match probability is presented without incorporating (or even estimating) the danger that a match might occur through lab error. On the other hand, it seems reasonable to consider one thing at a time (first what the evidence shows assuming no lab error, then the danger of lab error in the specific case). Compare Jonathan J. Koehler, Audrey Chia, and Samuel Lindsey, *The Random Match Probability in DNA Evidence: Irrelevant and Prejudicial?* 35 JURIMETRICS J. 201 (1995) and Jonathan J. Koehler, *Why DNA Likelihood Ratios Should Account for Error (Even When a National Research Council Report Says They Should Not)*, 37 JURIMETRICS J. 425 (1997) with Margaret A. Berger, *Laboratory Error Seen Through the Lens of Science and Policy*, 30 U.C. DAVIS L. REV. 1081 (1997), and Edward J. Imwinkelried & D.H. Kaye, *DNA Typing: Emerging or Neglected Issues*, 76 WASH. L. REV. 413, 460-64 (2001).

¹⁸⁰ Finkelstein & Fairley, *supra* note ____.

¹⁸¹ Tribe pointed out that to some jurors a probability of .50 might stand for what the chances were before any evidence was introduced, while to others it might mean that the search has to be narrowed to two suspects. Tribe, *supra* note ____, at 1358-59.

of the theorem would be more confusing than helpful.¹⁸² If there is a danger that the expert might be mistaken about the frequency of the print, for example, that would have to be taken into account somehow, but it would not be easy for a fact-finder to adjust the statistics. Tribe's analysis was in turn criticized by psychologists for making faulty psychological assumptions¹⁸³ as well as by statisticians,¹⁸⁴ philosophers,¹⁸⁵ and legal scholars.¹⁸⁶

Bayesian skeptics have continued to point out problems with the use of Bayes in the trial process. Progressive updating using Bayes' theorem throughout the trial would be so computationally complex that it would be beyond the capacity of fact-finders. Moreover, trials are structured in such a way that the jury does not receive information in a way that facilitates Bayesian updating. For example, the jury might have difficulty formulating priors when it does not get instructions about the law until the end, and hence does not know exactly what proposition it is supposed to decide.¹⁸⁷

To some extent, Bayesian enthusiasts and Bayesian skeptics seem to be talking past each other. The skeptics have demonstrated that it is not practical to use Bayesian analysis very often in the course of trial, but most of the enthusiasts do not argue for

¹⁸² For Bayes' Theorem to be helpful in the hypothetical described by Finkelstein and Fairley, one has to assume that the person whose handprint is on the knife is guilty of being the killer, and that the handprint expert is accurate. A factfinder who believed that these two facts were not certain would have to discount the probability estimate obtained by using Bayes's theorem, and giving instructions about how to do this would be too complicated to be feasible. Tribe, *supra* note __, at 1361-65.

¹⁸³ Michael J. Saks & Robert F. Kidd, *Human Information Processing and Adjudication: Trial by Heuristics*, 15 L. & SOC'Y REV. 123 (1980); David Faigman & A.J. Baglioni, *Bayes' Theorem in the Trial Process*, 12 L. & HUM. BEHAV. 1 (1988); Brian C. Smith, Steven D. Penrod, Amy L. Otto, and Roger C. Park, *Jurors' Use of Probabilistic Evidence*, 20 L. & HUM. BEHAV. 49 (1996).

¹⁸⁴ Richard D. Friedman, *Answering the Bayesioskeptical Challenge*, 1 INT'L J. EVID. & PROOF 276 (1997).

¹⁸⁵ Cohen, 1977, *supra* note __.

¹⁸⁶ David Kaye, *The Paradox of the Gatecrasher and Other Stories*, 1979 ARIZ. ST. L. J. 101.

¹⁸⁷ See Ronald J. Allen, *Rationality, Algorithms and Judicial Proof: A Preliminary Inquiry*, 1 INTERNATIONAL JOURNAL OF EVIDENCE AND PROOF 254 (1997); Mike Redmayne, *Presenting Probabilities in Court: The DNA Experience*, 1 INTERNATIONAL JOURNAL OF EVIDENCE AND PROOF 187, 199-208 (1997) (although Professor Redmayne suggests with "some trepidation" that Bayesian presentations will be appropriate in some cases, *id.* at 213, his article acknowledges and illustrates practical and conceptual difficulties). See also Alex Stein, *Judicial Fact-Finding and the Bayesian Method: The Case for Deeper Skepticism About Their Combination*, 1 INTERNATIONAL JOURNAL OF EVIDENCE AND PROOF 25 (1997).

such a use.¹⁸⁸ Bayesians often emphasize the value of Bayesian reasoning outside the heat of trial, in assessing the value of rules excluding evidence or in weighing the probative value of certain types of evidence. A basic Bayesian perspective asks about the degree to which new evidence changes our estimate of the odds of a fact being true, and tells us to compare the likelihood of finding the evidence if the fact were true with the likelihood of finding it if the fact were false. It can, for example, help us understand why prior convictions have little probative value in impeaching the testimony of the accused in a criminal case. Learning of the evidence of the prior conviction does not change our prior estimate of the odds that the defendant would lie. If he is guilty of a serious charge, the situational pressures to lie are so strong that even a generally honest person likely would lie to escape punishment; learning the fact that he was dishonest on an earlier occasion does not change the odds much. If he is innocent of the charged crime, the situational pressures are likely to push him toward telling the truth even if he is a veteran liar, and, anyway, it does not much matter if he lies his way out of a false charge.¹⁸⁹ Of course, one could arrive at the same insight without knowledge of Bayes' theorem, but it seems likely that exposure to the approach helps us be sensitive to the right factors in assessing probative value.¹⁹⁰ Similarly, a Bayesian perspective can help us understand why evidence that a man accused of murdering his wife had beaten her on previous occasions is probative, despite the fact that spousal abuse is common and that very few abusers progress to murder.¹⁹¹ Scholars have also relied on Bayesian models to analyze such diverse evidentiary issues as the meaning of the concept of relevance,¹⁹² the value of

¹⁸⁸ Friedman, *supra* note ____.

¹⁸⁹ See Richard D. Friedman, *Character Impeachment Evidence: Psycho- Bayesian [!?] Analysis and a Proposed Overhaul*, 38 UCLA L. REV. 637 (1991) (the bracketed punctuation [!?] is contained in Professor Friedman's original title); Richard A. Posner, *An Economic Approach to the Law of Evidence*, 51 STAN. L. REV. 1477 (1999). Professor Uviller arrived at a similar conclusion without reference to Bayes' theorem. See H. Richard Uviller, *Evidence of Character to Prove Conduct: Illusion, Illogic, and Injustice in the Courtroom*, 130 U.PA.L. REV. 845, 867-68(1982) ("it seems quite likely that a guilty person without prior convictions will lie on the stand as readily as will a guilty veteran, while innocent people with extensive criminal histories will testify as truthfully as the innocent novice").

¹⁹⁰ See Richard Lempert, *Of Flutes, Oboes, and the As If World of Evidence Law*, 1 INTERNATIONAL JOURNAL OF EVIDENCE AND PROOF 253 (1997); Lempert, *Modeling Relevance*, 75 MICH. L. REV. 1021 (1977).

¹⁹¹ Richard A. Friedman and Roger C. Park, *Sometimes What Everybody Thinks They Know is True*, 27 LAW AND HUMAN BEHAVIOR 29 (2003).

¹⁹² Lempert, *Modeling Relevance*, *supra* note ____.

forensic identification evidence,¹⁹³ the proper interpretation of DNA evidence,¹⁹⁴ the value of expert testimony in child abuse cases,¹⁹⁵ the value of hearsay,¹⁹⁶ and the appropriateness of questioning children in a suggestive manner.¹⁹⁷

A number of Bayesian enthusiasts believe that the theorem can be useful at trial, though few would argue for a step-by-step updating by applying Bayesian analysis in a multi-step fashion to discrete pieces of evidence. For example, in a DNA case one might aggregate all the non-DNA evidence, assign prior odds, and then multiply by a likelihood ratio derived from the DNA evidence. How to combine the evidence might be demonstrated to the jury either by using a chart showing prior and posterior odds under different assumptions about prior odds, or by telling them to multiply the prior odds by the likelihood ratio.¹⁹⁸ This approach is still problematic, for reasons discussed by

¹⁹³ BERNARD ROBERTSON & G.A. VIGNAUX, INTERPRETING EVIDENCE (1997).

¹⁹⁴ IAN EVETT & BRUCE WEIR, INTERPRETING DNA EVIDENCE: STATISTICAL GENETICS FOR FORENSIC SCIENTISTS (1998).

¹⁹⁵ Thomas Lyon & Jonathan J. Koehler, *The Relevance Ratio: Evaluating the Probative Value of Expert Testimony in Child Sexual Abuse Cases*, 82 CORNELL L. REV. 43 (1996).

¹⁹⁶ Peter Tillers & David A. Schum, *Hearsay Logic*, 76 MINN. L. REV. 813 (1992).

¹⁹⁷ Stephen Ceci & R.D. Friedman, *The Suggestibility of Children: Scientific Research and Legal Implications*, 86 CORNELL L. REV. 33 (2000); Thomas D. Lyon, *The New Wave in Children's Suggestibility Research: A Critique*, 84 CORNELL L. REV. 1004 (1999). For other examples of applications of Bayes in evaluation of evidence law, see Richard Friedman, *Towards a (Bayesian?) Convergence*, 1 INTERNATIONAL JOURNAL OF EVIDENCE AND PROOF 348 (1997).

¹⁹⁸ See Dale A. Nance, *Naturalized Epistemology and the Critique of Evidence Theory*, 87 VIRGINIA LAW REVIEW 1551, 1610-16 (2001); Dale A. Nance & Scott B. Morris, *An Empirical Assessment of Presentation Formats for Trace Evidence with a Relatively Large and Quantifiable Random Match Probability*, 42 JURIMETRICS J. 403 (2002). For advocacy of the approach of telling the trier about the likelihood ratio and explaining how to use it to combine DNA evidence with other evidence, see Geoffrey K. Chambers, Stephen J. Cordiner, John S. Buckleton, Bernard Robertson, & G.A. Vignaux, *Forensic DNA Profiling: The Importance of Giving Accurate Answers to the Right Questions* 8 CRIM. L.F. 445 (1997). See also Bernard Robertson & G.A. Vignaux, INTERPRETING EVIDENCE 20-21, 51-65 (1997) (supporting an approach under which a forensic expert explain to the trier of fact, "This evidence is R times more probable if the accused left the mark than if someone else did. This evidence therefore [very][strongly] supports the proposition that the accused left the mark." *Id.* at 65).

Tribe¹⁹⁹ and Allen,²⁰⁰ and courts have been slow to adopt it.²⁰¹ Proponents of decision aids have nonetheless made some progress. In paternity cases, for example, some courts have allowed charts to be provided to jurors showing how a prior probability of paternity should be revised in light of the paternity index (likelihood ratio) associated with a genetic test.²⁰² Experts sometimes use Bayes' theorem in calculating an estimate of paternity that they then offer to the trier as the probability of paternity, though this procedure is controversial among scholars because it involves making an artificial assumption that the prior probability of paternity is 50%. This assumption is made without taking into account the actual non-test evidence, so that the 50% figure would be used even if there were convincing evidence that the defendant was sterile.²⁰³ The National Research Council has suggested that experts might compute posterior probabilities to show jurors the power of DNA evidence for establishing identity²⁰⁴ although this proposal also remains controversial.²⁰⁵

When Bayesian skeptics and Bayesian enthusiasts address common ground, there seems to be a degree of convergence. Professor Allen, a leading Bayesian

¹⁹⁹ Tribe, *Precision and Ritual*, *supra* note ____.

