

EVIDENCE SCHOLARSHIP RECONSIDERED: RESULTS OF THE INTERDISCIPLINARY TURN

ROGER C. PARK*
MICHAEL J. SAKS**

Abstract: Evidence scholarship has developed a permanent interdisciplinary aspect, involving a variety of different disciplinary themes. These include: the psychology of witnesses and factfinders, forensic science, theories of probability and proof, feminist perspectives on evidence law, and the law and economics perspective. After first assessing the status of traditional doctrinal scholarship, we review each of the major interdisciplinary braids, compare them, and evaluate their relative contributions. We conclude by developing a thesis about the utility of different types of evidence scholarship, arguing that interdisciplinary evidence scholarship is more promising and useful to the extent that it helps to explain or advance the truth-seeking function of trials, rather than to posit or seek extrinsic effects of rules that traditionally have been understood as protecting the accuracy of verdicts.

INTRODUCTION

In this article, we examine the changing field of evidence scholarship, which has become decidedly interdisciplinary. The importance of these endeavors is bound to the importance of evidence law: the rules by which we adjudicate facts are as important as the interpretation of substantive law, and perhaps more so.¹ While each of

* Distinguished Professor and James Edgar Hervey Chair of Litigation, University of California Hastings College of the Law, J.D., Harvard Law School.

For their helpful comments on earlier drafts, the authors thank Professors Ronald Allen, Craig Callen, David Faigman, Daniel Farber, Richard Friedman, Steven Goldberg, David Kaye, Dale Nance, Aviva Orenstein, Jeffrey Rachlinski, Chris Sanchirico, Peter Tillers, and Joan Williams.

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

¹ See William Twining, *Recent Trends in Evidence Scholarship*, in PROCEEDINGS OF THE FIRST WORLD CONFERENCE ON NEW TRENDS IN CRIMINAL INVESTIGATION AND EVIDENCE 13, 15-16 (J.F. Nijboer & J.M. Reijntjes eds., 1997); WILLIAM TWINING, *Taking Facts Seriously*, in RETHINKING EVIDENCE: EXPLORATORY ESSAYS 12, 23 (1990); cf. John D. Jackson, *Modern Trends in Evidence Scholarship: Is All Rosy in the Garden?*, 21 QUINNIPIAC L. REV. 893,

those interdisciplinary domains has its own literature, wherein issues of relevance to that interdisciplinary intersection are addressed and debated, the present Article is the first to consider all of the major strains, compare them, and evaluate the comparative contribution of the different approaches.²

We start with an assessment of the status and value of traditional doctrinal scholarship (Part I). We then review interdisciplinary inquiries into the psychology of witnesses and factfinders (Part II), forensic science (Part III), theories of probability and proof (Part IV), the implications of feminism for evidence law (Part V), and the contributions of the law and economics perspective (Part VI). In the process, we develop a thesis about the utility of different types of evidence scholarship; namely, interdisciplinary evidence scholarship is more promising and useful to the extent that it helps to explain or advance the truth-seeking functions of trials, rather than to posit or seek extrinsic effects from rules that traditionally have been understood as protecting the accuracy of verdicts.³

I. DOCTRINAL SCHOLARSHIP ON EVIDENCE

Evidence scholarship has a distinguished history. It attracted one of the great minds of the nineteenth century, Jeremy Bentham, whose evidence writings, after being edited by John Stuart Mill, were published in 1827 in a five-volume treatise entitled *Rationale of Judicial Evidence*.⁴ Bentham's treatise urged radical utilitarian reform to the

906 (2003) (arguing that modern evidence scholarship has embraced interdisciplinary approaches within the rationalist tradition).

² For other treatments, see authorities cited *supra* note 1, and Ronald J. Allen & Brian Leiter, *Naturalized Epistemology and the Law of Evidence*, 87 VA. L. REV. 1491, 1493 (2001) (supporting the usefulness of social epistemology to the study of evidence); Richard Lempert, *The Economic Analysis of Evidence Law: Common Sense on Stilts*, 87 VA. L. REV. 1619, 1623 (2001) (evaluating the intersection of evidence and economics and dismissing the relevance of economics to the field of evidence); Roger C. Park, *Grand Perspectives on Evidence Law*, 87 VA. L. REV. 2055, 2079 (2001) (favoring empirical legal scholarship over legal epistemology or legal economics); Richard A. Posner, *Comment on Lempert on Posner*, 87 VA. L. REV. 1713, 1721 (2001) (supporting the importance of economics to evidence law).

³ In other words, we extend to the interdisciplinary realm Bentham's vision of trial fact finding. Bentham saw the "direct end" of legal procedure to be "rectitude of decision" and the "collateral ends" to be "the avoidance of unnecessary delay, vexation, and expense." 1 JEREMY BENTHAM, *RATIONALE OF JUDICIAL EVIDENCE* 34 (London, Hunt & Clark 1827). We agree with this proposition, though we recognize that in exceptional circumstances goals such as protection of privacy should be considered in framing evidence law. For a summary of Bentham's views, see WILLIAM TWINING, *THEORIES OF EVIDENCE: BENTHAM AND WIGMORE* 27-100 (1985).

⁴ BENTHAM, *supra* note 3.

misguided evidence rules of the time. Despite its prevailing tone of sarcasm and ridicule, Bentham's treatise seems to have been quite effective in speeding the abolition of its principal targets, such as rules disqualifying witnesses for interest.⁵ The early twentieth century brought an evidence treatise that was hailed by eminent scholars as the best written on any subject⁶—John Henry Wigmore's monumental and enormously influential *Evidence in Trials at Common Law* collected and systematized virtually all of the common law of evidence.⁷ Both works had great influence in resolving logical contradictions, making evidence doctrine more coherent, and reforming archaic rules that compelled judges to make senseless rulings that led to unjust results.

Despite the wide-ranging interests of Bentham and Wigmore, much of their writing on evidence consisted of what would now be called doctrinal scholarship. Doctrinal scholarship focuses on analyzing and synthesizing rules and urging that they be reconceptualized or otherwise improved.⁸ That is not to say that it is formal or purely analytic. Doctrinal scholarship long has had a prescriptive element. Its practitioners have sought to improve the law and have been concerned with the social consequences of law. But their policy analysis

⁵ 1 JOHN HENRY WIGMORE, *EVIDENCE IN TRIALS AT COMMON LAW* § 8, at 611 (Tillers rev. 1983). Wigmore stated:

Mature experience constantly inclines us to believe that the best results on human action are seldom accomplished by sarcasm and invective. . . . But Bentham's case must always stand out as a proof that sometimes the contrary is true—if conditions are meet. No one can say how long our law might have waited for regeneration if Bentham's diatribes had not lashed the legal community into a sense of its shortcomings.

Id.

⁶ Joseph Beale wrote, "It is hardly too much to say that this is the most complete and exhaustive treatise on a single branch of our law that has ever been written." J.H. Beale, Book Review, 18 HARV. L. REV. 478, 478 (1905) (reviewing JOHN HENRY WIGMORE, *A TREATISE ON THE SYSTEM OF EVIDENCE IN TRIALS AT COMMON LAW* (1905)). Of the third edition, Edmund Morgan wrote: "Not only is this the best, by far the best, treatise on the Law of Evidence, it is also the best work ever produced on any comparable division of Anglo-American Law." Edmund M. Morgan, Book Review, 20 B.U. L. REV. 776, 793 (1940) (footnote omitted) (reviewing JOHN HENRY WIGMORE, *A TREATISE ON THE SYSTEM OF EVIDENCE IN TRIALS AT COMMON LAW* (3d ed. 1940)). William Twining has noted that "one of the difficulties of debating with Wigmore was that, so great was his influence, once he had perpetrated a doctrine on the basis of little or no authority, precedents would soon follow to fill the gap." TWINING, *supra* note 3, at 111.

⁷ WIGMORE, *supra* note 5.

⁸ For examples of doctrinal scholarship and discussion of its usefulness, see Roger C. Park, *Evidence Scholarship, Old and New*, 75 MINN. L. REV. 849, 859–71 (1991).

has rested mainly on history, experience, and fireside inductions⁹—not on knowledge of the scholarly literatures of disciplines outside the law that address many of evidence law's concerns.

In leading law reviews, there has been a steep decline in doctrinal scholarship on evidence law. Examination of the top continuously-published journals shows a dramatic reduction in the proportion of pages devoted to such scholarship.¹⁰ At the turn of the twentieth century, doctrinal articles constituted 93% of the evidence articles in these journals. By mid-century that percentage had fallen to 79%, and by century's end only 20% of evidence articles were doctrinal. Doctrinal inquiries—that is, analyses of the rules themselves, their coherence, their organization, their emergence and disappearance—have been replaced by inquiries of other kinds.¹¹ These newer inquiries seek to cross law with some other discipline.

⁹ "Fireside inductions," a term coined by the psychologist Paul Meehl, describes inferences and deductions we draw from everyday experience, from introspection, from anecdotal evidence, and from culturally transmitted ideas, rather than from rigorous, systematic inquiry. Paul Meehl, *Law and the Fireside Inductions: Some Reflections of a Clinical Psychologist*, 27 J. Soc. Issues 65, 65–66 (1971).

¹⁰ 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 fifty 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 and 1996 disclosed ten law journals that appeared among the top twenty in both studies. See Scott Finet, *The Most Frequently Cited Law Reviews and Legal Periodicals*, 9 LEGAL REFERENCE SERVICES Q. 227, 229 tbl.1 (1989) (presenting material from Douglas B. Maggs, *Concerning the Extent to Which the Law Review Contributes to the Development of the Law*, 3 S. CAL. L. REV. 181 (1930)); James Lindgren & Daniel Seltzer, *The Most Prolific Law Professors and Faculties*, 71 CHI.-KENT L. REV. 781, 789 tbl.2 (1996). 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 *Minnesota Law Review*. Research assistants compiled a list of evidence articles from these journals at approximately fifty-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 determined not to be evidence law articles.)

¹¹ This certainly is the shared perception of evidence scholars themselves. See Richard Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof*, in *PROBABILITY AND INFERENCE IN THE LAW OF EVIDENCE: THE USES AND LIMITS OF BAYESIANISM* 61, 61–62 (Peter Tillers & Eric D. Green eds., 1988).

That does not mean that doctrinal scholarship is dying out. Though it is true that, with important exceptions,¹² it is becoming less common in elite journals, it flourishes more than ever in treatises¹³ and there is still ample space for it in journals due to the sheer proliferation of law reviews.¹⁴ Though the change in the proportion of doctrinal and interdisciplinary evidence scholarship in often-cited journals is striking, in absolute terms there is still plenty to go around, perhaps more pages than in earlier periods.

It is clear, however, that doctrinal evidence scholarship is less often published in elite journals, less respected among top scholars, and less rewarded as a career choice than in earlier times.¹⁵ What accounts for this shift in emphasis? There are several reasons. Some of them are not limited to evidence scholarship, but common to all forms of legal scholarship. For example, the realist perspective has triumphed, and hence there is more emphasis on considering consequences and discovering social facts. Law schools are more inclined to hire faculty with advanced degrees in other fields. There have been well-financed attempts to spread the law and economics perspective.¹⁶

¹² See generally Richard D. Friedman & Bridget McCormack, *Dial-In Testimony*, 150 U. PA. L. REV. 1171 (2002) (primarily doctrinal, with incidental citation to empirical work); Richard A. Nagareda, *Reconceiving the Right to Present Witnesses*, 97 MICH. L. REV. 1063 (1999) (doctrinal analysis of the right to present witnesses). See also the recent blossoming of doctrinal discussion about *Crawford v. Washington*, *infra* note 22.

¹³ A count supervised by one of the authors of this article indicates that there were twenty-three 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 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 were available. These estimates counted for 26,232 of the 150,833 pages counted in the second period.

¹⁴ See Michael J. Saks, Howard Larsen & Carol J. 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, 373–74 (1996). The authors found that there were nearly twice as many “practical” articles in 1985 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 the host law school’s library). *Id.* at 374.

¹⁵ See, e.g., George L. Priest, *Social Science Theory and Legal Education: The Law School as University*, 33 J. LEGAL EDUC. 437, 437 (1983) (treatise-like work no longer a credit to those “competing on the leading edge of legal thought”). See generally Richard A. Posner, *The Decline of Law as an Autonomous Discipline: 1962–1987*, 100 HARV. L. REV. 761 (1987).

¹⁶ Lempert, *supra* note 2, at 1636–37.

The decline of doctrinal scholarship on evidence also is related to particular features of evidence law. Doctrinal scholarship thrives when the law is plainly unjust, seriously confused, or rapidly changing. Evidence law is not as foolish as it was in Bentham's time nor as disorganized as it was in Wigmore's time. Moreover, on most topics, the climate for reform is not as good as it was in the eras of Wigmore and Bentham. The mildly radical reform attempted in Edmund Morgan's *Model Code of Evidence*¹⁷ failed completely, and today's judges and lawyers seem generally satisfied with the Federal Rules of Evidence, which were largely a codification of common-law rules extant in the mid-twentieth century. Finally, the confusion that existed in Wigmore's time and before has been largely tamed by the Federal Rules, so there is no enthusiasm for ground-breaking reclassifications. And, compared to other fields, doctrinal change is infrequent. The Advisory Committee has proven conservative in its role of proposing amendments. Nor have the evidence rules that are primarily focused on achieving accuracy attracted much congressional activity.¹⁸ Congress tends to have a substantive agenda, and even the rules of privilege, which attract the attention of interest groups and which are frequently the subject of state-level legislation, have not attracted much attention from Congress at the national level—at least not since the disputes in the 1970s over the content of the newly codified Federal Rules of Evidence.¹⁹

¹⁷ MODEL CODE OF EVIDENCE (1942). Edmund M. Morgan, Professor at Harvard Law School, was the Chief Reporter and wrote the foreword. Edmund M. Morgan, *Foreword to MODEL CODE OF EVIDENCE* 1-70 (1942).

¹⁸ An arguable exception has occurred in the area of expert testimony, where concern about the alleged "litigation explosion" and the perceived impacts of tort judgments against manufacturers on American competitiveness inspired efforts to codify and toughen the *Daubert* rule. See Nancy S. Farrell, Comment, *Congressional Action to Amend Federal Rule of Evidence 702: A Mischievous Attempt to Codify Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 13 J. CONTEMP. HEALTH L. & POL'Y 523, 543-51 (1997). This effort encouraged the Advisory Committee to the Federal Rules of Evidence to propose its own codification in what became the 2000 Amendment to Federal Rule of Evidence 702. See FED. R. EVID. 702 (amended 2000) advisory committee's note; Derek L. Mogck, Note, *Are We There Yet? Refining the Test for Expert Testimony Through Daubert, Kumho Tire and Proposed Federal Rule of Evidence 702*, 33 CONN. L. REV. 303, 324-25 (2000).

¹⁹ The primary exception is the Crime Bill of 1994, which made an exception to the ban on character evidence of other serious sex crimes by the alleged perpetrator in cases charging sexual assault or child molestation. Violent Crime Control and Law Enforcement Act of 1994, Pub. L. No. 103-322, § 320935, 108 Stat. 1796, 2136-37 (codified as FED. R. EVID. 413-415 at 28 U.S.C. app. (2000)). Whatever its merits, this change seems to have been driven by a substantive agenda of making it easier to convict sexual predators. See CHRISTOPHER B. MUELLER & LAIRD C. KIRKPATRICK, EVIDENCE 279-83 (3d ed. 2003). See generally Katharine K. Baker, *Once a Rapist? Motivational Evidence and Relevancy in Rape Law*,

The lack of doctrinal change has two effects. First, there are fewer new developments to explain and critique; over time it becomes harder to generate new ideas about old doctrine. Second, the fact that proposals for change tend to die discourages scholars from advocating change. This discouragement is important because modern doctrinal scholarship is prescriptive. While it was possible 100 years ago to publish an article that merely described and organized doctrine,²⁰ modern doctrinal scholarship invariably makes the case for reform and improvement.²¹ If reform is unlikely, there is less incentive to argue for improvements. When change does occur—as when the Supreme Court changed confrontation doctrine in its 2004 decision in *Crawford v. Washington*—an outpouring of doctrinal scholarship results.²²

The era of the great doctrinal analysts and treatise-writers can be seen as a continuation of Bentham's project of eliminating artificial distinctions and obstacles to free proof. Though current evidence law may have anomalies,²³ they are not as striking as they were in Bentham's time. One does not see judges straining against evidence rules that they themselves consider to be unjust.

110 HARV. L. REV. 563 (1997); James Joseph Duane, *The New Federal Rules of Evidence on Prior Acts of Accused Sex Offenders: A Poorly Drafted Version of a Very Bad Idea*, 157 F.R.D. 95 (1994); Aviva Orenstein, *No Bad Men! A Feminist Analysis of Character Evidence in Rape Trials*, 49 HASTINGS L.J. 663 (1998); Roger C. Park, *The Crime Bill of 1994 and the Law of Character Evidence: Congress Was Right About Consent Defense Cases*, 22 FORDHAM URB. L.J. 271 (1995).

²⁰ For an example of such an article, see generally David Torrance, *Evidence of Character in Civil and Criminal Proceedings*, 12 YALE L.J. 352 (1903).

²¹ See, e.g., Richard A. Nagareda, *Reconceiving the Right to Present Witnesses*, 97 MICH. L. REV. 1063, 1064–69 (1999).

²² A search using the term “Crawford v. Washington” yielded 204 citations in the LEXIS ALLREV database on February 13, 2006. The citations include a symposium in Volume 71 of the *Brooklyn Law Review*, introduced by Robert M. Pider, *Crawford and Beyond: Exploring the Future of the Confrontation Clause in Light of Its Past: Introduction*, 71 BROOK. L. REV. 1 (2005), and several other articles by law professors in well-recognized law reviews. See generally Stephanos Bibas, *Originalism and Formalism in Criminal Procedure: The Triumph of Justice Scalia, the Unlikely Friend of Criminal Defendants?*, 94 GEO. L.J. 183 (2005); Daniel J. Capra, *Amending the Hearsay Exception for Declarations Against Penal Interest in the Wake of Crawford*, 105 COLUM. L. REV. 2409 (2005); Tom Lininger, *Prosecuting Batterers After Crawford*, 91 VA. L. REV. 747 (2005); Miguel A. Méndez, *Crawford v. Washington: A Critique*, 57 STAN. L. REV. 569 (2004).

²³ The rules about impeaching the character of witnesses are an example. See Richard Friedman, *Character Impeachment Evidence: Psycho-Bayesian [?] Analysis and a Proposed Overhaul*, 38 UCLA L. REV. 637 *passim* (1991) (proposing revisions to the rules of character impeachment); H. Richard Uviller, *Evidence of Character to Prove Conduct: Illusion, Illogic, and Injustice in the Courtroom*, 130 U. PA. L. REV. 845, 867–68, 887–88 (1982) (arguing that prior convictions should be excluded).

Another possible reason that doctrinal evidence scholarship may decline is that the trial is becoming more of a rarity.²⁴ Evidence doctrine is most important in jury trials, where there are two decision-makers—one a referee who screens the evidence, and the other a factfinder that weighs the merits. But if jury trials (or trials in general) become less important, rules that have application mainly to the courtroom become correspondingly less important. This may contribute to the study of evidence as a topic, not of courtroom rules, but of how to determine the truth.

That is not to say that evidence rules always will remain static. Indeed, structural changes eventually might make the current rules obsolete. If Mirjan Damaška is right in thinking that the pillars of evidence law are crumbling because of structural changes—a drift away from the jury-centered time-concentrated adversarial model²⁵—then at some point doctrinal rejuvenation might re-open the door to doctrinal analysis. But if the rejuvenation comes, as Damaška predicted, because the pillars of the exclusionary rules deteriorate,²⁶ then the rules excluding evidence will themselves become less important, an effect that is sure to reduce the importance of scholarship describing them.

Doctrinal scholarship has some significant benefits. It produces material that can be easily understood by lawyers and judges on the basis of their law school training and that helps them in understanding and systematizing the law. Judge Harry Edwards's eloquent complaint that he finds little of use in elite law reviews suggests that journals may be losing readership among judges and policymakers, the very people who are in the best position to systematize and improve the law.²⁷ But this service to the profession also can be accomplished by producing interdisciplinary scholarship that is useful and accessible. And that seems to be what is happening.

On the whole, then, the increased amount and prestige of interdisciplinary scholarship is a welcome development because of the value of functional approaches to the analysis and criticism of law. Rules of law need to be assessed in light of their social impact. Light

²⁴ See Marc Galanter, *The Vanishing Trial: An Examination of Trials and Related Matters in Federal and State Courts*, 1 J. EMPIRICAL LEGAL STUD. 459, 459 (2004) (exhaustively reviewing the evidence on the trend toward fewer and fewer trials). See also the entire issue of the *Journal of Empirical Legal Studies* in which Galanter's article appears, which is devoted to the subject of the vanishing trial.

²⁵ See MIRJAN R. DAMAŠKA, EVIDENCE LAW ADRIFT 125–42 (1997).

²⁶ *Id.* at 149–52.

²⁷ See Harry T. Edwards, *The Growing Disjunction Between Legal Education and the Legal Profession*, 91 MICH. L. REV. 34, 35 (1992).

from other fields can be an aid in assessing the impact of law, but scholars in those other departments usually do not have the knowledge of legal doctrine and legal institutions needed to deliver well-crafted analyses. On the other side of the same coin, evidence scholars have a special need to become conversant with those other disciplines as well as with doctrinal analysis. This is perhaps especially true of the scientific method because both the social and natural sciences are increasingly relevant to the law's attempt to reach accurate verdicts. It is hard to know whether a rule helps triers reach accurate results without knowing something about human reasoning, or to know whether a forensic technique is valuable without knowing something about the scientific method.

We turn next to an examination of the principal strains of contemporary interdisciplinary scholarship on evidence.

II. PSYCHOLOGY AND EVIDENCE

For obvious reasons, psychology is the most important of the interdisciplinary threads that can be woven into evidence law scholarship. Evidence law is much concerned with the abilities of witnesses to perceive, to remember, and to report what they have observed. It is also concerned 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.²⁸ Experimental studies that address topics such as memory, perception, judgment, inference, decisions under conditions of uncertainty, and jury behavior are plainly relevant to evidence law.

A. *Early Examples*

The history of experimental psychology and law has been one of bursts of enthusiasm followed by periods of disenchantment. The first scholar to look at the interconnections of law and psychology was Hugo Münsterberg, a Harvard professor who, in the course of invent-

²⁸ These areas of psychology of primary interest to the scholarship of evidence law are the offspring of the marriage of philosophy and experimental biology that took place in the later 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* (2d ed. 1929).

ing one field of applied psychology after another—applications of psychology to industry, education, medicine, psychotherapy, business—undertook the first exploration of legal issues through a psychological lens. His 1908 book, *On the Witness Stand*,²⁹ dealt with problems of the perception and memory of witnesses, crime detection, confessions (especially false ones), influences on the examination of witnesses, hypnotism and crime, and the prevention of crime. Some of his points were sound, others flawed or quite speculative.³⁰ Unfortunately for the fledgling field, our evidence law forebear, John Henry Wigmore, took a strong dislike to Münsterberg's book, and published a somewhat bizarre, but unmistakably scathing, critique of it.³¹ Apparently Münsterberg never replied to Wigmore's attack,³² and we do not know whether Wigmore's assault caused him to abandon his work on law and psychology. Certainly it could have deterred others with an interest in the intersections where the two fields might profitably have met. One wonders what the body of psychology and evidence law scholarship might have developed into, and how much sooner, if the

²⁹ HUGO MÜNSTERBERG, *ON THE WITNESS STAND* (1908).

³⁰ Münsterberg's first two chapters, on memory, perception, and the fallibility of eye-witnesses, foreshadowed much later work and basically were sound, though written in a more casual fashion than contemporary reports of experimental work. For an assessment, see Ludy T. Benjamin, Jr., *Hugo Münsterberg, Portrait of an Applied Psychologist*, in 4 *PORTRAITS OF PIONEERS IN PSYCHOLOGY* 113, 113–30 (Gregory A. Kimble & Michael Wertheimer eds., 2000). Münsterberg then argued, however, that psychologists could detect lies about bad acts through a process of word association in which they present a suspect with words that are neutral and words that are related to the crime and measure the respective response times. MÜNSTERBERG, *supra* note 29, at 73–110. Wigmore effectively demolished this assertion. See John H. Wigmore, *Professor Münsterberg [sic] and the Psychology of Testimony*, 3 U. ILL. L. REV. 399, 427–31 (1909) (presenting a mock trial in which members of the bar brought a claim of libel against Professor Münsterberg for overstating the usefulness of psychologists as experts at trial). Other assertions by Münsterberg also were highly speculative, including that post-hypnotic suggestion could cause someone to bequeath all his money and then commit suicide, or that flashing lights increase suggestiveness and could cause false confessions.

³¹ See generally Wigmore, *supra* note 30.

³² Part of the reason might be that Münsterberg was busy with other projects. The same year that Wigmore's article appeared, Münsterberg published two more books, one on values and the other on the psychology of teaching. HUGO MÜNSTERBERG, *THE ETERNAL VALUES* (1909); HUGO MÜNSTERBERG, *PSYCHOLOGY AND THE TEACHER* (1909). The next year he was appointed an exchange professor from Harvard to the University of Berlin and sent on a quasi-diplomatic mission to establish an American Institute. BORING, *supra* note 28, at 427–28. 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. *Id.*; MATTHEW HALE, JR., *HUMAN SCIENCE AND SOCIAL ORDER: HUGO MÜNSTERBERG AND THE ORIGINS OF APPLIED PSYCHOLOGY* 103–05, 165–68 (1980); William Stern, *Hugo Münsterberg: In Memoriam*, 1 J. APPLIED PSYCHOL. 186, 186 (1917).

initial encounter had not been hobbled by the overreaching of its first important contributor and the overreaction of its first important critic.