²⁰⁰ Allen, *Algorithms*, *supra* note ____.

²⁰¹ On the infrequency of use of these Bayesian presentation methods except in paternity cases, see 1 DAVID L. FAIGMAN ET AL., *MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY* §§ 15-5.4 to 15-5.5, at 656-59 (1997) (with footnotes citing illustrative cases).

²⁰² David Kaye, Plemel as a Primer on Proving Paternity, 24 WILLAMETTE L. J. 867 (1988).

²⁰³ See, e.g., Kammer v. Young, 535 A.2d 936 (Ct. Spec. Appeals Md. 1988) (affirming judgment in case in which expert, pursuant to command of Maryland statute, gave opinion as to probability of paternity based on Bayesian calculation with assumed prior odds of .5). For criticism of this approach, see Jonathan J. Koehler, *DNA Matches and Statistics: Important Questions, Surprising Answers*, 76 JUDICATURE 222 (1993); BERNARD ROBERTSON & G.A. VIGNAUX, *INTERPRETING EVIDENCE* 25-27 (1995). Of course, in many cases the 50% assumed prior probability is generous to the defendant, and in any event the artificiality of the assumption could be brought out on cross-examination and the expert asked to calculate the probability of paternity based on other assumptions.

²⁰⁴ NATIONAL RESEARCH COUNCIL, *THE EVALUATION OF FORENSIC DNA EVIDENCE* (1996).

²⁰⁵ William C. Thompson, *Accepting Lower Standards: The National Research Council's Second Report on Forensic DNA Evidence*, 37 JURIMETRICS 405 (1997).

skeptic, has allowed that “there may very well be situations involving virtually purely statistical evidential bases in which Bayes’ theorem would be a useful analytic tool,” and that “the Bayesian sceptic does not deny a use for Bayes’ theorem as an analytical tool.”²⁰⁶ Professor Friedman, a leading enthusiast, has written that, at least in cases in which statistical evidence does not otherwise play a substantial role, there is “usually no substantial reason to make an explicit presentation of probability theory; fact-finders can deal with the evidence much as they deal with ordinary questions in their everyday lives.”²⁰⁷

Some of the Bayesian debaters are more prone to thought experiments than to empirical research, but there have been efforts to test whether jurors are intuitive Bayesians and whether Bayesian charts or instructions will aid decision-making. Numerous studies have shown that most of the time human beings do not behave according to the Bayesian ideal.²⁰⁸ Generally, humans are too conservative, failing to adjust their prior estimates to the degree required by Bayesian models of rational decision-making. In a more fundamental way, Bayesian modeling fails to capture the decision processes of human decision-makers. Jurors appear to evaluate evidence in a trial not by sequential updating but by constructing plausible narratives that might account for the evidence.²⁰⁹ As some researcher have concluded: humans are story tellers, not meter readers. But in cases in which the human story is supplemented by forensic match evidence, there is some empirical evidence that human decision-making might be improved, or at least come closer to the Bayesian norm, by use of a Bayesian

²⁰⁶ Allen, *Rationality, Algorithms and Judicial Proof*, supra note ___, at 258.

²⁰⁷ Richard D. Friedman, *Answering the Bayesioskeptical Challenge*, 1 INTERNATIONAL JOURNAL OF EVIDENCE AND PROOF 276, 290 (1997). See also John A. Michon, *The Time Has Come to Put This Debate Aside and Move On To Other Matters*, 1 INTERNATIONAL JOURNAL OF EVIDENCE AND PROOF 331-334 (1997).

²⁰⁸ Jason Schklar & Shari Seidman Diamond, *Juror Reactions to DNA Evidence: Errors and Expectancies*, 23 LAW & HUM. BEHAV. 159 (1999); Jane Goodman, *Jurors' Comprehension and Assessment of Probabilistic Evidence*, 16 Am. J. Trial Advoc. 361 (1992); Brian C. Smith, Steven D. Penrod, Amy L. Otto, and Roger C. Park, *Jurors' Use of Probabilistic Evidence*, 20 L. & HUM. BEHAV. 49 (1996); David L. Faigman & A. J. Baglioni, Jr., *Bayes' Theorem in the Trial Process: Instructing Jurors on the Value of Statistical Evidence*, 12 LAW & HUM. BEHAV. 1 (1988); William C. Thompson & Edward L. Schumann, *Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the Defense Attorney's Fallacy*, 11 LAW & HUM. BEHAV. 167 (1987).

²⁰⁹ Nancy Pennington & Reid Hastie, *A Cognitive Theory of Juror Decision Making: The Story Model*, 13 CARDOZO L. REV. 519 (1991); Nancy Pennington & Reid Hastie, *Explaining the Evidence: Tests of the Story Model for Juror Decision Making*, 62 J. PERSONALITY & SOC. PSYCHOL. 189 (1991).

chart showing the trier how estimates of prior probability should be changed by the forensic evidence.²¹⁰

Another thrust of “the new evidence scholarship” is concerned with inference and decision-making in litigation, with the problem of processing evidence and drawing inferences from it in order to prepare for trial, to try cases, and to decide them. It seeks to develop theories of inference in the litigation context.²¹¹ This area of scholarship is concerned with the problem of how lawyers can organize the mass of evidence in a case in a meaningful and effective way if it is to be persuasive to factfinders. How are lawyers to sort through the maze of evidence to determine which propositions that need to be proved are supported by what evidence, in a complex interconnected hierarchy of raw facts, intermediate inferences, and ultimate conclusions?

Wigmore²¹² developed the first system for organizing and assessing evidence for litigation by employing careful logic to trace the factual support for inferences. “Wigmore analysis is an attempt to capture the way we think when we think at our best.”²¹³ Wigmore charting was a major milestone in lawyerly thinking about facts, but it still was somewhat crude. Anderson and Twining²¹⁴ not only resurrected Wigmore charting, but improved upon it, such as by enabling it to take into account the applicable substantive law and by expanding it beyond requiring the chartist to have a single “ultimate probandum” in mind before starting. Instead, the chartist is able to explore alternative conclusions to which the evidence might lead. Wigmore and related kinds of charting²¹⁵ clarify the elements of evidence and their inter-relationships. Moreover, the addition of an element to the chart is an implicit probability judgment of the element's

²¹⁰ See Dale A. Nance, *Naturalized Epistemology*, *supra* note ___, at 1610-16; Dale A. Nance & Scott B. Morris, *An Empirical Assessment of Presentation Formats for Trace Evidence with a Relatively Large and Quantifiable Random Match Probability*, 42 JURIMETRICS J. 403 (2002).

²¹¹ William Twining, *The New Evidence Scholarship*, 13 CARDOZO L. REV. 295 (1991). For a leading example, see Peter Tillers & David Schum, *A Theory of Preliminary Fact Investigation*, 24 U.C. DAVIS L. REV. 931 (1991).

²¹² JOHN H. WIGMORE, *THE SCIENCE OF PROOF: AS GIVEN BY LOGIC, PSYCHOLOGY, AND GENERAL EXPERIENCE AND ILLUSTRATED IN JUDICIAL TRIALS* (1913; 2nd ed. 1931; 3rd ed. 1937).

²¹³ Bernard Robertson & G.A. Vignaux, *Taking Fact Analysis Seriously*, 91 MICH. L. REV. 1442, 1447 (1993). See also BERNARD ROBERTS & G.A. VIGNAUX, *INTERPRETING EVIDENCE* (1995).

²¹⁴ TERENCE ANDERSON & WILLIAM TWINING, *ANALYSIS OF EVIDENCE: HOW TO DO THINGS WITH FACTS BASED ON WIGMORE'S SCIENCE OF JUDICIAL PROOF* (1991; 1998).

²¹⁵ Namely, the improvements offered by ANDERSON & TWINING, *id.*, and by such devices as decision trees, influence diagrams, and Bayes networks. See ROBERTSON & VIGNAUX, 1995, *supra* note __.

importance. It is a small step to add quantified probability statements either of the empirical kind or the chartist's subjective probability of the element.²¹⁶

We should add a brief mention of evidence scholars who have addressed philosophical issues about the foundations of knowledge. At least three distinguished evidence scholars have assessed the significance of philosophical skepticism to the law of evidence. Despite some differences, they have essentially concluded that that lawyers and evidence scholars need not worry about the implications of profound skepticism.²¹⁷ The basic suppositions of a system of litigation require rejecting profound skepticism, even if one sees goals other than truth-finding to be central.

There is, however, a debate about the extent to which we should enthrone truth-finding as the central goal of evidence law and believe that the goal can be accomplished. Professor William Twining has noted and described the tradition of “optimistic rationalism” in evidence law and evidence scholarship.²¹⁸ Professor Michael Siegel has argued that philosophical pragmatism is a better foundation for thinking about evidence than optimistic rationalism. He posits that the rationalist tradition has caused too great an emphasis on truth-finding, causing theorists to underestimate the value of other goals, such as making verdicts more acceptable to the public and ending disputes in an efficient way.²¹⁹ In contrast, and without relying to the same degree upon the

²¹⁶ Robertson & Vignaux, 1993, *supra* note ___, at 1456.

²¹⁷ See William Twining, *Some Scepticism about Some Scepticisms*, in William Twining, *RETHINKING EVIDENCE: EXPLORATORY ESSAYS* 92-134 (1990 reprint of essay published in 1984) (few philosophers consistently maintain philosophical skepticism, and the literature “poses few threats to the centrality of concepts of Truth, Reason and Justice in any theory of Evidence” Id. at 134 (Notes)); Mirjan Damaska, *Truth in Adjudication*, 49 HASTINGS LAW JOURNAL 289 (1998) (when we engage in adjudication, we presuppose a world outside our statements; radical postmodern thought has no bearing on evidence law); Ronald J. Allen, *Truth and Its Rivals*, 49 HASTINGS LAW JOURNAL 309, 315, 317, 318 (1998) (“There is indeed an important philosophical puzzle, but it is primarily how to explain what is obviously true”; the implications of scepticism for the legal system are “marginal at best”; all the rivals to truth as a goal of litigation themselves involve rejection or cabining of the philosophical problem of scepticism). But cf. D. Nicolson, *Truth, Reason, and Justice: Epistemology and Politics in Evidence Discourse*, 57 MOD. L. REV. 726 (1994), critiqued in John D. Jackson, *Modern Trends in Evidence Scholarship: Is All Rosy in the Garden?*, 21 QUINNIPIAC L. REV. 893, 898-99 (2003)

²¹⁸ See William Twining, *The Rationalist Tradition of Evidence Scholarship*, in William Twining, *RETHINKING EVIDENCE: EXPLORATORY ESSAYS* 32-91 (1990).