Wigmore also tried to derail the second major event in the history of evidence law and psychology. When a bright young Yale law professor, Robert M. Hutchins, who in 1926 had just begun to teach evidence, delivered a paper on psychology and the law of evidence, Wigmore sent him a letter registering his disapproval and referring Hutchins to Wigmore's earlier critique of Münsterberg.³³ For good measure, Wigmore also sent a complaint to the president of Yale University.³⁴ Hutchins was not the least deterred. He took on psychologist Donald Slesinger as a collaborator and together they wrote prolifically on psychology and evidence law.³⁵ The articles drew from the extant body of psychological research to scrutinize evidence doctrine, identify weaknesses, and suggest improvements. Professor Schlegel's appraisal of them is that:

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.³⁶

Like other legal realist initiatives in law and social science, Hutchins and Slesinger's work just ran out of steam. The whole field of law and any social science was somewhat dormant for a time, suffering from lack of funding during the Great Depression and a redirection

³³ See John Henry Schlegel, *American Legal Realism and Empirical Social Science: From the Yale Experience*, 28 BUFF. L. REV. 459, 480 n.101 (1979).

³⁴ See *id.* at 474 n.83.

³⁵ Robert M. Hutchins & Donald Slesinger, *Some Observations on the Law of Evidence—The Competency of Witnesses*, 37 YALE L.J. 1017 (1928); Robert M. Hutchins & Donald Slesinger, *Some Observations on the Law of Evidence—Consciousness of Guilt*, 77 U. PA. L. REV. 725 (1929); Robert M. Hutchins & Donald Slesinger, *Some Observations on the Law of Evidence—Family Relations*, 13 MINN. L. REV. 675 (1929); Robert M. Hutchins & Donald Slesinger, *Some Observations on the Law of Evidence—Memory*, 41 HARV. L. REV. 860 (1928); Robert M. Hutchins & Donald Slesinger, *Some Observations on the Law of Evidence—Spontaneous Exclamations*, 28 COLUM. L. REV. 432 (1928); Robert M. Hutchins & Donald Slesinger, *Some Observations on the Law of Evidence—State of Mind in Issue*, 29 COLUM. L. REV. 147 (1929); Robert M. Hutchins & Donald Slesinger, *Some Observations on the Law of Evidence—State of Mind to Prove an Act*, 38 YALE L.J. 283 (1929).

³⁶ Schlegel, *supra* note 33, at 482.

of the energies of realist leaders during the Depression and the Second World War.³⁷ Hutchins left Yale in 1929 to become president of the University of Chicago, and four years later, at the ripe age of thirty-four, delivered an address in which he expressed doubts that the psychology of the era could teach the law enough to answer the necessary questions and resolve uncertainties about numerous evidence rules and their applications.³⁸

B. *Three Contemporary Stories of Success and Its Alternatives*

1. Research Relevant to Eyewitness Identification

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.³⁹ Today there are literally hundreds of studies on the subject of eyewitness testimony; one bibliography of eyewitness research lists 2000 entries, most of them scientific studies.⁴⁰ The body of research, both field studies and laboratory studies, has been growing at an increasing rate.⁴¹ On many questions, the findings show a high degree of convergence. Moreover, the research, unlike that of Münsterberg or that reviewed by Hutchins and Slesinger in their studies on psychology and evidence, 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.⁴² Do eyewitness experts help jurors understand? That has also been the subject of controlled experiments.⁴³ Do judges and lawyers understand problems with wit-

³⁷ *Id.* at 585–86.

³⁸ Robert Maynard Hutchins, *The Autobiography of an Ex-Law Student*, 1 U. CHI. L. REV. 511, 513 (1934).

³⁹ ELIZABETH F. LOFTUS, *EYEWITNESS TESTIMONY* (1979).

⁴⁰ BRIAN L. CUTLER & STEVEN D. PENROD, *MISTAKEN IDENTIFICATION: THE EYEWITNESS, PSYCHOLOGY, AND THE LAW* 68 (1995).

⁴¹ *Id.*

⁴² *Id.* at 173–80.

⁴³ *Id.* at 213–24.

nessing 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 field⁴⁴—exactly what Wigmore and Hutchins were calling for and bemoaning the absence of.

The topic had become so popular that one of us, while an editor of the journal *Law and Human Behavior*, wrote an editorial urging psycholegal scholars to tackle more than just eyewitness identification.⁴⁵ The dearth of scientific scholarship in the legal profession was noted even in the pages of *The New Yorker*, where the author, Atul Gawande, struck a tone similar to Münsterberg's:

[T]he 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"—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.⁴⁶

In contrast to a century ago, today's research on eyewitness testimony is far more complete, 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.⁴⁷

⁴⁴ *Id.* at 143–68, 255–64.

⁴⁵ Michael J. Saks, *The Law Does Not Live by Eyewitness Testimony Alone*, 10 LAW & HUM. BEHAV. 279, 279 (1986). The author states:

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.

Id.

⁴⁶ Atul Gawande, *Under Suspicion*, NEW YORKER, Jan. 8, 2001, at 50, 50.

⁴⁷ Thus, Gawande's criticism is unfair, at least as it pertains to eyewitness identification research.

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 identifiers than others.⁴⁸ 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.⁴⁹

What have been the legal consequences of this research? One important question that evidence casebooks address is whether expert testimony is admissible to alert jurors to the dangers of eyewitness identifications by pointing out the factors that increase or decrease eyewitness accuracy.⁵⁰ Most courts continue to hold that it is within the discretion of trial judges to exclude this evidence, but there are exceptions.⁵¹ Some trial judges, of course, exercise their discretion in favor of admitting it, especially when the eyewitness identification is crucial and when it fits the facts of the case.⁵²

More recently, eyewitness identification research has had an even more important effect on the administration of justice. It has suggested policy for the conduct of eyewitness identification procedures to minimize the risk of erroneous convictions without increasing the risk of erroneous failures to identify. In 1998, the American Psychology-Law Society produced a white paper on eyewitness identification

⁴⁸ For examples of these studies, see those cited in Peter Shapiro & Steven Penrod, *A Meta-Analysis of the Facial Identification Literature*, 100 PSYCHOL. BULL. 139, 154-56 (1986).

⁴⁹ See *id.* at 171-209.

⁵⁰ See CHRISTOPHER B. MUELLER & LAIRD C. KIRKPATRICK, *EVIDENCE UNDER THE RULES: TEXT, CASES, AND PROBLEMS* 599-600 (4th ed. 2000); JON R. WALTZ & ROGER C. PARK, *CASES AND MATERIALS ON EVIDENCE* 935-45 (10th ed. 2004).

⁵¹ 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, 1223-24 (Ariz. 1983) and *People v. McDonald*, 690 P.2d 709, 726-27 (Cal. 1984).

⁵² See *United States v. Hines*, 55 F. Supp. 2d 62, 72-73 (D. Mass. 1999).

written by a number of leading eyewitness researchers.⁵³ Their recommendations included:

- selecting lineup foils (or fillers) to resemble the witness's description (rather than the suspect's);
- blind administration of lineups (because police officers who interact 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 judgment);⁵⁴
- 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, which is also designed to reduce relative judgment); and
- 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).⁵⁵

The U.S. Department of Justice has adopted most of these recommendations (with the notable exception of blind testing)⁵⁶ and has advised all police agencies throughout the United States to follow them.⁵⁷ Other jurisdictions have gone further. The Illinois Governor's Commission on Capital Punishment recommended double-blind lineups and the state's General Assembly funded a limited program of

⁵³ Gary L. Wells, Mark Small, Steven Penrod, Roy S. Malpass, Solomon M. Fulero & C.A.E. Brimacombe, *Eyewitness Identification Procedures: Recommendations for Lineups and Photospreads*, 22 LAW & HUM. BEHAV. 603 (1998).

⁵⁴ "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.

⁵⁵ Wells et al., *supra* note 53, at 627-29, 632, 635, 639.

⁵⁶ Police representatives on the Department of Justice 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 argued 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 L. Wells et al., *From the Lab to the Police Station: A Successful Application of Eyewitness Research*, 55 AM. PSYCHOLOGIST 581, 594 (2000).

⁵⁷ U.S. DEP'T OF JUSTICE, EYEWITNESS EVIDENCE: A GUIDE FOR LAW ENFORCEMENT, at iii-iv, 9 (1999).

implementation and study of the procedure.⁵⁸ New Jersey has adopted double-blind lineups on a statewide basis.⁵⁹

2. Research Relevant to Character Evidence

The legal literature on character evidence has varied in its use of insights from academic literature on psychology.⁶⁰ Professor Uviller produced an influential article on character evidence without expressly relying upon insights from the psychology literature.⁶¹ Other articles do cite and discuss the literature on psychology, but are relatively cautious in their use of it.⁶² Still others have been more daring, using personality theory to argue that character evidence is worthless or virtually worthless.⁶³

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.⁶⁴ Some legal

⁵⁸ Thomas P. Sullivan, *Preventing Wrongful Convictions—A Current Report from Illinois*, 52 *DRAKE L. REV.* 605, 608–09 (2004).

⁵⁹ Gary L. Wells, *Police Lineups: Data, Theory, and Policy*, 7 *PSYCHOL. PUB. POL'Y & L.* 791, 791 (2001).

⁶⁰ For an extensive examination of the complexities of the character evidence doctrine, see generally EDWARD J. IMWINKELRIED, *UNCHARGED MISCONDUCT EVIDENCE* (2d ed. 1999).

⁶¹ Uviller, *supra* note 23.

⁶² See, e.g., Friedman, *supra* note 23, at 646–54; Miguel Angel Mendez, *California's New Law on Character Evidence: Evidence Code Section 352 and the Impact of Recent Psychological Studies*, 31 *UCLA L. REV.* 1003, 1041–60 (1984); Miguel A. Méndez, *The Law of Evidence and the Search for a Stable Personality*, 45 *EMORY L.J.* 221, 226–36 (1996); Roger C. Park, *Character at the Crossroads*, 49 *HASTINGS L.J.* 717 *passim* (1998).

⁶³ 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); David P. Leonard, *The Use of Character to Prove Conduct: Rationality and Catharsis in the Law of Evidence*, 58 *U. COLO. L. REV.* 1, 26–30 (1986–87); Robert G. Spector, *Rule 609: A Last Plea for Its Withdrawal*, 32 *OKLA. L. REV.* 334, 351 (1979). But see Susan Marlene Davies, *Evidence of Character to Prove Conduct: A Reassessment of Relevancy*, 27 *CRIM. L. BULL.* 504, 533–36 (1991) (using personality theory to argue that opinion and specific-acts evidence on character should be admissible during the prosecution's case in chief).

⁶⁴ See John M. Darley & C. Daniel Batson, *"From Jerusalem to Jericho": A Study of Situational and Dispositional Variables in Helping Behavior*, 27 *J. PERSONALITY & SOC. PSYCHOL.* 100, 107–08 (1973) (study of effect of being in a hurry on seminarians' behavior in helping strangers).

scholars have used "situationist" personality theory to argue that character evidence ought to be broadly excluded.⁶⁵ 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.⁶⁶ 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.⁶⁷

Along with the benefits of informing legal thought with interdisciplinary materials such as those discussed above come certain risks. One is the risk of applying literature that is too scant, too selective, and too old.⁶⁸ Another is the danger of using research that does not generalize to the legal situation. The first risk 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, such as avoiding the temptation to take one's psychology from law review articles and the empirical studies cited in them, and checking one's conclusions against recent editions of standard college texts on the subject.⁶⁹

⁶⁵ See Foster, *supra* note 63, at 32 ("Social psychology data reflect the conclusion that prior convictions have virtually no probative value as a predictor for determining a witness' in-court veracity."); Leonard, *supra* note 63, at 26 ("[A]s currently conceptualized by the law, character evidence often fails the test of logical relevance."); Spector, *supra* note 63, at 351 ("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."); cf. Robert G. Lawson, *Credibility and Character: A Different Look at an Interminable Problem*, 50 NOTRE DAME L. REV. 758, 783 (1974-75) (describing a less radical theory that considers both a person's trait and the situation).

⁶⁶ For a description of these experiments, see LEE ROSS & RICHARD E. NISBETT, *THE PERSON AND THE SITUATION: PERSPECTIVES OF SOCIAL PSYCHOLOGY* 119-44 (1991). Here, incidentally, guided by the work of Ross & Nisbett, is where the debate about psychology and the character rule might have been directed most usefully: not the question of whether stable traits exist but whether factfinders over-attribute behavior to traits (and under-attribute the influence of situations) or too easily infer a trait from learning about a small sample of conduct.

⁶⁷ *Id.*

⁶⁸ Sometimes old research is every bit as good as, 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.

⁶⁹ 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 "[b]oth 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." *Id.* 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.

Another possibility is to follow Hutchins's example and seek a psychologist as a co-author.

The second challenge is to make sure that the interdisciplinary literature "fits"—that is, that it generalizes to the relevant legal context. This 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 articles⁷⁰ were Hartshorne and May's massive 1928 work⁷¹ and Walter Mischel's 1968 book, *Personality and Assessment*.⁷² 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 think about its generalizability 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.⁷³ 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 correlations between different kinds of dishonest or deceptive behaviors were very modest.⁷⁴

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 groundbreaking book focuses on 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,

⁷⁰ See articles cited *supra* notes 62, 63, 65.

⁷¹ HUGH HARTSHORNE & MARK A. MAY, *STUDIES IN DECEIT* (1928).

⁷² WALTER MISCHEL, *PERSONALITY AND ASSESSMENT* (1968).