²¹⁹ Michael L. Seigel, *A Pragmatic Critique of Modern Evidence Scholarship*, 88 NW. U. L. REV. 995 (1994). For comments on critiques of the rationalist tradition, see John D. Jackson, *Modern Trends in Evidence Scholarship: Is all Rosy in the Garden?* 21 QUINNIPIAC L. REV.

literature on philosophy, Professor Nesson has argued that the legal system already gives primacy to making verdicts acceptable instead of to finding the truth.²²⁰

Professors Allen and Leiter have sought to bring lawyers up to date on epistemology in an exploration of contemporary work on naturalized epistemology. . Though their work is tough sledding for readers without a background in philosophy, it reaches conclusions that should be comforting to interdisciplinary evidence scholars. As Allen and Leiter themselves note, "For the great bulk of evidentiary scholars, then, this article merely solidifies the ground beneath their feet."²²¹ They maintain that philosophy should not be an a priori discipline, but one that is continuous with a posteriori inquiry in the empirical sciences.²²² They approve of a functional approach to evidence law, assessing the wisdom of rules in light of their social effects, and using the tools of social science to study consequences of legal rules: hypotheses should be tested empirically and discarded them in the face of disconfirming data. Their illustrations of this proposition include criticism of some of Judge Posner's speculative conclusions about the economics of evidence law.²²³ Finally, because naturalized epistemology is instrumental, evidence rules should only require intellectual performances that fact-finders are capable of doing – "ought implies can" – a position that they see as militating against searches for formal "algorithms," such as Bayes theorem, for use in fact-finding at trial.²²⁴

IV. Feminist Evidence Scholarship

Since 1990, an increasing number of law review articles have examined evidence law from a feminist perspective. Most of this scholarship has dealt with evidence

893, 898-901 (2003).

²²⁰Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357 (1985). For critique, see Ronald J. Allen, *Rationality, Mythology, and the 'Acceptability of Verdicts' Thesis*, 66 B.U. L. REV. 541 (1986).

²²¹87 VA. L. REV. 1491, 1493 (2001)

²²²*Id.* at 1494, 97.

²²³*Id.* at 1503, 1511, 1516-25.

²²⁴*Id.* at 1492-93, 1499. For a rebuttal, see Dale A. Nance, *Naturalized Epistemology and the Critique of Evidence Theory*, 87 VA. L. REV. 1551 (2001). Cf. Brian Leiter, *The Epistemology of Admissibility: Why Even Good Philosophy of Science Would Not Make for Good Philosophy of Evidence* 1997 B.Y.U.L. REV. 803 (arguing that *Daubert* makes unrealistic demands on the epistemic capacities of adjudicators)

doctrines that relate to sexual assault and spousal abuse. The sexual assault topics include rape shield rules,²²⁵ exceptions to the rule against character evidence for prior crimes of the alleged perpetrator,²²⁶ and admissibility of rape trauma syndrome testimony.²²⁷ Spousal abuse topics have included whether victims should be forced to testify over a claim of spousal immunity,²²⁸ whether prior acts of domestic violence against other victims should be admissible,²²⁹ and admissibility of social science evidence in spousal abuse cases, including testimony about battered woman's syndrome.²³⁰ Feminist scholars have also addressed evidence issues the impact of

²²⁵ See, e.g., Aviva Orenstein, *No Bad Men!: A Feminist Analysis of Character Evidence in Rape Trials*, 49 HASTINGS L.J. 663 (1998) (supporting rape shield legislation but opposing legislation admitting prior sex crimes of the alleged perpetrator).

²²⁶ See Katharine K. Baker, *Once a Rapist? Motivational Evidence and Relevancy in Rape Law*, 110 HARV. L. REV. 563 (1997); Orenstein, *supra* note ____.

²²⁷ Compare Toni M. Massaro, *Experts, Psychology, Credibility, and Rape: The Rape Trauma Syndrome Issue and its Implication for Expert Psychological Testimony*, 69 MINN. L. REV. 395,398 (1985) ("RTS evidence can help the fact finder resolve difficult issues of guilt or innocence and that such evidence can educate jurors and judges, which may help correct erroneous social attitudes about the crime of rape") with Susan Stefan, *The Protection Racket: Rape Trauma Syndrome, Psychiatric Labeling, and Law*, 88 NW. U. L. REV. 1271, 1277 (1994) ("The introduction of rape trauma syndrome evidence in criminal trials has probably assisted in convicting rapists, but at the cost of reinforcing the very myths and assumptions that early feminists fought so hard to eliminate").

²²⁸ For scholars who have urged that the victim not be allowed to invoke spousal immunity to refuse to testify against a battering husband, see Malinda Seymore, *Isn't It a Crime: Feminist Perspectives on Spousal Immunity and Spousal Violence*, 90 NW. U. L. REV. 1032 (1996) (the husband is likely to coerce the wife into not testifying, and treating the matter as "private" violates feminist norms) and Cheryl Hanna, *No Right to Choose: Mandated Victim Participation in Domestic Violence Prosecutions*, 109 HARV. L. REV. 1849 (1996) (favoring mandated victim participation in domestic violence cases for a variety of reasons, including deterrence and the educative function of the law).

²²⁹ See Lisa Marie DeSanctis, *Bridging the Gap Between the Rules of Evidence and Justice for Victims of Domestic Violence*, 8 YALE J.L. & FEMINISM 359 (1996). (arguing for admission to combat gender bias, educate jury about nature of batterers, and overcome problems caused by lack of other evidence and coerced nonparticipation by victims). See also Lisa A. Linsky, *Use of Domestic Violence History Evidence in the Criminal Prosecution: A Common Sense Approach*, 16 PACE L. REV. 73 (1995); Myrna S. Raeder, *The Admissibility of Prior Acts of Domestic Violence: Simpson and Beyond*, 69 S. CAL. L. REV. 1463 (1996).

²³⁰ Battered women syndrome testimony, as represented in the work of Lenore Walker, is sometimes offered into evidence to explain the conduct of women who kill their husbands. See generally LENORE WALKER, *THE BATTERED WOMAN* (1979); LENORE WALKER, *THE BATTERED WOMAN SYNDROME* (1984). Walker posits that women who do not leave their battering husbands are experiencing "learned helplessness," a phenomenon similar to that experienced by dogs in an experiment

which upon women is less obvious, including the excited utterance exception,²³¹ the

where dogs which could not control the electric shocks they were receiving simply stopped trying to avoid them. BWS theory has been criticized both as degrading to women and as based on poor social science methodology, a criticism to which the present authors subscribe. See Erica Beecher-Monas, *Domestic Violence: Competing Conceptions of Equality in the Law of Evidence*, 47 LOY. L. REV. 81 (2001) (criticizing BWS as degrading and methodologically flawed, and advocating admission of social context evidence and PTSD testimony) and Regina Schuller & Sara Rzepa, *The Battered Woman Syndrome and Other Psychological Effects of Domestic Violence Against Women*, § 11, in MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY (David L. Faigman, David H. Kaye, Michael J. Saks & Joseph Sanders eds., 2002) (reviewing the empirical research on which BWS theory is based); David L. Faigman & Amy J. Wright, *The Battered Woman Syndrome in the Age of Science*, 39 ARIZONA L. REV. 67 (1997); Janet C. Hoeffel, *The Gender Gap: Revealing Inequities in Admission of Social Science Evidence in Criminal Cases*, 24 U. ARK. LITTLE ROCK L. REV. 41 (Fall 2001) (arguing that feminists seek admission of social science evidence that fails the *Daubert* test, and asserting that this goal is pursued at the expense of socially disadvantaged criminal defendants, such as young African-American males). Cf., Myrna S. Raeder, *Battered Woman and Batterer Syndrome: The Double-edged Sword: Admissibility of Battered Woman Syndrome by and Against Batterers in Cases Implicating Domestic Violence*, 67 U. COLO. L. REV. 789, 797 (1996) (noting deficiencies in empirical support for Walker's cycle of violence thesis, but saying that "acceptance of the cycle is justified in cases where the factual foundation supports its existence in the particular relationship in question"). Some authors have suggested substitutes, such as evidence about other social science findings about domestic abuse. See *id.*; Myrna S. Raeder, *The Better Way: The Role of Batterers' Profiles and Expert "Social Framework" Background in Cases Implicating Domestic Violence*, 68 U. COLO. L. REV. 147 (1997).

²³¹ Aviva Orenstein, "My God!": A Feminist Critique of the Excited Utterance Exception to the Hearsay Rule, 85 CALIF. L. REV. 159 (1997) (excited utterance exception reflects male perspective; calm statements of sexual assault victims should also be admissible, subject to conditions).

party admission doctrine,²³² use of inconsistent statements to impeach,²³³ and the increased rigor *Daubert* demands of expert testimony.²³⁴

A number of feminist evidence scholars have classified feminist legal theory into three categories, noting that scholars from different strands of feminism might have

²³² Aviva Orenstein, *Apology Excepted: Incorporating a Feminist Analysis into Evidence Policy Where You Would Least Expect It*, 28 SW. U. L. REV. 221 (1999) (apologies should, like offers of compromise and remedial measures, be protected from use against the apologizer). The role of apologies has become a topic of considerable interest at the intersection of evidence and torts. *See, e.g.*, Jennifer K. Robbennolt, *Apologies and Legal Settlement: An Empirical Examination*, 102 MICHIGAN L. REV. 460 (2004).

²³³ In two articles, Kim Lane Schepple has argued that the law gives too much credence to initial stories to the derogation of revised versions of events, one example being the use of prior statements to impeach. *See* Kim Lane Schepple, *Just the Facts Ma'am: Sexualized Violence, Evidentiary Habits, and the Revision of Truth*, 37 N.Y.L. SCH. L. REV. 123 (1992) and Kim Lane Schepple, *The Ground-Zero Theory of Evidence*, 49 HASTINGS L.J. 321 (1998).

²³⁴ *See* Andrew E. Taslitz, *A Feminist Approach to Social Science Evidence: Foundations*, 5 MICH. J. OF GENDER & LAW 1 (1998); Andrew E. Taslitz, *What Feminism has to Offer Evidence Law*, 28 SOUTHWESTERN U. L. REV. 171, 211 (1999) (favoring an approach to expert testimony that allows "all relevant voices and perspectives" to be heard; "For example, informing the jury about weaknesses in certain social science evidence and allowing it to be admitted subject to certain reliability safeguards may often be preferable to a flat exclusionary rule if the social science helps convey an excluded group's voice").

Some feminist scholars question the methodological rigor to which good science aspires and they advocate – to greater or lesser degrees – more intuitive, contextual, narrative, and relational paths to knowledge. One of the more extreme statements of this view is offered by Catherine MacKinnon:

If feminism is a critique of the objective standpoint as male, then we also disavow standard scientific norms as the adequacy criteria for our theory, because the objective standpoint we criticize is the posture of science. In other words, our critique of the objective standpoint as male is a critique of science as a specifically male approach to knowledge. With it, we reject male criteria for verification.

CATHERINE A. MACKINNON, *FEMINISM UNMODIFIED* (1987), at 50. One can also find traces of an ascientific attitude in articles that assess the validity or social value of battered womens' testimony on grounds other than standard measures of scientific validity. *See* Martha R. Mahoney, *Legal Images of Battered Women: Redefining the Issue of Separation*, 90 MICH. L. REV. 1, 42 (1991). *Cf.* Kit Kinports, *Evidence Engendered*, 1991 ILL. L. REV. 413, 445-47 (1991).