⁷³ See HARTSHORNE & MAY, *supra* note 71, at 90–103.

⁷⁴ See *id.* at 382–84. 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.

such as progress of the patient.⁷⁵ 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.⁷⁶ His views about the lack of utility of trait-state theories are tied to his preference for "social behavior" therapy—for instance, 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.⁷⁷ 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.⁷⁸ However, experimental testing indicates that there was no substantial correlation between attitudes toward fathers and attitudes toward bosses.⁷⁹ He obtained similar results in examining attitudes towards peers and comparing them to attitudes toward siblings.⁸⁰ Thus, Mischel concluded that "there was little evidence for generality of attitudes either toward authority or toward peers."⁸¹

Mischel also expressed doubt about the generality of the psychological construct of an "authoritarian personality."⁸² 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.⁸³ He noted similar results with the posited personality trait of "rigidity."⁸⁴ Similarly, pencil-and-paper anxiety scales correlated substantially with each other, but not with psychological measures of anxiety.⁸⁵ He noted generally that when behaviors are measured by tests with similar formats and content there are substantial correlations but the cor-

⁷⁵ MISCHEL, *supra* note 72, at 106–48.

⁷⁶ *Id.* at 118–23.

⁷⁷ *Id.* at 149–50.

⁷⁸ *Id.* at 21.

⁷⁹ *Id.* at 22.

⁸⁰ MISCHEL, *supra* note 72, at 22.

⁸¹ *Id.*

⁸² *Id.* at 28.

⁸³ *Id.*

⁸⁴ *Id.* at 29.

⁸⁵ MISCHEL, *supra* note 72, at 81.

relation often does not hold when the same behaviors are measured by different means.⁸⁶ 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.⁸⁷ He noted that experiments in which children's 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 who modeled patience.⁸⁸

The question for the law of evidence is whether the findings of the studies reviewed and reported by Hartshorne and 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 make one think that what is learned from those studies tells us about the behavior of adults?⁸⁹ Similarly, valid research about the predictive value of ordinary behavior does not necessarily generalize to the prediction of extreme behavior.⁹⁰

When the legal issue is whether evidence showing propensity for criminal violence should be admitted, there is an additional problem of "fit": none of the underlying research in the studies referred to dealt with criminal or violent behavior. Although decades of research have carried the Hartshorne and 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

⁸⁶ *Id.* at 101.

⁸⁷ *Id.* at 23-25.

⁸⁸ *Id.* at 151-52.

⁸⁹ This has been itself a research problem in 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 & ROBERT PLOMIN, *A TEMPERAMENT THEORY OF PERSONALITY DEVELOPMENT* 7-8 (1975); Jerome Kagan, *Temperamental Contributions to Social Behavior*, 44 *AM. PSYCHOLOGIST* 668, 668 (1989).

⁹⁰ See ROSS & NISBETT, *supra* note 66, at 116-17; Park, *supra* note 62, at 737 & n.67.

lack of cross-situational consistency of behavior have been based on studies that generally examine nonviolent, noncriminal behavior of normal research participants.⁹¹ This is not a criticism of that body of research, but simply an 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.⁹² At the end of the day, external validity (generalizability) can only be determined empirically, not logically.

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."⁹³ 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 behavior."⁹⁴ By focusing on the behavior, as a matter of sheer prediction, it does not matter

⁹¹ For a useful review of these studies, see generally ROSS & NISBETT, *supra* note 66.

⁹² 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 Control Theory Interpretation of Psychological Research on Aggression*, in AGGRESSION AND VIOLENCE: SOCIAL INTERACTIONIST PERSPECTIVES 47, 50 (Richard B. Felson & James T. Tedeschi eds., 1993) ("[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" (quoting L. Rowell Huesmann et al., *Stability of Aggression over Time*, 20 DEVELOPMENTAL PSYCHOL. 1120, 1133 (1984))); Dan Olweus, *Stability of Aggressive Reaction Patterns in Males: A Review*, 86 PSYCHOL. BULL. 852, 872-73 (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 ROSS & NISBETT, *supra* note 66, at 52-58; Gottfredson & Hirschi, *supra*, at 53 (Milgram's studies "measure compliance, obedience, or . . . acquiescence," rather than aggression).

⁹³ MISCHEL, *supra* note 72, at 135-48.

⁹⁴ *Id.*

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 Mischel 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.⁹⁵

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 eighteen and twenty-one years of age.⁹⁶ 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,⁹⁷ and among non-prison populations, boys and men with higher testosterone levels were more likely to respond aggressively to provocation.⁹⁸ When aggression has been rewarded, especially intermittently, it tends to persist.⁹⁹ Some men have long records of criminal violence, which has been found to serve one of two major purposes: to command respect among associates and for more instrumental purposes.¹⁰⁰ A review of a large body of research on people who participate in violent and other criminal activity found

⁹⁵ 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 *United States v. Lindstrom*, 698 F.2d 1154, 1162 n.6 (11th Cir. 1983).

⁹⁶ See generally JAMES Q. WILSON & RICHARD HERRNSTEIN, *CRIME AND HUMAN NATURE* (1985).

⁹⁷ James M. Dabbs Jr., Timothy S. Carr, Robert L. Frady & Jasmin K. Riad, *Testosterone, Crime, and Misbehavior Among 692 Male Prison Inmates*, 18 *PERSONALITY & INDIVIDUAL DIFFERENCES* 627, 631 (1995); James M. Dabbs Jr., *Testosterone Measurements in Social and Clinical Psychology*, 11 *J. SOC. & CLINICAL PSYCHOL.* 302, 309 (1992).

⁹⁸ John Archer, *The Influence of Testosterone on Human Aggression*, 82 *BRIT. J. PSYCHOL.* 1, 13-14 (1991); James M. Dabbs, Jr. & Robin Morris, *Testosterone, Social Class, and Antisocial Behavior in a Sample of 4,462 Men*, 1 *PSYCHOL. SCI.* 209, 209 (1990).

⁹⁹ See generally Gerald R. Patterson, Richard A. Littman & William Bricker, *Assertive Behavior in Children: A Step Toward a Theory of Aggression*, 32 *MONOGRAPHS SOC'Y FOR RES. CHILD DEV.* 39 (1967).

¹⁰⁰ HANS TOCH, *Violent Men* 133-34 (1992).

that such conduct clusters among genetically closely-related individuals.¹⁰¹

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.¹⁰² That persistent finding has been helpful in leading researchers away from reliance on clinical assessment to study specific risk factors¹⁰³ 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."¹⁰⁴ 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.¹⁰⁵ Prediction tools using the risk factor approach have greatly improved predictive power, though it remains less than a pre-

¹⁰¹ WILSON & HERRNSTEIN, *supra* note 96, at 90–100. 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? The Answer May Depend on Who Chooses What Evidence*, SCI. AM., Feb. 1986, at 22, 22 (criticizing Wilson and Herrnstein's lack of empirical basis for their arguments and their failure to distinguish between causation and correlation); Heathcote W. Wales, *Tilting at Crime: The Perils of Eclecticism*, 74 GEO. L.J. 481, 489–91 (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 *60 Minutes: Murder Gene* (CBS television broadcast Feb. 10, 1999) (profiling Jeff Landrigan, an Arizona death row prisoner with an extensive criminal record matching that of his biological father, Darrel Hill, who is also on death row in Arkansas, even though the son was adopted out of his biological family in infancy). In the meantime, the U.S. Supreme Court will soon decide whether Jeff Landrigan made a colorable claim that he was prejudiced by ineffective assistance of counsel because his lawyer failed to submit during sentencing evidence of his (partly genetic) predisposition to violence. See generally *Landrigan v. Schriro*, 441 F.3d 638 (9th Cir. 2006) (en banc), cert. granted, 75 U.S.L.W. 3162 (U.S. Sept. 26, 2006) (No. 05-1575). 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.

¹⁰² 2 MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY §§ 12:1, 12:20 (David L. Faigman, David H. Kaye, Michael J. Saks & Joseph Sanders eds., 2005) [hereinafter MODERN SCIENTIFIC EVIDENCE] ("[T]he sober conclusion that clinicians are 'modestly better than chance' at predicting violence appears to be becoming the consensus view.").

¹⁰³ See generally JOHN MONAHAN ET AL., *RETHINKING RISK ASSESSMENT: THE MACARTHUR STUDY OF MENTAL DISORDER AND VIOLENCE* (2001).

¹⁰⁴ 2 MODERN SCIENTIFIC EVIDENCE, *supra* note 102, § 12:12.

¹⁰⁵ *Id.* § 12:19.

cise art or science.¹⁰⁶ The occurrence of past violent acts is predictive; the more of the acts and the more serious they are, the higher the predictive accuracy.¹⁰⁷

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) to the more inclusive proposals (making character evidence in some form, as well as other-acts evidence, more available). To reach such conclusions, the law must learn more about the actual utility of "character" and prior acts in predicting behavior (or 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 is 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.¹⁰⁸ 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

¹⁰⁶ *Id.* § 12:13.

¹⁰⁷ See MONAHAN ET AL., *supra* note 103, at 44-47.

¹⁰⁸ Though we say this with some trepidation, a Bayesian format for presenting the information might be most appropriate. See *infra* note 183 and accompanying text.

with care and determine whether they are applicable or distinguishable. They should do no less when using psychology authority.

3. Research Relevant to the Hearsay Rule

The first empirical study of the juror use of hearsay was published in 1991.¹⁰⁹ Since then there have been at least twenty-six other studies, most of them original empirical work as opposed to reviews or commentary on empirical work by others.¹¹⁰ 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 two 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.¹¹¹ 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, these studies 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.¹¹² For example, if the out-of-court declarant's statements

¹⁰⁹ Stephan Landsman & Richard F. Rakos, *Research Essay: A Preliminary Empirical Enquiry Concerning the Prohibition of Hearsay Evidence in American Courts*, 15 LAW & PSYCHOL. REV. 65 (1991).

¹¹⁰ For a bibliography, see Roger C. Park, *Visions of Applying the Scientific Method to the Hearsay Rule*, 2003 MICH. ST. L. REV. 1149, 1171-74.

¹¹¹ *Id.* In addition to arguing appellate cases involving high-tech forensic science, Thompson was co-counsel for O.J. Simpson in *People v. Simpson*.

¹¹² Maithilee K. Pathak & William C. Thompson, *From Child to Witness to Jury: Effects of Suggestion on the Transmission and Evaluation of Hearsay*, 5 PSYCHOL. PUB. POL'Y & L. 372, 381-86 (1999).

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?¹¹³ Some studies created mock cases, the "truth" of which is not only not known but does not exist;¹¹⁴ others created a set of true underlying facts, and the jurors' conclusions about the event could be compared to its reality.¹¹⁵ Some of the studies presented the "case" by way of brief written transcripts or even briefer summaries,¹¹⁶ others by audio or videotaped presentations of mock trials.¹¹⁷ In some stud-

¹¹³ *Id.*

¹¹⁴ See, e.g., Jonathan M. Golding, Rebecca Polley Sanchez & Sandra A. Sego, *The Believability of Hearsay Testimony in a Child Sexual Assault Trial*, 21 LAW & HUM. BEHAV. 299, 304 (1997); Landsman & Rakos, *supra* note 109, at 73.

¹¹⁵ Peter Miene, Roger C. Park & Eugene Borgida, *Juror Decision Making and the Evaluation of Hearsay Evidence*, 76 MINN. L. REV. 683, 688-89 (1992). 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. See Margaret Bull Kovera, Roger C. Park & Steven D. Penrod, *Jurors' Perceptions of Eyewitness and Hearsay Evidence*, 76 MINN. L. REV. 703, 708 (1992). The results of such studies can provide insights into the ability of mock jurors to reach accurate conclusions about the underlying events that are otherwise completely unavailable.

¹¹⁶ See, e.g., Golding et al., *supra* note 114, at 304; Landsman & Rakos, *supra* note 109, at 73; Richard F. Rakos & Stephan Landsman, *Researching the Hearsay Rule: Emerging Findings, General Issues, and Future Directions*, 76 MINN. L. REV. 655, 659 (1992).

¹¹⁷ See, e.g., Miene et al., *supra* note 115, at 689; Angela Paglia & Regina A. Schuller, *Jurors' Use of Hearsay Evidence: The Effects of Type and Timing of Instructions*, 22 LAW & HUM. BEHAV. 501, 506 (1998). 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:

[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.

ies the hearsay witness was a layperson,¹¹⁸ in others an expert witness.¹¹⁹ In some the firsthand witness was cross-examined¹²⁰ and in others not.¹²¹ And so on.

Put succinctly, some of these studies appear to offer greater verisimilitude than others.¹²² 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.¹²³ In other studies, the jurors credited the hearsay as much as they did the firsthand testimony.¹²⁴ 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.

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 in-

¹¹⁸ See Golding et al., *supra* note 114, at 304.

¹¹⁹ Regina A. Schuller & Angela Paglia, *An Empirical Study: Juror Sensitivity to Variations in Hearsay Conveyed via Expert Evidence*, 23 LAW & PSYCHOL. REV. 131, 137-38 (1999) (finding no differences in verdicts as a function of the amount of hearsay underlying an expert opinion, though other measures suggested that the participants were able to distinguish between hearsay that was supported by other evidence presented at trial and hearsay that was not supported in that fashion).