One need not look far for examples of errors resulting from poor research design that have had profoundly detrimental effects on women. A recent example is provided by developments in knowledge about hormone replacement therapy. Less rigorous research design (observational) led to the belief that estrogen replacement was beneficial. Better research design (experimental) led to the conclusion that such treatment was dangerous, and that over the decades of its use had caused women tens of thousands of unnecessary breast cancers, heart attacks, and strokes. JERRY AVORN, *POWERFUL MEDICINES: THE BENEFITS RISKS AND COSTS OF PRESCRIPTION DRUGS* (2004).

different views about law.²³⁵ The first is liberal feminism, sometimes dubbed “sameness” feminism. Liberal feminists are said to favor formal equality with men, assuming that women can compete in the same way as men, and not emphasizing differences between men and women. Radical feminists, sometimes called “dominance” feminists, see patriarchy and male domination of women, especially sexual domination, as the key to understanding modern society and its laws. A third strain of feminism, sometimes called “difference” feminism or “cultural” feminism, notes differences between men and women and argues for the legitimacy or superiority of the female perspective. These authors express a wide variety of views, some of them shared with strains of critical legal theory or postmodernism. These views include valuing intuition and emotion; believing in contextualized thinking instead of abstract, rigid rules; and emphasizing the value of relationships over market-like competition.²³⁶

In some writers, one also sees an emphasis on narratives over, or as an equal partner with, quantitative social science.²³⁷ Others have reservations about the public-private distinction, arguing that it has been used to shield the abuse of women by treating what goes on in the home as a “private” matter, beyond the reach of the law.²³⁸ There is also more attention to the law’s constitutive role in helping define society and perpetuate views of the world.²³⁹ Sometimes one also sees a belief in the pervasiveness

²³⁵ See Seymore, *supra* note __, at 1066-67 (dividing the feminist legal movement into “liberal feminism,” “relational feminism,” and “radical feminism”). See generally Andrew E. Taslitz, *What Feminism Has to Offer Evidence Law*, 28 SW. U. L. REV. 171 (1999). Cf. Orenstein, *Apology Excepted*, *supra* note __, at 226-27 (difference feminism, dominance feminism, and postmodern feminism).

²³⁶ Kinports summarized some of these differences as follows:

Men tend to value autonomy, abstract reasoning, individual rights, hierarchical organization, and detachment from others, [feminists] said, whereas women are more likely to value relationships, contextual reasoning, interdependence, and connection and responsibility to others.

Kit Kinports, *Evidence and Procedure for the Future: Evidence Engendered*, 1991 U. ILL. L. REV. 413, 417 (1991). See generally Taslitz, *supra* note __.

²³⁷ As an example, see generally Martha R. Mahoney, *Legal Images of Battered Women: Redefining the Issue of Separation*, 90 MICH. L. REV. 1 (1991).

²³⁸ “Feminists have long realized that the absence of the state, of law, from the private sphere has itself contributed to male dominance and female subordination. . . . The battered women’s movement has been, in the past twenty years, enormously successful in bringing the ‘private’ problem of wife abuse to public attention. . . .” Seymore, *supra* note __ at 1071.

²³⁹ See, e.g., Orenstein, *No Bad Men*, *supra* note __, at 692 (opposing Rule 413, in part because it perpetuates misconception that rape is not pervasive); Baker, *supra* note __, at 591 (discussing whether

and inevitability of politics, and a result-oriented approach that deprecates the wisdom or feasibility of achieving objective solutions to social problems.²⁴⁰ The “cultural” strand of feminism seems to present the most obvious challenge to what William Twining called the tradition of “optimistic rationalism” in evidence scholarship.²⁴¹

The influence of this strand of feminism on specific doctrinal reform is likely to be limited until the rest of the world changes, however, because of a tension between the results of a conflict resolution model based on cultural feminism and the more immediate goal of fair outcomes for women litigants. An emphasis on contextualized decision-making that explores all of the relational nuances of a situation militates against having fixed rules, and in favor of discretionary decision-making. But if decision-makers have attitudes tainted by sexism, then fixed rules are needed to protect women, since discretion will be exercised against them. Hence it is helpful to have as rigid a rule as possible against admission of the complainant’s sexual history in rape cases. Similarly, the hierarchical and competitive features of the adversarial system might be softened by feminist influences into a proof-taking system that encouraged cooperation and mediation.

Moreover, evidence law is adjective law, and it is hard to predict what substantive effect a particular evidence proposal will have. For example, an approach toward tough screening of “junk science” supported by large manufacturers seeking to escape liability in products liability cases can have the unexpected effect of hurting the prosecution in criminal cases. Similarly, liberal admission of rape trauma syndrome testimony can backfire when defendants in rape cases want to put in evidence of absence of symptoms to support the conclusion that the complainant was not raped. Thus, it may be harder for a substantive agenda, such as that of liberal feminism or dominance feminism, to be reflected in evidence law, because one can see ahead of time that it might cut both ways.

Perhaps for that reason, the areas in which reforms advocated by feminists, and by others concerned with fair treatment of women, seem to have been most effective and widely accepted are those in which the beneficial impact of the reform in helping women is predictable because women are disproportionately the victims of a particular crime. These include “rape shield” statutes protecting sexual assault victims from

Rule 413 spreads or diminishes rape myths, and concluding that it perpetuates a stereotype of the chronic rapist); Taslitz, *supra* note __, at 180; Marilyn MacCrimmon, *The Social Construction of Reality and the Rules of Evidence*, 25 U. BRIT. COLUM. L. REV. 36 (1991); Mahoney, *supra* note __, at 7 (noting interaction of law and culture and advocating “separation assault” as a legal concept that can “reshape cultural understanding”).

²⁴⁰ See Taslitz, *supra* note __, at 213-17.

²⁴¹ See WILLIAM TWINING, *RETHINKING EVIDENCE: EXPLORATORY ESSAYS* 32-91 (1990).

revelation of sexual history and exceptions to the character evidence rule for prior crimes committed by the accused in sex crime and domestic violence cases.

Much of the feminist writing on evidence is consistent with all three of the above-described strands of feminism, including liberal feminism, and hence its method is not too different from what might be obtained by a conventional legal analyst concerned with fair treatment of women. But sometimes dominance feminism or cultural feminism seems to have led in directions that a conventional legal analyst might not follow. Without attempting a comprehensive review of the literature, we will refer to four leading articles in which dominance feminism or cultural feminism seems to have been particularly influential on the analysis or result.

Our first examples are two articles about Rule 413, which allows evidence that the accused committed other sexual assaults to be admitted in a sexual assault prosecution. Aviva Orenstein and Katherine Baker, working independently, both came to the conclusion that the legislation was unwise and unjustifiable. Strains of cultural feminism and dominance feminism can be seen in both works. The legislation is seen as unwise because it decontextualizes the situation by stereotyping rapists, treating them as pathological outlaws rather than normal men engaged in situational conduct. It is also seen as perpetuating rape “myths” about rape being strange and deviant, whereas the authors see it as common and widespread, and as a way that society controls women. Finally, because rape is so common, men who have raped do not particularly stand out from other men, and hence rape has less probative force than would be the case were it a rare phenomenon.²⁴²

Another example of scholarship that seems influenced by strands of cultural feminism is Martha Mahoney’s article about spousal abuse cases.²⁴³ She addresses the question why women do not leave abusive relationships. Much of her article consists of narratives, including her own, of battered women.²⁴⁴ She finds nothing to criticize about Lenore Walker’s methodology, though she has mixed feelings about the message sent

²⁴² Orenstein, *No Bad Men*, *supra* note __, at 692; Baker, *supra* note __. The rape statistics are controversial. See Roger C. Park, *Character at the Crossroads*, 49 HASTINGS L. J. 717, 765-69 (1998); Neil Gilbert, *Miscounting Social Ills: Sexual Assault and Advocacy Research*, WELFARE JUSTICE 84-129 (1995).

²⁴³ Mahoney, *supra* note __.

²⁴⁴ “[T]his article offers narratives and poems from the lives of survivors of domestic violence, and a few from the stories of non-survivors, as part of its analysis and argument. Seven women’s stories have come to me through their own accounts. Five of these have at some time identified themselves as battered women. . . . The other women’s voices in this paper are drawn from identified published sources. . . . One of these stories is my own.” *Id.* at 7.

by learned helplessness,²⁴⁵ and advocates using the concept of “separation assault” as being more central to understanding why the battered woman, who is likely to be assaulted upon separating, does not leave the abusive relationship.²⁴⁶

Reliance on narratives is sometimes said to be a characteristic of feminist writing, and Mahoney’s article is a good example of extensive use of narratives. She uses stories from acquaintances, from the facts of reported cases, and from other published sources as social fact evidence to support her views of domestic violence. While she does not reject quantitative social science, she seems to have an attitude toward proof of social facts that would differ from that of many social scientists and *Daubert*-era evidence experts. She is indifferent to defects in the methodology of Walker’s battered woman research, assessing the theory primarily in terms of its utility in telling a story of oppression and the countervailing danger that the story may degrade women. She uses anecdotal evidence extensively, and regards it as a source of convincing proof instead of a source of hypotheses to be tested. Stories from the author’s acquaintances and from reported cases would be viewed with suspicion by many social scientists, because of the small sample size and obvious problems of selection bias. (For one thing, the typicality of facts stated in reported cases is highly suspect, because trial of a question of fact is itself an aberration; the typical dispute never reaches the legal system, much less survive to trial and appeal.) Narratives and counter-narratives can be produced on almost any issue. It is not clear that feminists rely on anecdotal evidence more than the

²⁴⁵ “I do not mean to criticize here the psychological theory underlying battered woman syndrome, or even the particular theory of learned helplessness. First, the collection of experience and perception summed up in battered woman syndrome are descriptively true of many women. [Here Mahoney cites two amicus briefs.] Lenore Walker’s defense of expert testimony is also correct: it helps women’s stories be brought into court by bringing together fragments that women experience as part of a whole relationship. [Citation omitted.] Finally, I would not choose to discard such a major tool in the effort to explain women’s experience in court, just because it has proved vulnerable to distortion in culture and law – we need more, not less, explanation. However, as long as explanation emphasizes ‘helplessness’ in the psychology of individual women, it runs into the danger of contributing to stereotyping.” *Id.* at 42. For an argument that the admission of BWS evidence reflects a political judgment in favor of battered women who kill rather than a judgment about the scientific validity of the syndrome, and that recognition that the evidence is admitted in pursuit of a “solution to a specific social problem rather than entirely neutral applications of basic evidentiary principles” may “inhibit the expansion of the principles developed in battered woman self-defense cases to more problematic situations” see Robert P. Mosteller, *Syndromes and Politics in Criminal Trials and Evidence Law*, 46 DUKE L.J. 461, 478-515 (1996)

²⁴⁶ “Because of the interactive relationships between law and culture in this area, law reform requires such an approach to simultaneously reshape cultural understanding. Separation assault is particularly easy to grasp because it responds to prevailing cultural and legal inquiry (‘why didn’t she leave’) with a twist emphasizing the batterer’s violent quest for control.” *Id.* at 7.

conventional fireside policy analyst, which includes most lawyers, though feminist writers seem more ready to reveal their reliance on it and even revel in it. In some cases, this may be due to suspicion of the motives and funding sources of social scientists,²⁴⁷ or simply to belief in the power of general feminist theory to guide the way in deciding which narratives to believe.

The final example is an article by Kim Lane Scheppelle entitled *Just the Facts, Ma'am: Sexualized Violence, Evidentiary Habits, and the Revision of Truth*.²⁴⁸ This article, using the Anita Hill hearings as its point of departure, posits that impeachment of witnesses by showing that they delayed in making a complaint or that they gave inconsistent versions of a complaint at different times hurts women disproportionately.²⁴⁹ In cases of sexualized violence, such as rape, sexual harassment, or spousal abuse, victims often delay in reporting and do not report a full or accurate version the first time.²⁵⁰ They then suffer when impeached with evidence of delay in reporting or of prior inconsistent statements. But victims have legitimate reasons for delay and revision. They are traumatized by the attack. Or they fear being criticized by the dominant culture for being a victim, for example, for provoking the attack.²⁵¹ Or they do not perceive the full implications of, say, sexual harassment until after they have thought about it or perhaps had therapy. Instead, they may present initial accounts that try to repair relationships and make things normal again.²⁵²

In addition to her points that are specific to women and sexualized violence, Scheppelle has a more general point rooted in postmodern epistemology. Invoking Wittgenstein, she notes that accounts of events (“stories”) are narratives that are influenced by interpretive frameworks. A woman who interprets her husband’s violence toward her as expressing his love does not see an event of battering the same way that a feminist lawyer sees it. But the difference in accounts that the feminist lawyer and the battered woman would give is not, Scheppelle writes, a difference “between truth and falsehood,” but a difference in interpretive frameworks. Consciousness raising may cause the same person to see the same event in different ways. The second

²⁴⁷ See *id.* at 27-28 (noting “split between social scientists and feminist activists on domestic violence issues”; complaining about “gender-neutral approach” of some social scientists; arguing that funding sources have affected how issues were explored and what research was done, and that “conscious use of feminist methodology in research is rare.”)

²⁴⁸ Scheppelle, *Just the Facts, Ma'am*, *supra* note ____.

²⁴⁹ *Id.* at 128.

²⁵⁰ *Id.* at 138-39.

²⁵¹ *Id.* at 142-43.

²⁵² *Id.* at 138-41.

interpretation may be better, just as a revised paper is better than the original. Understanding how accounts of facts are “socially constituted” is necessary to liberate women from sexualized violence:

[F]act-finders need to understand that early narratives about sexualized violence may reveal not some deeper truth, but rather the effects of oppression on women. Not allowing women to reinterpret their own experiences as they learn to oppose the abuse is a way of furthering that oppression.²⁵³

Scheppele’s solution for this perceived unfairness is not entirely clear. In her first article on the subject she complained about judicial “exclusion” of revised accounts, which would be manifestly unfair treatment if “exclusion” meant exclusion from evidence.²⁵⁴ The revised accounts are not excluded from evidence, of course, but merely subject to impeachment by prior inconsistent statements.²⁵⁵ Though an operational solution is not offered in any detail, Scheppele’s general message is clear – that “much more sympathy and belief” should be given the revised stories, even when they contradict what was said at the time of the events that led to the litigation.²⁵⁶ Though she relies on examples from cases involving sexual violence and harassment, she seems to call for application of her perspective to all kinds of cases.

The importance of freshness of memory is a psychological insight that is not often questioned, and Scheppele is effective in making us think twice about it. However, she could have done more to combine feminist social science with other studies of perception and memory. In her principal article, Scheppele virtually ignores the extensive body of literature on eyewitness testimony, except for one unexplained cite to Elizabeth Loftus.²⁵⁷ Some of the psychology scholarship on eyewitnesses would be helpful to her argument – studies suggesting, for example, that consistency in description of suspects

²⁵³ *Id.* at 172.

²⁵⁴ “Courts’ exclusion of revised stories works disproportionately against women because women are disproportionately the victims of a socialization that masks the immediate recognition of sexualized abuse *as* abuse.” *Id.* at 169-70.

²⁵⁵ Scheppele recognizes this point in a later article on the same subject, in which she says that, “What I have argued so far is that the Federal Rules of Evidence show a strong preference for acquiring information as close in time and space to the events in issue as possible. This does not mean that all other information is excluded. Certainly not. It means, however, that whatever is said and done at the time of the trouble will always have a place in the evidence that must be considered.” Scheppele, *Ground-Zero*, *supra* note ___, at 330.

²⁵⁶ *Id.* at 334.

²⁵⁷ Scheppele, *Just the Facts, Ma’am*, *supra* note ___, at 167 n.174.

is not a strong predictor of accuracy.²⁵⁸ Other parts of this body of scholarship would not strengthen Scheppele's argument – studies of the contaminating effect of post-event information and suggestive interviewing are examples.²⁵⁹

In summary, feminist evidence scholarship sometimes reflects an attitude that favors qualitative anecdotal data (such as narratives) over systematic quantitative analysis, that views science as irredeemably political, or that at least views general feminist theory as a better guide to social facts than the sorts of expertise and data generally favored by scientists and by *Daubert*. For example, Professor Taslitz opposes using a “flat exclusionary rule” for evidence which “helps to convey an excluded group’s voice,” even if the evidence fails more conventional criteria such as whether the methodology that produces the evidence is valid by the conventions of science.²⁶⁰ This attitude may explain why discussions of battered women’s syndrome sometimes seem result-oriented, as if the portrait that the theory paints of women is more important in judging its acceptability than the validity of the research methods that produced it.

V. Economics and Evidence

Until quite recently it could be said that law and economics scholars had virtually nothing to say about evidence law.²⁶¹ This is not exactly surprising. Evidence scholarship is focused on understanding, explaining, evaluating, and suggesting improvements for rules that are concerned principally with the goal of maximizing the ability of trials to

²⁵⁸ See BRIAN L. CUTLER & STEVEN D. PENROD, *MISTAKEN IDENTIFICATION: THE EYEWITNESS, PSYCHOLOGY, AND THE LAW* 93-95 (1995).

²⁵⁹ In the latter article, Scheppele seems more receptive to the idea that reflection can contaminate memories, and is somewhat broader in her citation of social science literature. In Scheppele, *Ground-Zero*, *supra* note ___, at 325-26, she notes psychological experiments on the decay of memory, writing that “subjects in experiments often show that their memories can in fact be predictably altered by the introduction of new information, and they unproblematically (even unconsciously) take into account the new information as if it were part of the original memory. . . .” This, of course, is a very good reason to distrust factual accounts given after there has been some time for reflection, for time affords possibilities for distortion even if one assumes that reflection itself does not distort.

²⁶⁰ Andrew E. Taslitz, *What Feminism has to Offer Evidence Law*, 28 SOUTHWESTERN U. L. REV. 171, 211(1999)(passage quoted n. ___, *supra*)

²⁶¹ See Roger C. Park, *Evidence Scholarship: Old and New*, 75 MINN. L. REV. 849, 849 n. 2 (1991). Some might say that Jeremy Bentham was an exception. But Bentham was a long time ago; his *RATIONALE OF JUDICIAL EVIDENCE* was published in 1827. He was less an economist than a utilitarian philosopher (and cognitivist). His solution to any problem of evidence was to abolish rules and rely on the discretion of judge and jury. And his work has been all but ignored by evidence scholars of the 20th Century. See WILLIAM TWINING, *THEORIES OF EVIDENCE: BENTHAM AND WIGMORE* (1985).

discover the truth of a matter in dispute.²⁶² Various fields might contribute to this endeavor: logic, cognitive science, psychology, philosophy, feminism, statistics, and so on. But economics? The first great law and economics movement, of the late nineteenth century,²⁶³ involved the macroeconomics of law. It was concerned with political economy, the behavior of markets, economic systems; and was reflected in such law as that of anti-trust, taxation, and banking regulation. The second great law and economics movement – the one with which readers of this article will be more familiar, involved the microeconomics of law, the pursuit of efficiency and wealth maximization. It has been concerned with the effects on individual behavior of varying incentive structures, and has been reflected in economic analyses of torts, contracts, property, and criminal law. What could marginalism or wealth maximization have to do with the truth-seeking goals of evidence law?

Only recently has there been a broad-gaged attempt to apply microeconomics to evidence law. To be sure, there were occasional articles on incentives and disincentives for gathering evidence,²⁶⁴ how those incentives affect the evidence received by courts,²⁶⁵ the resulting outcomes of trials and the impact of those outcomes on behavior

²⁶² See, e.g., William Twining, *Evidence and Legal Theory*, 47 MOD. L. REV. 261, 272 (1984):

The most striking feature of [evidence scholarship] is how homogeneous it is. Nearly all of the Anglo-American writers from Gilbert to Cross have shared essentially the same basic assumptions about the nature and ends of adjudication and about what is involved in proving facts in this context.... It can be re-stated simply in some such terms as these: the primary end of adjudication is rectitude of decision, that is the correct application of rules of substantive law to facts that have been proved to an agreed standard of truth or probability. The pursuit of truth in adjudication must at times give way to other values and purposes, such as the preservation of state security or of family confidences; disagreements may arise as to what priority to give to rectitude of decision as a social value and to the nature and scope of certain competing values.... But the end of the enterprise is clear: the establishment of truth.

See other citations and quotations in Chris Sanchirico, *Character Evidence and the Object of Trial*, 101 COLUM. L. REV. 1227 (2001).

²⁶³ See Herbert Hovenkamp, *The First Great Law and Economics Movement*, 42 STAN. L. REV. 993 (1990).

²⁶⁴ Andrew F. Daughety and Jennifer F. Reinganum, *On the Economics of Trials: Adversarial Process, Evidence, and Equilibrium Bias*, 16 J. L., ECON. AND ORG. 365 (2000).

²⁶⁵ Jeffrey S. Parker, *Daubert's Debut: the Supreme Court, the Economics of Scientific Evidence and the Adversarial System*, 4 SUPREME COURT ECON. REV. 1 (1995).

(especially economic activity) outside of court.²⁶⁶ These focused mainly on the problems of assembling evidence for trials and the effects of verdicts. Recently, we have seen broader attempts to use apply economics to evidence law and the philosophy of evidence. We will focus on three articles which illustrate this new contribution to evidence scholarship.

Perhaps fittingly, the first broad major law and economics treatment of evidence law is by Richard Posner, tackling a wide range of evidence topics.²⁶⁷ In his article, Judge Posner's assumptions about rational planning often lead him to inferences about that evidence rules' strong *ex ante* effects. His perspective entails implicit assumptions about pervasive knowledge of the rules among the general population, and about the friction-free willingness of actors to change customary ways of doing things in order to obtain an advantage if they ever wind up in litigation, assumptions which sometimes seem unrealistic. As might be expected, Judge Posner also brings to his study of evidence law a sensitivity to costs, trade-offs, and substitutions.

Posner's rational choice, *ex ante* perspective, would not appear to be the best starting point when drawing inferences about the issues that are at the core of traditional American evidence law. Many rules, such as the hearsay and character evidence bans, have long been viewed as being tailored with cognitive biases in mind, and aim to control reasoning at trial rather than future primary conduct.²⁶⁸ They seek to protect against mistaken, unreasonable, or lawless interpretations of evidence by factfinders. In assessing the value of these rules, the psychological literature seems a more obvious starting point. But Posner is fearless in applying his perspective, even to seemingly unpromising topics such as character evidence.²⁶⁹

We will describe and discuss some of those ideas. First, some of the good ones.

In his discussion of search and seizure, Posner assesses the value of sanctions other than exclusion of evidence, such as damage remedies for illegal searches. Posner argues that if these alternative sanctions were effective, there would be evidentiary gain because the searches would not be made in the first place. Therefore,

²⁶⁶ David Crump, *Evidence, Economics, and Ethics: What Information Should Jurors Be Given to Determine the Amount of a Punitive-damage Award?*, 57 MD. L. REV. 174 (1998).

²⁶⁷ Richard A. Posner, *An Economic Approach to the Law of Evidence*, 51 STAN. L. REV. 1477 (1999).

²⁶⁸ But see the discussion, *infra*, of Sanchirico's work.

²⁶⁹ For more thorough critiques of Posner's article, see: Richard Lempert, *The Economic Analysis of Evidence Law: Common Sense on Stilts*, 87 VA. L. REV. 1619 (2001); Ronald J. Allen & Brian Leiter, *Naturalized Epistemology and The Law of Evidence*, 87 VA. L. REV. 1491 (2001); Roger C. Park, *Grand Perspectives on Evidence Law*, 87 VA. L. REV. 2055 (2001).

those who oppose *Mapp v. Ohio*²⁷⁰ ought to be arguing about the definition of illegal search rather than about the sanction.²⁷¹ Scholars from the law and economics perspective seem able to come up with that sort of realization much more readily than others.