¹²⁰ See Rakos & Landsman, *supra* note 116, at 658-60.

¹²¹ See Kovera et al., *supra* note 115, at 707-10.

¹²² It is far from clear whether realistic simulations are necessary to get correct answers that are generalizable to actual trial settings. A comparison of simulations with more and with less verisimilitude found that the results did not differ. Brian H. Bornstein, *The Ecological Validity of Jury Simulations: Is the Jury Still Out?*, 23 LAW & HUM. BEHAV. 75, 88 (1999). But it almost certainly is the case that lawyers and judges are more persuaded by studies that share many superficial similarities with trial procedures.

¹²³ See Kovera et al., *supra* note 115, at 719; Miene et al., *supra* note 115, at 691.

¹²⁴ See Golding et al., *supra* note 114, at 318; Jonathan M. Golding, Mary C. Alexander & Terri L. Stewart, *The Effect of Hearsay Witness Age in a Child Sexual Assault Trial*, 5 PSYCHOL. PUB. POL'Y & L. 420, 433 (1999).

stead. The policymaker 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.¹²⁵ 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 television 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.¹²⁶ Thus it is obvious that they are of limited usefulness in assessing the value of hearsay testimony as a substitute for cross-examined testimony.

b. Whether Hearsay Should Be Received when There Is No Choice

A different issue of hearsay policy is presented when the declarant is unavailable, 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

¹²⁵ The substitution has other consequences, such as precluding physical confrontation between the witness and the accused in criminal cases and allowing evidence to be received from out-of-court declarants who were not under oath or subject to observation for demeanor cues.

¹²⁶ See Kovera et al., *supra* note 115, at 707-10; Pathak & Thompson, *supra* note 112, at 375-77, 379-80.

go in the dark because the light is not perfect. But the hearsay rule sometimes bars such testimony.¹²⁷ 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.¹²⁸ In two of the experiments, however, the investigators did have accurate knowledge of the ground truth.¹²⁹ 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 reconstruct accurately what actually happened.

Neither of those experiments simulated cross-examination.¹³⁰ 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

¹²⁷ See *United States v. Day*, 591 F.2d 861, 881–83 (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, 850 (Ct. App.), petition for review granted, 102 P.3d 930 (2004); *State v. Meeks*, 88 P.3d 789, 794 (Kan. 2004) (holding that defendant forfeited the right to confrontation by murdering victim (citing Richard D. Friedman, *Confrontation and the Definition of Chutzpa*, 31 *Isr. L. Rev.* 506 (1997))).

¹²⁸ See *supra* note 114 and accompanying text.

¹²⁹ Kovera et al., *supra* note 115, at 707–09; Pathak & Thompson, *supra* note 112, at 375–81.

¹³⁰ Kovera et al., *supra* note 115, at 707–09; Pathak & Thompson, *supra* note 112, at 375–81.

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 or whether the declarant was paying attention).

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, appellate courts refuse to reverse.¹³¹ The hearsay studies are relevant to this policy issue even when they do not address the question whether hearsay aids in reaching accurate verdicts. If juries heavily discount hearsay, as some of the studies suggest,¹³² then the erroneous admission of hearsay (at least where it is substituted for live testimony) should be considered prejudicial error less often.

The hearsay experiments reach divergent results, and without further research we cannot be sure why. 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 (for instance, 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. Indeed, the divergent results might have been due to something else. In any event, the divergent results, the wide range of circumstances in which hearsay issues can arise, and the differences between the simulations and the actual courtroom environment, caution against 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

¹³¹ See ROGER C. PARK, TRIAL OBJECTIONS HANDBOOK § 12.2 (2d ed. 2001).

¹³² See Kovera et al., *supra* note 115, at 719; Miene et al., *supra* note 115, at 691.

more narrow questions. For example, the experiment by Warren and Woodhall¹³³ suggests the danger in admitting hearsay in the child witness context—persons interrogating child witnesses do not remember accurately their mode of interrogation or whether they were suggestive. The Pathak and Thomson experiment¹³⁴ 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. Hearsay issues arise in many different contexts. 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-examination. Hearsay might be an adequate substitute in situations in which the purpose of cross-examination 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.¹³⁵

Consequently, it makes sense to approach the vast problem of hearsay and its exceptions bit by bit, studying specific situations such

¹³³ Amye R. Warren & Cara E. Woodall, *The Reliability of Hearsay Testimony: How Well Do Interviewers Recall Their Interviews with Children?*, 5 PSYCHOL. PUB. POL'Y & L. 355, 365–70 (1999).

¹³⁴ Pathak & Thompson, *supra* note 112, at 386.

¹³⁵ Suppose, for example, that a 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 one does not lose anything when one uses hearsay instead of a neutral witness would not mean that one does not lose anything when one uses hearsay in lieu of cross-examining Detective Fuhrman in the O.J. Simpson trial.

as child witness interrogations 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, policymakers 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.

In the preceding examples, psychological research and theory were employed to test the assumptions underlying evidence doctrine. That is, indeed, what most legal scholars have used psychology for. But another role, one that awaits future development, is the possibility that psychology could help to identify and explain existing regularities in evidence law: why, in the hands of common-law evidence rule-makers, certain rules evolved and others did not.¹³⁶

III. LAW AND FORENSIC SCIENCE

Law and forensic science has long been a topic of evidence scholarship—the history of articles about new ways of gathering evidence to prove the identity of suspected perpetrators is a long one.¹³⁷ But the flourishing of this field has been remarkable since the Supreme Court's 1993 decision in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*¹³⁸ *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,¹³⁹ what the results of testing have been,¹⁴⁰ and how well judges have been doing with the law and the science.¹⁴¹

¹³⁶ An article that reflects this use of psychological knowledge to understand law, including some aspects of the law of evidence, is Jeffrey J. Rachlinski, *A Positive Psychological Theory of Judging in Hindsight*, 65 U. CHI. L. REV. 571 (1998).

¹³⁷ See generally SIMON A. COLE, SUSPECT IDENTITIES: A HISTORY OF FINGERPRINTING AND CRIMINAL IDENTIFICATION (2001).

¹³⁸ 509 U.S. 579 (1993).

¹³⁹ See generally COMM. ON SCI. ASSESSMENT OF BULLET LEAD ELEMENTAL COMPOSITION COMPARISON, NAT'L RESEARCH COUNCIL, FORENSIC ANALYSIS: WEIGHING BULLET LEAD EVIDENCE (2004) [hereinafter WEIGHING BULLET LEAD EVIDENCE] (citing works); 4 MODERN SCIENTIFIC EVIDENCE, *supra* note 102 (various chapters on various forensic sciences); Paul Giannelli & Edward Imwinkelried, *Scientific Evidence: The Fallout from Supreme Court's Decision in Kumho Tire*, 14 CRIM. JUST. 12 (2000); Jane Campbell Moriarty & Michael J. Saks, *Forensic Science: Grand Goals, Tragic Flaws, and Judicial Gatekeeping*, 44 JUDGES' JOURNAL 16 (Fall 2005); Michael J. Saks & Jonathan J. Koehler, *The Coming Paradigm Shift in Forensic Identification Science*, 309 SCIENCE 892 (2005).

Daubert already has led to an astonishing amount of scholarship¹⁴² 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*.¹⁴³ These discussions focus not only on issues of the legal admissibility of scientific evidence, or other limitations that might be placed on it, but also on the scientific basis for the evidence as required by *Daubert* and *Kumho Tire Co. v. Carmichael*.¹⁴⁴ In the post-*Daubert* world, judges, lawyers, and legal scholars cannot do their work properly without becoming scientifically literate.¹⁴⁵

What lawyers, scholars, and the courts are discovering 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

Ironically, at the same time that judges, legal scholars, and scientists are inquiring into the scientific foundations of forensic science, often finding them to be surprisingly weak, popular culture is experiencing an upsurge of credulousness. See Janine Robben, *The 'CSI' Effect: Popular Culture and the Justice System*, 66 OR. ST. B. BULL. 9, 9 (Oct. 2005).

¹⁴⁰ 4 MODERN SCIENTIFIC EVIDENCE, *supra* note 102, §§ 91:33–45.

¹⁴¹ See, e.g., Sophia I. Gatowski, Shirley A. Dobbin, James T. Richardson, Gerald P. Ginsburg, Mara L. Merlino & Veronica Dahir, *Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-Daubert World*, 25 LAW & HUM. BEHAV. 433, 451–55 (2001); Jennifer L. Groscup, Steven D. Penrod, Christina A. Studebaker, Matthew T. Huss & Kevin M. O'Neil, *The Effects of Daubert on the Admissibility of Expert Testimony in State and Federal Criminal Cases*, 8 PSYCHOL. PUB. POL'Y & L. 339, 363–71 (2002); Joseph Sanders, Shari S. Diamond & Neil Vidmar, *Legal Perceptions of Science and Expert Knowledge*, 8 PSYCHOL. PUB. POL'Y & L. 139, 152 (2002). See generally DAVID H. KAYE, DAVID E. BERNSTEIN, JENNIFER L. MNOOKIN & RICHARD D. FRIEDMAN, *THE NEW WIGMORE, A TREATISE ON EVIDENCE: EXPERT EVIDENCE* (2005); 4 MODERN SCIENTIFIC EVIDENCE, *supra* note 102.

¹⁴² A search of the Westlaw JLR database (journals and law reviews) on February 25, 2006, found 1492 articles citing *Kumho Tire Co. v. Carmichael* and 5173 articles citing *Daubert*.

¹⁴³ See generally FED. JUDICIAL CTR., REFERENCE MANUAL ON SCIENTIFIC EVIDENCE (2d ed. 2000); MODERN SCIENTIFIC EVIDENCE, *supra* note 102; MICHAEL J. SAKS, DAVID L. FAIGMAN, DAVID H. KAYE & JOSEPH SANDERS, ANNOTATED REFERENCE MANUAL ON SCIENTIFIC EVIDENCE, SECOND (2004).

¹⁴⁴ *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 147–53 (1999) (applying *Daubert* to engineers and other experts who are not scientists).

¹⁴⁵ Efforts to assist them have taken the form of conferences, courses, CLE programs, checklists, and works of the kind cited *supra* note 143. A good example of the need of judges and lawyers to learn to think conceptually about scientific research is provided by Edward J. Imwinkelried, *Coming to Grips with Scientific Research in Daubert's "Brave New World": The Courts' Need to Appreciate the Evidentiary Differences Between Validity and Proficiency Studies*, 61 BROOK. L. REV. 1247, 1282–84 (1995).

of the new test.¹⁴⁶ On the other side of the coin, and perhaps ironically, the new legal test might breathe new life into proffered polygraph examinations. Once courts must ask seriously about scientific foundations of fields, they discover that more science and more research exist concerning polygraph examination than about most or all of the traditional forensic sciences.¹⁴⁷ What courts should do with the new realizations, what are the exact requirements of *Daubert*, and what are or should be the contours of *Daubert*'s exclusionary zone are the subjects of countless articles.

Another important development in the area has been the rise and success of DNA evidence. At first DNA evidence was accepted blindly,¹⁴⁸ like many other forensic sciences.¹⁴⁹ Then came the *People v. Castro* case—real scientists got involved and criticized the methodology of the DNA labs, successfully.¹⁵⁰ Subsequently, the issues of methodology and fit were addressed and cleaned up. It now appears that DNA analysis may, by example, fuel improvements in the rest of forensic science, because its more careful and explicit scientific approach puts to shame so much of what previously had been accepted as forensic expertise.¹⁵¹ The advent of DNA typing has made the shaky scientific foundations of the traditional forensic individualization sciences all the more apparent. In addition, DNA-based exonerations of persons who had been erroneously convicted are exposing the flaws of the traditional forensic sciences,¹⁵² just as they have already reinforced, in the eyes of policymakers, the findings of scientific studies about the fallibility of eyewitness identification.¹⁵³

There already have been successful challenges to a form of expertise that, for most of the twentieth century, had been regarded as

¹⁴⁶ It turns out that these traditional identification techniques often are little more than observation plus intuition (look and opine), with little and sometimes no scientific underpinnings or testing of any of these fields' foundational questions, which they merely beg. See 4 MODERN SCIENTIFIC EVIDENCE, *supra* note 102, §§ 31:33-45; Michael J. Saks, *Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science*, 49 HASTINGS L.J. 1069, 1094-127 (1998); Saks & Koehler, *supra* note 139, at 892.

¹⁴⁷ See 4 MODERN SCIENTIFIC EVIDENCE, *supra* note 102, § 40.

¹⁴⁸ *Id.* § 32:3 ("rapid and sometimes uncritical acceptance").

¹⁴⁹ See generally Saks, *supra* note 146.

¹⁵⁰ 545 N.Y.S.2d 985, 999 (Sup. Ct. 1989); see Jennifer L. Mnookin, *Fingerprint Evidence in an Age of DNA Profiling*, 67 BROOK. L. REV. 13, 52-53 (2001).

¹⁵¹ Michael J. Saks & Jonathan J. Koehler, *What DNA "Fingerprinting" Can Teach the Law About the Rest of Forensic Science*, 13 CARDOZO L. REV. 361, 361 (1991).

¹⁵² See Saks & Koehler, *supra* note 139, at 893.

¹⁵³ EDWARD CONNORS, THOMAS LUNDREGAN, NEAL MILLER & TOM McEWEN, CONVICTED BY JURIES, EXONERATED BY SCIENCE: CASE STUDIES IN THE USE OF DNA EVIDENCE TO ESTABLISH INNOCENCE AFTER TRIAL 24 (1996); U.S. DEP'T OF JUSTICE, *supra* note 57, at 3.

being of settled admissibility—handwriting identification.¹⁵⁴ Similar weaknesses can be found in microscopic hair identification, bite-marks, toolmarks, and other areas of forensic 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 and assumptions relied upon 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 that 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 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

Perhaps the most dramatic example of the impact of the law's new approach to scrutinizing purported scientific evidence has been bullet lead analysis. Bullet lead analysis, in use since the 1960s, involves the comparison of the chemical composition of a crime scene bullet to the chemical composition of bullets found in the possession of a suspect.¹⁶¹ Debate over the technique throughout the 1990s questioned the theory behind it, and pointed out that experts often testified beyond what could be supported by available data.¹⁶² The resulting debate ultimately led the Federal Bureau of Investigation (the

¹⁵⁴ See D. Michael Risinger, *Handwriting Identification*, in 4 MODERN SCIENTIFIC EVIDENCE, *supra* note 102, § 35.

¹⁵⁵ See generally 4 MODERN SCIENTIFIC EVIDENCE, *supra* note 102.

¹⁵⁶ See *id.* § 34; Robert Epstein, *Fingerprints Meet Daubert: The Myth of Fingerprint "Science" Is Revealed*, 75 S. CAL. L. REV. 605, 622–49 (2002). See generally COLE, *supra* note 137.

¹⁵⁷ See 4 MODERN SCIENTIFIC EVIDENCE, *supra* note 102, §§ 34:3–:13.

¹⁵⁸ *United States v. Llera Plaza*, No. CR. 98-362-10, 2002 WL 27305, at *517–18 (E.D. Pa. Jan. 7, 2002); see also *United States v. Crisp*, 324 F.3d 261, 272 (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, 560, 575–76 (E.D. Pa. 2002) (finding, nevertheless, that the "science" of fingerprint identification suffered from major shortcomings).

¹⁶⁰ See 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. *Id.* at 105.

¹⁶¹ WEIGHING BULLET LEAD EVIDENCE, *supra* note 139, at 8.

¹⁶² See William A. Tobin & Wayne Duerfeldt, *How Probative Is Comparative Bullet Lead Analysis?*, 17 CRIM. JUST. 26, 33–34 (2002).

only practitioner of the technique) to discontinue it in 2005, based on a review of the claims of this "science" produced by the National Research Council.¹⁶³

In addition, there are important questions directed at what might be regarded as the intersection of forensic science, law, and psychology. What inferences do jurors draw from admitted expert testimony on forensic science? Do they understand the evidence properly? Are they misled by the testimony? Can the testimony be cabined in ways that would correct any problems that are found?¹⁶⁴

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*, inevitably will lead to more evidence scholarship about law and science of all kinds.¹⁶⁵

IV. THE "NEW EVIDENCE SCHOLARSHIP"—PROBABILITY AND PROOF

The term "New Evidence Scholarship," coined by Richard Lempert,¹⁶⁶ is broad enough to 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

¹⁶³ See generally WEIGHING BULLET LEAD EVIDENCE, *supra* note 139.

¹⁶⁴ Jonathan J. Koehler, *DNA Matches and Statistics: Important Questions, Surprising Answers*, 76 JUDICATURE 222, 229 (1993); Jonathan J. Koehler & Daniel N. Shavero, *Veridical Verdicts: Increasing Verdict Accuracy Through the Use of Overtly Probabilistic Evidence and Methods*, 75 CORNELL L. REV. 247, 275-79 (1990); Jonathan J. Koehler, *When Are People Persuaded by DNA Match Statistics?*, 15 LAW & HUM. BEHAV. 493, 508-10 (2001); Dale A. Nance & Scott B. Morris, *Juror Understanding of DNA Evidence: An Empirical Assessment of Presentation Formats for Trace Evidence with a Relatively Small Random-Match Probability*, 34 J. LEGAL STUD. 395, 418-36 (2005); Gary L. Wells, *Naked Statistical Evidence of Liability: Is Subjective Probability Enough?*, 62 J. PERSONALITY AND SOC. PSYCHOL. 739, 748-50 (1992).

¹⁶⁵ In addition to the paradigm-shifting developments in traditional criminal identification, there will be new legal solutions as well as new problems brought by advances in biotechnology, nanotechnology, and digital informatics, among other fields.

¹⁶⁶ Richard Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof*, 66 B.U. L. REV. 439, 439-40 (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* note 168] is being rediscovered, and disciplines outside the law, like mathematics, psychology and philosophy, are being plumbed for the guidance they can give.").

scholarship that applies formal tools of probability theory, such as Bayes' Theorem.¹⁶⁷

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.¹⁶⁸ 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 scholar used the product rule in an attempt to calculate whether commonalities in two signatures could have been coincidental.¹⁶⁹

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 1968 article on decision theory,¹⁷⁰ Richard Eggleston's 1978 book entitled *Evidence, Proof and Probability*,¹⁷¹ and Jonathan Cohen's 1977 challenge to the applicability of standard probability reasoning in legal contexts.¹⁷² The 1968 case of *People v. Collins*,¹⁷³ 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*,¹⁷⁴ with Michael Finkelstein and William Fairley on one side and Laurence Tribe on the other, no doubt inspired other scholars to pursue the topic,

¹⁶⁷ See *infra* note 183 and accompanying text.

¹⁶⁸ See generally JOHN HENRY WIGMORE, *THE SCIENCE OF JUDICIAL PROOF AS GIVEN BY LOGIC, PSYCHOLOGY, AND GENERAL EXPERIENCE AND ILLUSTRATED IN JUDICIAL TRIALS* (3d ed. 1937).

¹⁶⁹ See Note, *The Howland Will Case*, 4 AM. L. REV. 625, 648-49 (1870) (recounting testimony of Benjamin Peirce and his son, mathematician Charles S. Peirce). Other early work included the rigorous and largely ignored efforts of Jerome Michael and Mortimer Adler to trace the logic of proof. See Jerome Michael & Mortimer J. Adler, *Real Proof: I*, 5 VAND. L. REV. 344 (1952); 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). William Twining detected a "continuous intellectual tradition from Bentham, through Wills, Best, Stephen, Thayer, Gulson, Wigmore, Michael and Adler to Leo Levin, Irvin Rutter, and contemporary teachers of law." See TWINING, *RETHINKING EVIDENCE: EXPLORATORY ESSAYS*, *supra* note 1, at 12, 15. 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. *Id.* at 341-42.

¹⁷⁰ John Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065 (1968).

¹⁷¹ RICHARD EGGLESTON, *EVIDENCE, PROOF AND PROBABILITY* (1978).

¹⁷² L. JONATHAN COHEN, *THE PROBABLE AND THE PROVABLE* 6 (1977).

¹⁷³ 438 P.2d 33, 36-37 (Cal. 1968).

¹⁷⁴ Michael O. Finkelstein & William B. Fairley, *A Bayesian Approach to Identification Evidence*, 83 HARV. L. REV. 489 (1970); Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971).

even though Tribe's skepticism toward the practicality of using Bayesian approaches at trial also may have had a dampening effect. William Twining's calls to action no doubt also helped,¹⁷⁵ as did conferences on New Evidence Scholarship organized by Peter Tillers.¹⁷⁶ David Schum, a nonlawyer, also deserves much credit for interesting law professors in the formal analysis of evidence.¹⁷⁷

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?¹⁷⁸ 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

¹⁷⁵ See TWINING, *RETHINKING EVIDENCE: EXPLORATORY ESSAYS*, *supra* note 1, *passim* and 12-31, 341-72 (1990).

¹⁷⁶ Professor Tillers was the chair (with Eric Green), the 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, the chair, and a panelist at the Cardozo Conference on Decision and Inference in Litigation, Cardozo School of Law, Yeshiva University, March 24-26, 1991. E-mail I from Peter Tillers, Professor of Law, Benjamin N. Cardozo School of Law, to Roger C. Park (July 8, 2005) (on file with the authors). 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. E-mail II from Peter Tillers, Professor of Law, Benjamin N. Cardozo School of Law, to Roger C. Park (July 8, 2005) (on file with the authors). For a European perspective on the development of the field, see TWINING, *RETHINKING EVIDENCE: EXPLORATORY ESSAYS*, *supra* note 1, *passim* and 349-50.

¹⁷⁷ See generally Peter Tillers, *Webs of Things in the Mind: A New Science of Evidence*, 87 MICH. L. REV. 1225 (1989) (review of David A. Schum's *EVIDENCE AND INFERENCE FOR THE INTELLIGENCE ANALYST* (1987), with comments on his contributions to the science of evidence).

¹⁷⁸ See generally COHEN, *supra* note 172 (describing anomalies that arise if Pascalian probability theory is applied in certain legal situations, and arguing that a "Baconian" approach would be more just).

the probability is actually over 50%, because a negative inference can be drawn from lack of other proof.¹⁷⁹

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 the plaintiff has established each element by a preponderance of the evidence.¹⁸⁰ 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.¹⁸¹ Moreover, the approach of telling juries to judge the 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

¹⁷⁹ See 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 Shavero, *Statistical-Probability Evidence and the Appearance of Justice*, 103 HARV. L. REV. 530, 533–36 (1989) [hereinafter Shavero, *Statistical-Probability Evidence and the Appearance of Justice*]. See generally 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) [hereinafter Shavero, *A Response to Professor Allen*]; Daniel Shavero, *A Response to Professor Callen*, 65 TUL. L. REV. 499 (1991) [hereinafter Shavero, *A Response to Professor Callen*].

¹⁸⁰ Ronald J. Allen, *A Reconceptualization of Civil Trials*, 66 B.U. L. REV. 401, 425–37 (1986).

¹⁸¹ 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, 949–52 (1986).

story, and the jury believes that a third story not offered by either party (but favoring the defendant) is the true one.¹⁸²

Another topic, the use of Bayes' Theorem to evaluate evidence and evidence law, has become one of the centerpieces of the New Evidence Scholarship.¹⁸³ 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 factfinder 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 (mtDNA) 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.

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, which is the random match probability in the general population. That does not mean, however, that there is a 1 in 100 probability that the defendant is the source, given the mtDNA 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, $\Pr(C|P)$, is obviously not the same as the probability that a person is in prison, given that he has committed a crime, $\Pr(P|C)$. 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 $\Pr(C|P)$ must be equal to $\Pr(P|C)$ 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 mtDNA test on the hair. Suppose there are 100,000 other people living in the vicinity and no reason except the mtDNA 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%. In other words, the

¹⁸² See Dale A. Nance, *Naturalized Epistemology and the Critique of Evidence Theory*, 87 VA. L. REV. 1551, 1575-80 (2001). For another discussion, see Posner, *supra* note 179, at 1512-15.

¹⁸³ For a useful introduction, see generally 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).

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%.¹⁸⁴ Of course, the probability that the defendant is not the source would decrease dramatically once other inculpatory evidence is offered, such as testimony that the defendant had a motive and was seen near the crime scene.

A factfinder who used Bayes' Theorem to take account of the mtDNA test could arrive at a probability that the defendant is the source without falling into transposition error. Suppose again that the factfinder estimates the prior odds that the defendant was the source (before the DNA test) to be 2 to 1. That factfinder 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 mtDNA 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%.¹⁸⁵ If the defendant is not the source, the probability of finding a match is 1/100 or 1% (the probability of a random match in the general population).¹⁸⁶ The likelihood ratio therefore is 100% divided by 1%, or 100. To calculate the posterior odds, one would mul-

¹⁸⁴ For purposes of the example, we are ignoring the possibility that someone not living in the vicinity is the source.

¹⁸⁵ 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 mtDNA profiles. But here we assume that we would find a match every time if the hair came from the defendant.

¹⁸⁶ 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 factfinder 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 & 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).

tively the prior odds of 2 to 1 by the likelihood ratio of 100. Thus, the factfinder using this method would conclude that the updated odds that the defendant was the source are 200 to 1.

In an article published early in the Bayesian debate, Finkelstein and Fairley 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.¹⁸⁷ 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.¹⁸⁸ 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.¹⁸⁹

In a celebrated reply, Professor Laurence Tribe, then an evidence teacher, raised several objections to this use of Bayes' Theorem in the trial process.¹⁹⁰ Jurors who are not proposition bettors might have mistaken or inconsistent understandings of the meaning of prior probability.¹⁹¹ There is a danger of "[d]warfing . . . [s]oft [v]ariables," 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).¹⁹² Moreover, uncertainty about facts upon which the Bayesian calculations would be based could require additional quantification decisions about so many issues that use of the Theorem would be more confusing than helpful.¹⁹³ 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 factfinder to adjust the statistics. Tribe's analysis was in turn criticized by psychologists for making

¹⁸⁷ See generally Finkelstein & Fairley, *supra* note 174.

¹⁸⁸ *Id.* at 496.

¹⁸⁹ *Id.* at 498-500.

¹⁹⁰ Tribe, *supra* note 174, at 1358-77.