Next, there is the famous Blue bus conundrum.²⁷² Suppose a plaintiff is negligently injured by a bus, but cannot determine what bus company owned the bus that hurt him. And suppose it can be learned that bus Company A runs 51 buses along the route where the accident occurred, while Company B runs 49 buses there. Should the “naked” statistical fact that the defendant Company A owns 51% of the buses be admissible and sufficient to establish by a preponderance of the evidence that one of the defendant’s buses caused the plaintiff’s injuries? Here is part of Posner’s analysis of this problem:

Suppose both parties do conduct a thorough investigation yet are unable to come up with any additional evidence bearing on the ownership of the bus. There is no longer a basis for suspicion that the plaintiff really believes that a bus owned by Company B hit him, or for punishing him for not having investigated more. The case may seem no different from any other one tried under the preponderance of the evidence standard in which the balance of probabilities tilts only slightly in favor of the plaintiff. But there is a difference. Suppose the legal system can identify an entire class of cases in which the balance of probabilities tilts as slightly in favor of the plaintiff as it does in the bus case. If there are 1000 such cases, then allowing them to be tried can be expected to yield 510 correct decisions (that is, 510 decisions in which the defendant was in fact the injurer) and 490 incorrect ones, while not allowing them to be tried can be expected to yield 490 correct decisions and 510 erroneous ones. The social benefits of the twenty additional correct decisions that allowing the 1000 cases to be tried would produce – benefits in more perfect deterrence of negligent accidents – would probably fall short of the social cost of 1000 trials.²⁷³

²⁷⁰ 367 U.S. 643 (1961).

²⁷¹ Posner, *supra* note __ at 1533.

²⁷² For commentary on the blue bus problem, see Daniel Shavero, *Statistical-Probability Evidence and the Appearance of Justice*, 103 HARV.L.REV. 530, (1989); Ronald J. Allen, *On the Significance of Batting Averages and Strikeout Totals: a Clarification of the "Naked Statistical Evidence" Debate, The Meaning of "Evidence," and the Requirement of Proof Beyond a Reasonable Doubt*, 65 TUL.L. REV. 1093 (1991); Craig R. Callen, *Adjudication and the Appearance of Statistical Evidence*, 65 Tul. L. Rev. 457 (1991); Daniel Shavero, *A Response to Professor Allen*, 65 TUL. L. REV. 1111 (1991); Daniel Shavero, *A Response to Professor Callen*, 65 TUL. L. REV. 499 (1991).

²⁷³ Posner, *supra* note __, at 1510.

Posner's analysis makes the useful policy point that admitting the evidence, and therefore allowing the trials to go forward, has the virtue of reaching more correct results, but there is a disproportionate social cost of that marginal improvement in accuracy.²⁷⁴

Now we turn to some bad ideas. Posner argues that if the bus case were allowed to go to the jury, bus company A would have a big incentive to be careful and bus company B would have little or no incentive. Over time, B would cause more accidents, though with fewer busses, since it would be less careful; moreover, a monopoly would eventually be created because A, burdened by higher liability costs, would withdraw from the route.²⁷⁵

Judge Posner seems to enjoy revving up his models and seeing where they go, and it is not clear that he is completely serious in his remark about the demise of bus company A.²⁷⁶ At any rate, the consequence predicted is speculative and that it fails to take account of likely changes in behavior. If the companies really reacted that strongly to the burden of having liability imposed on the basis of naked statistical evidence, then they might instead each reduce the number of busses in an attempt to have fewer than half, thus leading to a race to the bottom.²⁷⁷ But it seems more likely either that naked statistical cases would be so rare as not to affect conduct at all, or that if the companies did feel pressure to reduce the number of busses, that there would be countervailing incentives that lead to adjustments in conduct, such as use of safety measures, that would make it worthwhile for the dominant company to bear the litigation burden while continuing to operate. Consumers might even prefer the larger and safer company, paying a premium for its services. Only one thing is clear: the ex

²⁷⁴ On the other hand, if the analysis is correct, it would seem that the decision whether to allow the evidence would turn on what proportion of the Blue bus company's buses run on that route. At some point, the gain in accuracy becomes worth the administrative and transaction costs. So it would seem that a rule setting a higher threshold for admission of such evidence would be the efficient solution in such cases.

²⁷⁵ *Id.* at 1510.

²⁷⁶ The bus-monopoly comment occurs at one of two places in which Judge Posner puts an exclamation point after a speculative comment about *ex ante* effects. See Posner, *supra* n. ____, at 1510 (Posner's present comment that allowing naked statistical evidence to sustain "blue bus" verdict might lead to bus monopoly), 1532 (comment that abrogating attorney-client privilege might increase enrollment in law schools, because clients would seek to learn more about the law themselves) Professor Lempert has decoded this punctuation to mean that Judge Posner was joking. See Richard Lempert, *The Economic Analysis of Evidence Law: Common Sense on Stilts*, 87 VA. L. REV. 1619, 1671, 1690 (2001) If so, this is an unusual way to signal humor, and one is left puzzled about how to treat other passages that lacked exclamation points but also seemed far-fetched.

²⁷⁷ Allen & Leiter, *supra* note ____, at 1526 (citation omitted).

ante consequences of the rule, if they exist at all, are highly speculative and unpredictable.

Next, we look at questionable economic ideas about character evidence. Posner argues that prior-crime evidence “is only weakly probative, because repeat offenders are punished more heavily than first-time offenders in part precisely to offset any greater propensity to commit crimes that their previous convictions have revealed. If recidivists are punished severely enough, the propensity to commit a subsequent offense may be reduced to the same level as the propensity to commit a first offense.”²⁷⁸

This proposition – that previously convicted defendants, if punished severely enough, will not be any more likely to commit crime than persons with clean records – would seem to merit a look at the empirical evidence. It requires justification in view of data showing that previously convicted defendants are dozens or hundreds of times more likely to commit an offense than are persons chosen at random.²⁷⁹ It contradicts the empirically documented view that recidivism is not much affected by increases in severity of punishment.²⁸⁰ Just proposing new ideas based upon a rational choice model, under which potential offenders apparently make a reasonable assessment of the value of present gratification compared to future punishment,²⁸¹ can be positively misleading to

²⁷⁸ Posner *supra* note ____ at 1525.

²⁷⁹ See Roger C. Park, *Character at the Crossroads*, 49 HASTINGS L.J. 717, 758-63 (1998). Prediction systems using actuarial methods commonly use prior crimes as a predictor. See PETER HOFFMAN, PREDICTING CRIMINALITY, U.S. DEPT OF JUSTICE STUDY GUIDE (1988) and VERNON L. QUINSEY, GRANT T. HARRIS, MARNIE E. RICE & CATHERINE A. CORMIER, VIOLENT OFFENDERS: APPRAISING AND MANAGING RISK (1998). See also MICHAEL R. GOTTFREDSON & TRAVIS HIRSCHI, A GENERAL THEORY OF CRIME, *passim* and 107 (1990) (“[R]esearch regularly shows that the best predictor of crime is prior criminal behavior” and that the differences between people, with respect to the likelihood they will commit criminal acts, persist over time); JOHN MONAHAN, U.S. DEP’T OF HEALTH AND HUMAN SERVICES, THE CLINICAL PREDICTION OF VIOLENT BEHAVIOR 71-72 (1981) (“If there is one finding that overshadows all others in the area of prediction, it is that the probability of future crime increases with each prior criminal act.”)

²⁸⁰ “The figures suggest that recidivism rates cannot be affected by varying the severity of the punishment, at least within acceptable Limits.” Daniel Nagin, GENERAL DETERRENCE: A REVIEW OF THE EMPIRICAL EVIDENCE, NATIONAL ACADEMY OF SCIENCES, PANEL ON RESEARCH ON DETERRENT AND INCAPACITATIVE EFFECTS 96 (1978).

²⁸¹ It seems likely, for example, that persons who are prone to criminal acts are also prone to an unrealistically low estimate of the danger of getting caught and the cost of future punishment. See Michael R. Gottfredson & Travis Hirschi, A General Theory of Crime, *passim* and 107 (1990)

policymakers unless the scholar is willing to check his assumptions against potentially disconfirming data.²⁸²

Marital privilege is another area where Posner's analysis sometimes gets out of control, requiring more human foresight, knowledge and flexibility than is plausible. I doubt that even law professors consult the rules about marital privilege before confiding in their spouses or committing a crime, but Posner has ordinary people doing both.²⁸³

We agree with Allen and Leiter that Judge Posner's article is prone to rootless theorizing.²⁸⁴ We cannot say for certain that these are incorrect, flawed ideas. We can say that, if the reality of human behavior is important to evidence policy, and hypotheses are to be tested, then one has to have some idea about which hypotheses are plausible enough to be worth the effort. One could regard the passages discussed above as saying: I'm not asserting this is true, I'm just showing you where the model leads, I'm just throwing out ideas for you ordinary scholars to check out. But in deciding whether to seek empirical verification of ideas, one has to make some choices about what to test. Guidance could come from theory, analogous studies, fireside inductions from history and experience, and intuition. With those preliminary screening tools as our guide, not many of Posner's economic ideas about evidence law appear promising.

In contrast to Posner's shotgun approach to applying economic ideas to evidence law, the next two articles use a laser: they zero in on one doctrine and work it over extensively. The first reflects the characteristic concern of contemporary law and economics with the impact of incentives and disincentives on individual behavior by deliberately looking away from the fact-finding function of evidence rules to consider the arguable impact of that rule on (mostly) crime deterrence.²⁸⁵ More specifically, Professor Sanchirico argues that the rule prohibiting the use of character evidence offered to persuade a factfinder that (consistent with the trait of character) the person performed the act at issue is a rule that cannot be explained coherently or convincingly in terms of enhancing the accuracy of trial fact-finding. He reviews each of the major extant truth-focused explanations for the rule and argues that they are unconvincing, to wit: the limited probative value of character evidence, the strong tendency of the jury to overweight such evidence, the imposition of liability for the defendant's character rather than for the wrongful conduct charged, judicial efficiency, an effort to impel parties to produce more and better evidence directed toward the conduct at issue, and trial bias (a

²⁸² Park, *supra* note ___, at 2057-58.

²⁸³ Posner, *supra* note ___, at 1530-31.

²⁸⁴ Allen & Leiter, *supra* note ___ at 1521-27.

²⁸⁵ Sanchirico, *supra* note ___.

biased distribution of persons selected for prosecution which further reduces the inferential value of character evidence).²⁸⁶

Sanchirico then offers a new explanation for the rule, one rooted in the notion that trials in general, and this rule in particular, are devices for dispensing primary incentives. In order to explain the rule coherently, all we need do is analyze the character evidence ban in light of its ability to deter undesirable conduct, rather than its ability to lead factfinders closer to truth:

Character evidence... is one area in which the truth seeking approach and the primary incentives approach to trial point in very different directions. This Article makes use of that divergence to advance our understanding of both character evidence and trial. It demonstrates that many of the rules governing character evidence – so difficult to rationalize when trial is regarded as an isolated exercise in sorting out past events – fall easily into place when trial is viewed as but one component of the larger system by which the state regulates everyday out of court behavior. The Article draws from this stark disparity in explanatory power the important lesson that, despite most of what is said about the object of trial, our desire to find the truth is subordinate to our desire, in effect, to shape it through the provision of incentives.²⁸⁷

The essential argument of Sanchirico's theory about the character evidence ban is this: Character evidence has predictive and therefore probative value.²⁸⁸ But it has no incentive value: its presence or absence creates no incentive to refrain from proscribed acts. "Trace evidence,"²⁸⁹ on the other hand, generally comes into being by the commission of proscribed acts and generally does not when such acts are not performed. Thus, trace evidence has incentive value. By focusing on trace evidence, the law reinforces the disincentive to committing proscribed acts. Were character evidence permitted as evidence of conduct, the disincentive for performing proscribed acts would be dampened. Without a character evidence ban, a person with a "bad" character is in a "damned if you do and damned if you don't" situation. If he refrains from the proscribed

²⁸⁶ We need not delay with a critique of his critiques. Except to say that the traditional explanations do not strike us as so weak as Sanchirico asserts. Sanchirico, like most of us, holds rival theories to higher standards of validity than he holds his own. Interested readers will have to decide for themselves whether Sanchirico has reached the right verdict on each theory, including his own.