¹⁹¹ 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. *Id.* at 1358-59.

¹⁹² *Id.* at 1361-65.

¹⁹³ 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' Theorem, and giving instructions about how to do this would be too complicated to be feasible. *Id.*

faulty psychological assumptions¹⁹⁴ as well as by statisticians,¹⁹⁵ philosophers,¹⁹⁶ and legal scholars.¹⁹⁷

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 the factfinder. 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 prior probability assessments 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.¹⁹⁸

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 such a use.¹⁹⁹ 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.²⁰⁰ 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.²⁰¹ It can, for example, help us understand why prior convictions have little probative value in im-

¹⁹⁴ David L. Faigman & A.J. Baglioni, Jr., *Bayes' Theorem in the Trial Process*, 12 LAW & HUM. BEHAV. 1, 13-14 (1988); Michael J. Saks & Robert F. Kidd, *Human Information Processing and Adjudication: Trial by Heuristics*, 15 LAW & SOC'Y REV. 123, 123-45 (1980); Brian C. Smith, Steven D. Penrod, Amy L. Otto & Roger C. Park, *Jurors' Use of Probabilistic Evidence*, 20 LAW & HUM. BEHAV. 49, 78 (1996).

¹⁹⁵ Richard D. Friedman, *Answering the Bayesioskeptical Challenge*, 1 INT'L J. EVIDENCE & PROOF 276, 285-86 (1996).

¹⁹⁶ COHEN, *supra* note 172, at 53-56.

¹⁹⁷ Stephen E. Fienberg & Mark J. Schervish, *The Relevance of Bayesian Inference for the Presentation of Statistical Evidence and for Legal Decisionmaking*, 66 B.U. L. REV. 771, 784-86 (1986).

¹⁹⁸ See Ronald J. Allen, *Rationality, Algorithms and Judicial Proof: A Preliminary Inquiry*, 1 INT'L J. EVIDENCE & PROOF 254, 265-69 (1997); Mike Redmayne, *Presenting Probabilities in Court: The DNA Experience*, 1 INT'L J. EVIDENCE & 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 INT'L J. EVIDENCE & PROOF 25, 34-47 (1997).

¹⁹⁹ Friedman, *supra* note 195, at 290.

²⁰⁰ *Id.*

²⁰¹ *Id.*

peaching the testimony of the accused in a criminal case.²⁰² 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. 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.²⁰³ 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.²⁰⁴ Scholars also have relied on Bayesian models to analyze such diverse evidentiary issues as the meaning of the concept of relevance,²⁰⁵ the value of forensic identification evidence,²⁰⁶ the proper interpretation of DNA evidence,²⁰⁷ the value of expert testimony in child abuse cases,²⁰⁸ the value of hearsay,²⁰⁹ and the appropriateness of questioning children in a suggestive manner.²¹⁰

²⁰² See Friedman, *supra* note 23, at 659–66; Posner, *supra* note 179, at 1526–27, 1533–35. Professor Uviller arrived at a similar conclusion without reference to Bayes' Theorem. See Uviller, *supra* note 23, at 867–68 (“[I]t 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.”).

²⁰³ See generally Richard O. Lempert, *Modeling Relevance*, 75 MICH. L. REV. 1021 (1977) [hereinafter Lempert, *Modeling Relevance*]; Richard Lempert, *Of Flutes, Oboes, and the As If World of Evidence Law*, 1 INT'L J. EVIDENCE & PROOF 316 (1997).

²⁰⁴ Richard D. Friedman & Roger C. Park, *Sometimes What Everybody Thinks They Know Is True*, 27 LAW & HUM. BEHAV. 629, 637 n.15 (2003).

²⁰⁵ Lempert, *Modeling Relevance*, *supra* note 203, at 1021–32.

²⁰⁶ BERNARD ROBERTSON & G.A. VIGNAUX, INTERPRETING EVIDENCE 16–29 (1995).

²⁰⁷ See generally IAN W. EVETT & BRUCE S. WEIR, INTERPRETING DNA EVIDENCE: STATISTICAL GENETICS FOR FORENSIC SCIENTISTS (1998).

²⁰⁸ Thomas D. Lyon & Jonathan J. Koehler, *The Relevance Ratio: Evaluating the Probative Value of Expert Testimony in Child Sexual Abuse Cases*, 82 CORNELL L. REV. 43, 46–50 (1996).

²⁰⁹ Peter Tillers & David Schum, *Hearsay Logic*, 76 MINN. L. REV. 813, 832–52 (1992).

²¹⁰ Stephen J. Ceci & Richard D. Friedman, *The Suggestibility of Children: Scientific Research and Legal Implications*, 86 CORNELL L. REV. 33, 76–81 (2000). For other examples of applications of Bayes in evaluation of evidence law, see generally Richard D. Friedman, *Towards a (Bayesian) Convergence?*, 1 INT'L J. EVIDENCE & PROOF 348 (1997).

A number of Bayesian enthusiasts believe that the Theorem can be useful at trial, though few would argue for applying Bayesian analysis iteratively 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.²¹¹ This approach is still problematic, for reasons discussed by Tribe²¹² and Allen,²¹³ and courts have been slow to adopt it.²¹⁴ 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.²¹⁵ 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.²¹⁶ The National Research

²¹¹ See Nance, *supra* note 182, 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, 437–45 (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. Cordner, 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, 456–59 (1997); see also ROBERTSON & VIGNAUX, *supra* note 206, at 20–21, 51–65 (supporting an approach under which a forensic expert explains to the trier of fact that “[t]his evidence is R times more probable if the accused left the mark than if someone else did,” and that “[t]his evidence therefore [very strongly] supports the proposition that the accused left the mark”).

²¹² Tribe, *supra* note 174, at 1355–78.

²¹³ Allen, *supra* note 198, at 265–69.

²¹⁴ On the infrequency of use of these Bayesian presentation methods except in paternity cases, see D.H. KAYE ET AL., *supra* note 141, § 12.4.2(b), at 478 (“[L]ikelihood ratios are rarely introduced in criminal cases . . .”).

²¹⁵ D.H. Kaye, Plemel as a Primer on Proving Paternity, 24 WILLAMETTE L. REV. 867, 881 (1988).

²¹⁶ See, e.g., *Kammer v. Young*, 535 A.2d 936, 938–40 (Md. Ct. Spec. App. 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 ROBERTSON & VIGNAUX, *supra* note 206, at 25–27; Jonathan J. Koehler, *DNA Matches and Statistics: Important Questions, Surprising An-*

Council has suggested that experts might compute posterior probabilities to show jurors the power of DNA evidence for establishing identity,²¹⁷ although this proposal also remains controversial.²¹⁸

When Bayesian skeptics and Bayesian enthusiasts address common ground, there seems to be a degree of convergence. Professor Allen, a leading Bayesian 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 skeptic does not deny a use for Bayes' Theorem as an analytical tool."²¹⁹ 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; factfinders can deal with the evidence much as they deal with ordinary questions in their everyday lives."²²⁰

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.²²¹ 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 decisionmakers. Jurors appear to evaluate evidence in a trial not

swers, 76 JUDICATURE 222, 225-26 (1993). 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.

²¹⁷ NAT'L RESEARCH COUNCIL, THE EVALUATION OF FORENSIC DNA EVIDENCE 199-202 (1996).

²¹⁸ William C. Thompson, *Accepting Lower Standards: The National Research Council's Second Report on Forensic DNA Evidence*, 37 JURIMETRICS 405, 422-23 (1997).

²¹⁹ Allen, *supra* note 198, at 258.

²²⁰ Friedman, *supra* note 195, at 290; *see also* John A. Michon, *The Time Has Come to Put This Debate Aside and Move On to Other Matters*, 1 INT'L J. EVIDENCE & PROOF 331, 334 (1997).

²²¹ Faigman & Baglioni, *supra* note 194, at 13-16; Jane Goodman, *Jurors' Comprehension and Assessment of Probabilistic Evidence*, 16 AM. J. TRIAL ADVOC. 361, 384-89 (1992); Jason Schklar & Shari Seidman Diamond, *Juror Reactions to DNA Evidence: Errors and Expectancies*, 23 LAW & HUM. BEHAV. 159, 178-82 (1999); Smith et al., *supra* note 194, at 73-78; 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, 181-85 (1987).

by sequential updating but by constructing plausible narratives that might account for the evidence.²²² As some researchers 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 might come closer to the Bayesian norm, by use of a Bayesian chart showing the trier how estimates of prior probability should be changed by the forensic evidence.²²³

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.²²⁴ 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 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?

Wigmore²²⁵ developed the first system for organizing and assessing evidence for litigation by employing careful logic to trace the factual support for inferences. "Wigmoreian analysis is an attempt to capture the way we think when we think at our best."²²⁶ Wigmoreian charting was a major milestone in lawyerly thinking about facts, but it still was somewhat crude. Anderson and Twining²²⁷ not only resurrected Wigmoreian 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. Wig-

²²² Nancy Pennington & Reid Hastie, *A Cognitive Theory of Juror Decision Making: The Story Model*, 13 CARDOZO L. REV. 519, 549-50 (1991); Nancy Pennington & Reid Hastie, *Explaining the Evidence: Tests of the Story Model for Juror Decision Making*, 62 J. PERSONALITY & SOC. PSYCHOL. 189, 202-03 (1992).

²²³ See Nance, *supra* note 182, at 1610-16; Nance & Morris, *supra* note 211, at 437-45.

²²⁴ William Twining, *The New Evidence Scholarship*, 13 CARDOZO L. REV. 295, 295 (1991). For a leading example, see generally Peter Tillers & David Schum, *A Theory of Preliminary Fact Investigation*, 24 U.C. DAVIS L. REV. 931 (1991).

²²⁵ WIGMORE, *supra* note 168.

²²⁶ Bernard Robertson & G.A. Vignaux, *Taking Fact Analysis Seriously*, 91 MICH. L. REV. 1442, 1447 (1993).

²²⁷ TERENCE ANDERSON & WILLIAM TWINING, *ANALYSIS OF EVIDENCE: HOW TO DO THINGS WITH FACTS BASED ON WIGMORE'S SCIENCE OF JUDICIAL PROOF*, *passim* (1991).

morian and related kinds of charting²²⁸ 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 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.²²⁹

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 essentially have concluded that lawyers and evidence scholars need not worry about the implications of profound skepticism.²³⁰ 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.²³¹ Professor Seigel 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

²²⁸ Namely, the improvements offered by ANDERSON & TWINING, *id.*, and by such devices as decision trees, influence diagrams, and Bayes networks. See generally Robertson & Vignaux, *supra* note 226.

²²⁹ Robertson & Vignaux, *supra* note 226, at 1456.

²³⁰ See WILLIAM TWINING, *Some Scepticism About Some Scepticisms*, in *RETHINKING EVIDENCE: EXPLORATORY ESSAYS*, *supra* note 1, at 92, 94, 134 (stating that few philosophers consistently maintain philosophical skepticism, and the literature "poses few threats to the centrality of the concepts of Truth, Reason and Justice in any theory of Evidence"); Ronald J. Allen, *Truth and Its Rivals*, 49 HASTINGS L.J. 309, 315-318 (1998) ("There is indeed an important philosophical puzzle, but it primarily is how to explain what is obviously true." *Id.* at 315. "[T]he implications of scepticism for the legal system are marginal at best" because all the rivals to truth as a goal of litigation themselves involve rejection or cabining of the philosophical problem of scepticism. *Id.* at 317.); Mirjan Damaška, *Truth in Adjudication*, 49 HASTINGS L.J. 289, 290 (1998) (stating that "when we engage in . . . adjudication, we presuppose a world beside our statements," and arguing that radical postmodern thought has no bearing on evidence law). But cf. Donald Nicolson, *Truth, Reason, and Justice: Epistemology and Politics in Evidence Discourse*, 57 MOD. L. REV. 726, 743-44 (1994). Nicolson's arguments are critiqued in Jackson, *supra* note 1, at 898-99.

²³¹ See WILLIAM TWINING, *The Rationalist Tradition of Evidence Scholarship*, in *RETHINKING EVIDENCE: EXPLORATORY ESSAYS*, *supra* note 1, at 32, 33.

making verdicts more acceptable to the public and ending disputes in an efficient way.²³² In contrast, and without relying to the same degree upon the literature on philosophy, Professor Nesson has argued that the legal system already gives primacy to making verdicts acceptable instead of to finding the truth.²³³

Professors Allen and Leiter have sought to bring lawyers up to date on epistemology in an exploration of contemporary work on naturalized epistemology.²³⁴ Though their work is tough sledding for readers without a background in philosophy, it reaches conclusions that should be comforting to interdisciplinary evidence scholars. Allen and Leiter themselves note that "[f]or the great bulk of evidentiary scholars, then, this paper merely solidifies the ground beneath their feet."²³⁵ They maintain that philosophy should not be an a priori discipline, but one that is continuous with an a posteriori inquiry in the empirical sciences.²³⁶ 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 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.²³⁷ Finally, because naturalized epistemology is instrumental, evidence rules should only require intellectual performances that factfinders 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.²³⁸

²³² Seigel, *supra* note 117, at 996–99. For comments on critiques of the rationalist tradition, see Jackson, *supra* note 1, at 898–901.