²⁸⁷ *Id.* at 1232.

²⁸⁸ No one disagrees with this proposition. Even those who regard character evidence as a weak predictor of behavior do not say that it has zero predictive value.

²⁸⁹ Evidence left by the commission of an act, including the memories of eyewitnesses.

behavior, he still could be convicted of a relevant crime based on the evidence that he is more likely than others to commit such crimes. But with a character evidence ban in place, refraining from proscribed acts has greater power to prevent conviction (since the trace evidence necessary for conviction will not exist) and therefore the person is more likely to refrain. Thus, deterrence is stronger with the rule against character evidence than without it.

Sanchirico's theory is plausible in accounting for the relevant law, at least to the extent that the assumptions are valid. What is less plausible is that judges, facing an issue of whether evidence should be admitted or excluded from trial, presented with arguments from counsel about accuracy versus fairness, and despite writing opinions which consider the problem in those terms, are nevertheless, somehow, solving the dilemma with entirely different goals in mind, and that the most influential (if not most) judges were collectively engaged in the unexpressed enterprise (talk about in-court accuracy versus fairness but adopt rules calculated for out-of-court deterrence). Not impossible, but surely improbable.

Sanchirico uses the basic incentivist notion to explain the wisdom of an array of related rules: the admissibility of character evidence for impeachment, at sentencing, for punitive damages, exceptions to the character evidence rule that make sense from a deterrence perspective, and exceptions to the character evidence rule that do not make sense from a deterrence perspective.²⁹⁰

The article illustrates an important value of a "new" discipline's meta-theoretical imperatives. Because it is the concern of microeconomic analysis to search for incentives and consider their effects, Sanchirico was led to look away from the trial's apparent internal quest (for true facts) and look instead to the effects of the rule outside of the trial; from seeing the rule as part of a backward-looking search for truth to seeing it as a forward-looking tool of social control.

Although Sanchirico focuses on one rule, he argues that his analysis is illustrative of a larger truth about trials: they are more concerned with "influencing what happens in the future than discovering what happened in the past."²⁹¹ That is a bold departure from (and challenge to) the heretofore nearly unanimous evidence scholarship of the past several hundred years.²⁹² Whether it can be shown to be true we will have to wait and

²⁹⁰ The most prominent example is the relatively recent exceptions admitting evidence of an accused's propensity to engage in proscribed sexual conduct. According to the economic analysis of this article, these rules will increase rather than decrease the disincentive to commit such crimes.

²⁹¹ *Id.* at 1259.

²⁹² *See supra* note ____.

see. But who cannot be excited by the debates promised by so grand a claim, whether it ultimately proves to be true or false?

The final work of economic analysis of evidence doctrine that we will consider succeeds in combining traditional evidence scholarship's core concern of accurate factfinding with a "consequentialist, game-theoretic perspective."²⁹³ Professors Seidmann and Stein explore the question of whether the right of criminal defendants to remain silent benefits only guilty defendants or whether it creates conditions that assist factfinders in distinguishing innocent from guilty defendants and therefore benefits courts and society (by reducing the incidence of erroneous convictions). Critics of the right to silence dating back at least to Bentham have argued that innocent suspects and defendants do not need the right (they desire to offer true exculpatory evidence), and that only the guilty will avail themselves of it, thus increasing the incidence of erroneous acquittals. Defenders of the right to silence have defended it on a completely different plane, arguing that regardless of any costs in accuracy it may entail, the right is necessary to promote moral and ethical values of fair process.

Seidmann and Stein's economic analysis finds a defense of the right to silence that had been overlooked by both utilitarian critics and libertarian defenders. The critics of the right to silence made a series of economic miscalculations; the defenders may have turned to moral and ethical arguments because the erroneous reasoning of the critics appeared so persuasive. In any event, consequentialist game theory analysis has led to a new and utilitarian explanation for and defense of the right to silence.

To summarize: If they were compelled to submit to interrogations and to testify, guilty suspects and defendants would tell lies in an effort to avoid conviction. Whenever police or prosecutors are unable to expose those lies, the guilty would be indistinguishable from the innocent. Jurors and judges, aware that guilty (as well as innocent) defendants were offering exculpatory statements, would discount those statements, giving all of them less weight. With the right to silence, guilty suspects face the choice of telling lies that could be discovered (adding to the evidence against them) or exercising their right to silence. As more guilty defendants choose the option of remaining silent, they do not "pool with" innocent suspects and defendants. Consequently, innocent and guilty suspects and defendants become more distinguishable. Thus, the right to silence helps the innocent as well as the guilty.

Further analysis suggests that these effects are most likely to occur when the prosecution's evidence is moderately inculpatory (rather than weak or strong), and it works only when the standard of proof is "beyond a reasonable doubt" (rather than some lower threshold). Indeed, the authors argue that not only would a reduction in the

²⁹³ Daniel Seidmann and Alex Stein, *The Right to Silence Helps the Innocent: A Game-Theoretic Analysis of the Fifth Amendment Privilege*, 114 HARV. L. REV. 430 (2004).

standard of proof lead to more erroneous convictions, “but also serious indeterminacy in suspect identification and selection.”²⁹⁴

The authors argue that their analysis fits well with, and supports, or explains, not only the basic fifth amendment right of silence, but much of the jurisprudence that has grown up around it.²⁹⁵

The game theory analysis consists largely of thinking through the strategies of innocent and guilty suspects and defendants under varying conditions of evidence, standard of proof, and some other variables. These are plausible, reasonable arguments about what suspects and defendants and factfinders would do. But they rarely are

²⁹⁴ *Id.* at 449.

²⁹⁵ For examples: *Schmerber* (the right to silence protects only against compelled testimony, not against compelled production of physical evidence); *Jenkins-Fletcher* (permitting adverse inferences from pre-arrest silence); *Griffin* (prohibiting adverse inferences from refusal to testify); *Mitchell* (the right to silence applies in sentencing proceedings).

informed by empirical data about such behavior. They could be incorrect.²⁹⁶ This, of course, is typical of economic analysis of law.²⁹⁷

Uncharacteristically, Seidmann and Stein do offer a brief section which purports to test the implications of their theory against empirical data. Though the section seems to be an afterthought, and not much intellectual energy is put into it, at least it is there. But the data are tenuous²⁹⁸ and presented in a way that is confusing if not contradictory.

²⁹⁶ For examples: Regarding the assumption that innocent suspects tell exonerating truths: To what extent do innocent suspects lie also, in order to add a margin of safety to their factual innocence, only to get caught in the lie? (Doesn't that vitiate anti-pooling effects?) Regarding the assumption that suspects can assess the evidence against them: To what extent do police lie to suspects about the evidence they have? (If often, doesn't that make it difficult if not impossible for innocent as well as guilty suspects to evaluate what their "move" should be?) Regarding the assumption that only in the "rare" case, with abnormal people or abnormal circumstances, do the police so confuse or intimidate suspects that they cannot make rational calculations about their own best moves: To what extent do police interrogation techniques succeed in confusing or intimidating typical suspects into making foolish choices, including making inculpatory statements (which, after all, is exactly what they are designed to do)? Regarding the assumption that "the typical suspect confesses to crime only when confronted with evidence that he believes to be irrefutable." To what extent is this true? Will guilty subjects choose to continue to remain silent even if inferences based on silence are allowed – for example, because they have no convincing story, or because prior convictions will become admissible? If the guilty did speak instead of confessing or remaining silent, would their tales be as convincing as those of the innocent? See Gordon Van Kessel, *Quieting the Guilty and Acquitting the Innocent: A Close Look at a New Twist on the Right to Silence*, 35 INDIANA L. REV. 925, 956-60 (2002). Is it plausible to believe simultaneously that fact-finders obey the instruction not to draw adverse inferences from silence to an extent that encourages guilty defendants to remain silent, and also that the right to silence favors the innocent by making fact-finders more likely to believe their stories? See *Id.* at 942.

The answer the authors are likely to give is that so long as a plurality of the actual behavior is consistent with their assumptions, their theory still has predictive and explanatory value.

But it is worth reminding ourselves about the value of data. One illustration of the usefulness of combining data on actual behavior with game theory is provided by ROBERT AXELROD, *THE EVOLUTION OF COOPERATION* (1985). Axelrod conducted a game theory contest in which entrants submitted programs designed to elicit cooperative responses from the opponent in the game. Naturally, many entries were based on theories sans data. The winning entry, it turned out, was from a psychologist who knew from empirical experiments on game theory games what the most successful strategy is for eliciting cooperation from an opponent, and wrote a simple program to play that strategy. That winning strategy: tit-for-tat.

²⁹⁷ Though Posner has, in another context, urged that theoretical assumptions and hypotheses be empirically and rigorously tested with data. Richard Posner, *The Summary Jury Trial and Other Methods of Alternative dispute Resolution: Some Cautionary Observations*, 53 U. CHICAGO L. REV. 366 (1986).

²⁹⁸ Seidmann and Stein themselves declare half of the data to be useless: "too contaminated with measurement error (including inconsistent classification schemes) to draw any meaningful

From the data mish mosh they draw the conclusion that the two predictions they derive from their model are confirmed.

Perhaps the most telling finding (the import of which is not noted by Seidmann and Stein) is that few suspects refuse to answer police questions. Two American studies found 9.5% and 20.9%.²⁹⁹ British studies found about 10%, and the presence or absence of a prosecution right to argue adverse inferences changes that number only a few percent.³⁰⁰ If the actual figure is around 10%, and the claimed benefits are found only when the inculpatory evidence is moderate, then we are talking about a few percent of cases. If the vast majority of suspects talk even when they have a right to silence, then presumably most of them are offering the police lies which falsely tend to exonerate. Does that not create quite a lot of the very pooling that the article argues is prevented by the right to silence? But if so many (presumably) guilty suspects lie rather than avail themselves of their right to silence, then Bentham and his followers are also pretty far off the mark.

Which again suggests to us that economists of the law have not escaped the need to be more concerned about data. Economic models are not the royal road to truth, and need to be tested more earnestly.³⁰¹ Of course this is no different from most legal

conclusions....” *Supra* note ____, at 500.

²⁹⁹ *Id.* at n. 60.

³⁰⁰ *Id.* at 501.

³⁰¹ In the paper’s introduction, the authors dismissed the empirical approach:

A factual examination of these assumptions may follow two principal routes. One of these routes is empirical. By gathering and analyzing relevant empirical data, one can evaluate the workings of the right to silence without relying on sheer intuition. Such an approach might determine, statistically or by any other epistemologically plausible standard, whether the right aids only the guilty. The alternative route, which this Article follows, is behavioral modeling. Such modeling is usually, but not exclusively, based on rational-choice theory. Because reliable empirical evidence is often unavailable, the empirical approach is often problematic, as is the case with the factual assumptions examined by this Article. For example, it is extremely difficult, if not altogether impossible, to estimate the effect of the right to silence on the rate of true and false confessions. A suspect may confess to a crime for a variety of reasons. He may confess to a crime truthfully on finding the incriminating evidence irresistible. Alternatively, he may make a false or a truthful confession under the pressure of police questioning. *437 He may also make a false confession to exonerate the actual guilty party (for example, out of fear or love). A suspect deciding to remain silent during his interrogation may do so regardless of the right to silence: silence would be the best strategy for many guilty suspects even in the absence of a right.

An even greater problem inherent in the empirical approach lies in its limited

scholarship, and it would be unfair to lay any special criticism on the doorstep of Seidmann and Stein or of devotees of law and economics.³⁰²

A concluding comment on law and economics and evidence scholarship more generally, or perhaps a suggestion, is this. What we've seen so far usually uses a strong version of rational choice theory. But since evidence scholars are usually alert to the nonrational flaws in human reasoning, why not a weak version? Why not the bounded rationality of behavioral law and economics?³⁰³ Perhaps that will be the next turn.

Conclusion: The Interdisciplinary Future of Evidence Scholarship

Evidence scholarship in the leading law reviews is no longer predominantly doctrinal scholarship about rules. That is, it is no longer scholarship that focuses on analyzing and synthesizes rules, or on using skills taught in law school to argue that rules should be improved or reconceptualized.³⁰⁴ Although doctrinal scholarship lives on – indeed, flourishes – in treatises³⁰⁵ and in less prestigious law reviews,³⁰⁶ the best-

ability to produce determinate predictions when applied to human actions and decisions. There is no good reason to believe that uniformly observed actions and decisions will continue in the future. Reliance on statistical generalizations in forecasting human actions may prove perilous: recall Bertrand Russell's (in)famous chicken, conditioned to expect its daily feeding until the day the farmer interrupted this routine by butchering it for meat. One can make predictions about human actions only within some theoretical framework that imposes order on the empirically gathered facts. Generalizations about human actions acquire plausibility only by virtue of some explanatory theory that connects actions to reasons. Theoretical lenses may be microscopic or macroscopic, depending on the desired level of abstraction. In a search for a causal mechanism that explains numerous actions by their underlying motivations, theoretical lenses must be at a relatively high level of abstraction. This form of reductionism is necessary to tame "wild facts" and is, therefore, intrinsic to behavioral modeling. The compromised accuracy resulting from this reductionism is the price that any behavioral theory (and, perhaps, any theory) exacts in order to attain determinacy.

Id. at 436-437.

³⁰² Cf., Posner, *supra* note ____.

³⁰³ See Russell B. Korobkin and Thomas B. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*, 88 CALIF. L. REV. 1051 (2000).

³⁰⁴ For examples of doctrinal scholarship and comments on its usefulness, see Roger C. Park, *Evidence Scholarship, Old and New*, 75 MINN. L. REV. 849, 859-70 (1991)

³⁰⁵ A count supervised by one of the authors of this article indicates that there were 23 doctrinal texts and treatises on evidence law in print in 1957-58 (25,416 pages), compared to 207 texts and treatises in 2001-02 (150,833 pages), a page increase of 593%. (Casebooks and

known law reviews will probably continue to move away from doctrinal scholarship in favor of interdisciplinary work.

commercial outlines were not counted). The data collection was accomplished by identifying doctrinal evidence books that were listed in the 1957 and 1958 editions of LAW BOOKS IN PRINT (Glanville Publishers) and the 2001-02 edition of BOOKS IN PRINT (Bowker). Books that were unfamiliar to the principal investigator were judged by their titles. For the second period, page counting was not practical for some works, so the investigators estimated page length based on the mean length of the volumes for which data was available. These estimates counted for 26,232 of the 150,833 pages counted in the second period.

³⁰⁶See Michael Saks, Howard Larsen and Carol Hodne, *Is There a Growing Gap Among Law, Law Practice, and Legal Scholarship?: A Systematic Comparison of Law Review Articles One Generation Apart*, 30 SUFFOLK U. L. REV. 353 (1996). The authors found that there were nearly twice as many “practical” articles in 1986 as in 1960, a period during which the number of primary law reviews had nearly tripled. *Id.* at 373. The biggest change in content away from doctrinal scholarship occurred in the top quintile of law reviews (in terms of prestige as measured by the size of library). *Id.* at 374.

What is the future of interdisciplinary evidence scholarship? Let's consider some of the specific forms.

First, empirical studies, often the province of psychology research but not exclusively. The concern of some evidence scholars is that a variety of barriers will prevent this kind of interdisciplinary work from flourishing. These concerns are, among other things, that law professors have no training in empirical research and analysis, that even if they did they do not have students who are in a position to assist in carrying out such research, that funds are lacking to support such research, and that the traditions of single authorship and of grand theory scholarship militate against empirical work.³⁰⁷ In short, empirical research does not suit the scholarly habits of law professors – who rarely ventured beyond the law library, and now with so many resources online they need not even leave their offices – and support of various kinds is lacking for the doing of such research.³⁰⁸

There are two basic ways to overcome these sorts of problems. Evidence scholars who are sufficiently interested in testing evidence law assumptions empirically will either have to acquire the skills and the assistance to perform such work, or they will have to form collaborations with colleagues in other departments who already have those skills. Examples of both can be given. Availability of funding may be more of an obstacle. Federal funding agencies³⁰⁹ have not seen improvement in law or the trial system as a priority for the dollars they have to spend. Or perhaps there has been too little demand for those resources for the kinds of research needed to test and advance evidence doctrine. Lawyers usually are resourceful enough to overcome these kinds of difficulties, but there is no gainsaying such problems.

³⁰⁷ Professor Lawrence Friedman has written: "To begin with, empirical research is hard work, and lots of it; it is also nonlibrary research, and many law teachers are afraid of it; it calls for skills that most law teachers do not have; if it is at all elaborate, it is team research, and law teachers are not used to this kind of effort; often it requires hustling grant money from foundations or government agencies, and law teachers simply do not know how to do that. . . . Prestige is a factor too. Law schools . . . tend to exalt "theory" over applied research. Empirical research has an applied air to it, compared to "legal theory." Lawrence M. Friedman, *The Law and Society Movement*, 38 STAN. L. REV. 763, 774 (1986).

³⁰⁸ See Peter Schuck, *Why Don't Law Professors Do More Empirical Research?*, 39 J. LEGAL EDUC. 323 (1989).

³⁰⁹ The State Justice Institute, the National Institute of Justice. The National Science Foundation has been more supportive.

Richard Lempert raises a more specific methodological concern about empirical research in the field of evidence: “With the exception of some psychologists, few scholars have attempted to shed any empirical light on evidentiary issues. One reason for our lack of empirical knowledge is that it is hard to study the effects of evidence rules outside the laboratory, and laboratory studies raise substantial external validity problems.”³¹⁰ Whether the issue of generalizability is greater here than in other kinds of laboratory simulation research is not clear. And the problem sometimes exists and sometimes does not exist. It seems to us that numerous research questions can be studied without running into very serious questions of generalizability. One could do simulation studies about how, for example, people reason about evidence, such as studies of the prosecutor’s fallacy, the defense attorney’s fallacy,³¹¹ or Koehler’s experiments on how people reason about probability.³¹² Where the question is generalizability from the research participants, when the task is one requiring intellect and it is failed by undergraduates, one would think the failure results would generalize to lay juries. Other times, experiments can and have been done using jurors. In addition, one can do field studies on how juries reason about evidence. One example is a study of jury deliberations by Shari Diamond and Neil Vidmar, in which they found, by recording the deliberations, that jurors did not talk about using insurance coverage to find deep pockets, but rather they talked about whether the plaintiff had insurance that would provide a collateral source for compensation.³¹³

As more and more scholars and judges come to appreciate that without such research their rulemaking and their rulings are little more than guesses, we suspect that more ways will be found to conduct more of the necessary research.

Second, evidence law and forensic science (or science generally). This field will obviously keep flourishing with the scientization of society and the growing realization that much of what the courts had been accepting through the 20th Century was seriously inadequate. Moreover, *Daubert* might teach us something about the scientific method that will make us more able and willing to become involved in the empirical studies of evidence.

³¹⁰ Lempert, *supra* note ____ at 1709.

³¹¹ William C. Thompson & E.L. Schumann, *Interpretation of Statistical Evidence in Criminal Trials: the Prosecutor's Fallacy and the Defense Attorney's Fallacy*, 11 L. & Hum. Behav. 167 (1986).

³¹² See, e.g., Jonathan J. Koehler & Daniel Shaviro, *Veridical Verdicts: Increasing Verdict Accuracy Through the Use of Overtly Probabilistic Evidence and Methods*, 75 CORNELL L. REV. 75 (1990); Jonathan J. Koehler, *When Are People Persuaded by DNA Match Statistics?* 15 L. & HUM. BEHAV. 493 (2001).

³¹³ Shari Seidman Diamond and Neil Vidmar, *Jury Room Ruminations on Forbidden Topics*, 87 VA. L. REV. 1857, 1888-89 (2001).

Third, the “new evidence scholarship. Some topics seem to have run (or over-run) their course, such as the blue bus and gatecrasher problems. Topics involving the study of inference and decision-making have unrealized potential and will probably continue to be growth areas. For example, we expect to see further efforts to upgrade Wigmorean charting and further work on computer-aided pretrial fact analysis. Bayes will live on, both as an aid in thinking about evidence law and, among Bayesian enthusiasts, as a tool in fact analysis. But, because the decisions of humans (jurors and judges) are not well explained or predicted by Bayesian approaches, watch for the further incorporation of cognitive science and the arrival of artificial intelligence into these projects – to the extent that these new sciences provide more accurate predictions of how judges will rule and how jurors will infer. We also are unlikely to see Bayes' theorem in trials themselves, where many items of evidence are involved and an expert has to explain to the factfinder how to apply the theory.

Fourth, concerning evidence and feminism, we expect that this area will continue and grow to the same extent that an intellectual tradition regarded as distinctly feminist continues and grows. It may tend to concentrate on areas of evidence that are of special concern to women (sexual assault, obviously), but it will be useful both to the scholars and to the law of evidence if they move beyond those limited areas. There certainly is no reason why evidence law should not be as susceptible to continued feminist analysis as it is to analysis through the lens of any other field. In some ways, feminist legal analysis has the same advantage that traditional legal analysis and much law and economics analysis had, namely, that it is an armchair activity that can be carried on by taking a set of ideas and using them as a lens with which to examine the law.³¹⁴

Fifth, of Law and Economics, Richard Lempert sees little future. He writes:

[M]ajor funders of law and economics seem to have a pro-business, antiregulation, and/or generally conservative political agenda they wish to promote. Although the Olin Foundation's support of... intellectual activities that promote no coherent social agenda are contrary to my hypothesis, I still do not believe that evidence law will be a high priority for support among law and economics research funders. Not only does evidence law not deal with issues that are at the core of what funders hope to establish through economic research, but, as Posner points out, when the lamp of economics shines on evidence law, what it reveals is not necessarily compatible with conservative or big business political agendas. Posner, for example, argues that a law and economics perspective supports the institution of jury trial, a message that business supporters of law and economics are unlikely to relish.³¹⁵

³¹⁴ The contrast is to most kinds of empirical research, which require more qualitatively different activities.

³¹⁵ Lempert, *supra* note ____ at 1637-38 (citations omitted).

Lempert might be wrong, at least in the short term and at least concerning law and economics articles of the theoretical type (not law and economics articles involving empirical research). Theoretical law and economics articles do not require any more funding than doctrinal scholarship, and the elite law journals are fond of them. And there is no reason to think that scholars like Chris Sanchirico or Alex Stein will stop writing about it.