²³³ Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357, 1359–68 (1985). For critique, see generally Ronald J. Allen, *Rationality, Mythology, and the 'Acceptability of Verdicts' Thesis*, 66 B.U. L. REV. 541 (1986).

²³⁴ See generally Allen & Leiter, *supra* note 2.

²³⁵ *Id.* at 1493.

²³⁶ *Id.* at 1494–97.

²³⁷ *Id.* at 1503, 1511, 1516–25.

²³⁸ *Id.* at 1492–93, 1499. For a rebuttal, see generally Nance, *supra* note 182. Cf. Brian Leiter, *The Epistemology of Admissibility: Why Even Good Philosophy of Science Would Not Make for Good Philosophy of Evidence*, 1997 BYU L. REV. 803, 815–19 (arguing that *Daubert* makes unrealistic demands on the epistemic capacities of adjudicators).

V. 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 doctrines that relate to sexual assault and spousal abuse. Topics relating to sexual assault have addressed rape shield rules,²³⁹ exceptions to the rule against character evidence for prior crimes of the alleged perpetrator,²⁴⁰ and admissibility of rape trauma syndrome testimony.²⁴¹ Topics concerning spousal abuse have included whether victims should be forced to testify over a claim of spousal immunity,²⁴² whether prior acts of domestic violence against other victims should be admissible,²⁴³ and whether social science evidence in spousal abuse cases, including testimony about battered woman's syndrome, should be admissible.²⁴⁴ Feminist scholars also

²³⁹ See generally Orenstein, *supra* note 19 (supporting rape shield legislation but opposing legislation admitting prior sex crimes of the alleged perpetrator).

²⁴⁰ See generally Baker, *supra* note 19; Orenstein, *supra* note 19.

²⁴¹ 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) ("[Rape trauma syndrome] evidence can help the fact finder resolve difficult issues of guilt or innocence and . . . 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.").

²⁴² For scholars who have urged that the victim not be allowed to invoke spousal immunity to refuse to testify against a battering husband, see generally 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) and Malinda L. 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).

²⁴³ See generally Lisa Marie De Sanctis, *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); 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).

²⁴⁴ Battered woman's syndrome ("BWS") 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 E. WALKER, *THE BATTERED WOMAN* (1979); LENORE E. WALKER, *THE BATTERED WOMAN SYNDROME* (1984) [hereinafter WALKER, *THE BATTERED WOMAN SYNDROME*]. 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 where dogs that could not control the electric shocks they were receiving

have examined evidence issues whose impact on women is less obvious, including the excited utterance exception,²⁴⁵ the party admission doctrine,²⁴⁶ the use of inconsistent statements to impeach,²⁴⁷ and the increased rigor *Daubert* demands of expert testimony.²⁴⁸

simply stopped trying to avoid them. WALKER, THE BATTERED WOMAN SYNDROME, *supra*, at 86. 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. Cf. Myrna S. Raeder, *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"). See generally 2 MODERN SCIENTIFIC EVIDENCE, *supra* note 102, § 15 (reviewing the empirical research on which BWS theory is based); 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 post-traumatic stress disorder testimony); David L. Faigman & Amy J. Wright, *The Battered Woman Syndrome in the Age of Science*, 39 ARIZ. L. REV. 67 (1997); Janet C. Hoefel, *The Gender Gap: Revealing Inequities in Admission of Social Science Evidence in Criminal Cases*, 24 U. ARK. LITTLE ROCK L. REV. 41 (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). Some authors have suggested substitutes, such as admitting evidence of social science findings about domestic abuse. See 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, 151-52 (1997).

²⁴⁵ See generally Aviva Orenstein, "MY GOD!": A Feminist Critique of the Excited Utterance Exception to the Hearsay Rule, 85 CAL. L. REV. 159 (1997) (excited utterance exception reflects male perspective; calm statements of sexual assault victims should also be admissible, subject to conditions).

²⁴⁶ See generally Aviva Orenstein, *Apology Excepted: Incorporating a Feminist Analysis into Evidence Policy Where You Would Least Expect It*, 28 SW. U. L. REV. 221 (1999) (contending that 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 generally Jennifer K. Robbenolt, *Apologies and Legal Settlement: An Empirical Examination*, 102 MICH. L. REV. 460 (2004).

²⁴⁷ In two articles, Kim Lane Scheppelle 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 generally Kim Lane Scheppelle, *The Ground-Zero Theory of Evidence*, 49 HASTINGS L.J. 321 (1998) [hereinafter Scheppelle, *Ground-Zero*]; Kim Lane Scheppelle, *Just the Facts, Ma'am: Sexualized Violence, Evidentiary Habits, and the Revision of Truth*, 37 N.Y.L. SCH. L. REV. 123 (1992) [hereinafter Scheppelle, *Just the Facts, Ma'am*].

²⁴⁸ See Andrew E. Taslitz, *What Feminism Has to Offer Evidence Law*, 28 SW. U. L. REV. 171, 211 (1999) [hereinafter Taslitz, *What Feminism Has to Offer Evidence Law*] (favoring an approach to expert testimony that allows "all relevant voices on perspectives" to be heard, and stating that "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 to convey an excluded group's voice"). See generally Andrew E. Taslitz, *A Feminist Approach to Social Scientific Evidence: Foundations*, 5 MICH. J. GENDER & L. 1 (1998) [hereinafter Taslitz, *A Feminist Approach to Social Scientific Evidence*].

A number of feminist evidence scholars, noting that scholars from different strands of feminism might have different views about law, have classified feminist legal theory into three categories.²⁴⁹ The first is liberal feminism, sometimes dubbed "sameness" feminism. Liberal feminists are said to favor formal equality with men and assume that women can compete in the same way as men; they do not emphasize differences between men and women.²⁵⁰ 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.²⁵¹ 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.²⁵² These authors express a wide variety of views, some of them shared with strains of critical legal theory or postmodernism. These views include valuing

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 Catharine 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.

CATHARINE A. MACKINNON, *FEMINISM UNMODIFIED* 54 (1987). One can also find traces of an ascientific attitude in articles that assess the validity or social value of battered women's 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 U. ILL. L. REV. 413, 445-47.

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) eventually revealed that such treatment was dangerous, but not before decades of misguided estrogen replacement had caused tens of thousands of unnecessary breast cancers, heart attacks, and strokes in women. JERRY AVORN, *POWERFUL MEDICINES: THE BENEFITS, RISKS, AND COSTS OF PRESCRIPTION DRUGS* 23-38 (2004).

²⁴⁹ See Seymore, *supra* note 242, at 1066-70 (dividing the feminist legal movement into "liberal feminism," "relational feminism," and "radical feminism"); Taslitz, *What Feminism Has to Offer Evidence Law*, *supra* note 248, at 175-76 (dividing into "liberal feminism," "radical feminism," and "cultural feminism"); cf. Orenstein, *supra* note 246, at 226-27 (dividing into "difference feminism," "dominance feminism," and "postmodern feminism").

²⁵⁰ Taslitz, *What Feminism Has to Offer Evidence Law*, *supra* note 248, at 176.

²⁵¹ *Id.* at 175.

²⁵² *Id.*

intuition and emotion; believing in contextualized thinking instead of abstract, rigid rules; and emphasizing the value of relationships over market-like competition.²⁵³

Some feminist scholars emphasize the value of narratives over, or as an equal partner with, quantitative social science.²⁵⁴ 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.²⁵⁵ Many scholars are attentive to the law's constitutive role in helping define society and perpetuate views of the world.²⁵⁶ Sometimes one also sees a belief in the pervasiveness and inevitability of politics, and a result-oriented approach that deprecates the wisdom or feasibility of achieving objective solutions to social problems.²⁵⁷ 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.²⁵⁸

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 situa-

²⁵³ 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." Kinports, *supra* note 248, at 417. See generally Taslitz, *What Feminism Has to Offer Evidence Law*, *supra* note 248.

²⁵⁴ As an example, see generally Mahoney, *supra* note 248.

²⁵⁵ "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 242, at 1071.

²⁵⁶ See, e.g., Baker, *supra* note 19, at 591 (discussing whether Rule 413 spreads or diminishes rape myths, and concluding that it perpetuates a stereotype of the chronic rapist); Marilyn MacCrimmon, *The Social Construction of Reality and the Rules of Evidence*, 25 U.B.C. L. Rev. 36, 49-50 (1991); Mahoney, *supra* note 248, at 6-7 (noting interaction of law and culture and advocating "separation assault" as a legal concept that can "reshape cultural understanding"); Orenstein, *supra* note 19, at 692 (opposing Rule 413, in part because it perpetuates misconception that rape is not pervasive); Taslitz, *What Feminism Has to Offer Evidence Law*, *supra* note 248, at 180.

²⁵⁷ See Taslitz, *What Feminism Has to Offer Evidence Law*, *supra* note 248, at 213-17.

²⁵⁸ See WILLIAM TWINING, *The Rationalist Tradition of Evidence Scholarship*, in *RETHINKING EVIDENCE: EXPLORATORY ESSAYS*, *supra* note 1, at 32-33.

tion militates against having fixed rules, and in favor of discretionary decision making. But if decisionmakers have attitudes tainted by sexism, then fixed rules are needed to protect women, because 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.

Moreover, evidence law is adjective, 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 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.²⁵⁹

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 refer to four leading articles in which dominance feminism or cultural feminism appears to have been particularly influential on the analysis or result.²⁶⁰

²⁵⁹ See generally Baker, *supra* note 19; Orenstein, *supra* note 19.

²⁶⁰ Baker, *supra* note 19; Mahoney, *supra* note 248; Orenstein, *supra* note 19; Scheppele, *Just the Facts, Ma'am*, *supra* note 247.

Our first examples are two articles about Federal Rule of Evidence 413, which allows evidence that the accused committed other sexual assaults to be admitted in a sexual assault prosecution.²⁶¹ Aviva Orenstein and Katharine Baker, working independently, both came to the conclusion that the legislation was unwise and unjustifiable.²⁶² Strains of cultural feminism and dominance feminism can be seen in both works. The authors see the legislation as unwise because it decontextualizes the situation by stereotyping rapists, treating them as pathological outlaws rather than normal men engaged in situational conduct.²⁶³ The authors also believe the legislation perpetuates rape "myths" about rape being strange and deviant, whereas the authors see rape as common and widespread, and as a way that society controls women.²⁶⁴ 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.²⁶⁵

Another example of scholarship that seems influenced by strands of cultural feminism is Martha Mahoney's article about spousal abuse cases.²⁶⁶ She addresses the question of why women do not leave abusive relationships. Much of her article consists of narratives, including her own, of battered women.²⁶⁷ She avoids criticism of Lenore Walker's controversial work,²⁶⁸ even though she has mixed feelings about the

²⁶¹ FED. R. EVID. 413.

²⁶² Baker, *supra* note 19, at 623–24; Orenstein, *supra* note 19, at 690–97. The rape statistics are controversial. See NEIL GILBERT, *Miscounting Social Ills: Sexual Assault and Advocacy Research*, in WELFARE JUSTICE: RESTORING SOCIAL EQUALITY 84, 99–108 (1995); Park, *supra* note 62, at 765–69.

²⁶³ Baker, *supra* note 19, at 576–83; Orenstein, *supra* note 19, at 690–97.

²⁶⁴ Baker, *supra* note 19, at 576–78; Orenstein, *supra* note 19, at 692–93.

²⁶⁵ Baker, *supra* note 19, at 578–83; Orenstein, *supra* note 19, at 693.

²⁶⁶ Mahoney, *supra* note 248.

²⁶⁷ *Id.* at 7–8. She states:

[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. (citations omitted).

²⁶⁸ WALKER, *THE BATTERED WOMAN SYNDROME*, *supra* note 244. On Walker's methodology, see David L. Faigman, Note, *The Battered Woman Syndrome and Self-Defense: A Legal and Empirical Dissent*, 72 VA. L. REV. 619, 641–47 (1986).

message sent by learned helplessness,²⁶⁹ 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.²⁷⁰

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.²⁷¹ While she does not reject quantitative social science, she seems to have an attitude toward proof of social facts that differs 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.²⁷² She uses anecdotal evidence extensively, and regards it as a source of convincing proof instead of a source of hypotheses to be tested.²⁷³ 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

²⁶⁹ Mahoney, *supra* note 248, at 42. She states:

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. 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. 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. (citations omitted).

²⁷⁰ *Id.* at 7. She states:

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.

²⁷¹ *Id.* at 7–8.

²⁷² *See id.* at 29–30.

²⁷³ *Id.* at 20–24.

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, and those that do rarely 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 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,²⁷⁴ 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 Scheppele entitled *Just the Facts, Ma'am: Sexualized Violence, Evidentiary Habits, and the Revision of Truth*.²⁷⁵ This article, using the Anita Hill hearings as its point of departure, posits that impeachment of witnesses hurts women disproportionately because women delay in making a complaint or give inconsistent versions of a complaint at different times.²⁷⁶ In cases of sexualized violence, such as rape, sexual harassment, or spousal abuse, victims often delay in reporting the incident and do not report a full or accurate version the first time.²⁷⁷ 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 that the dominant culture might blame the victim, for example, by saying that she provoked the attack.²⁷⁸ 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.²⁷⁹

In addition to her points that are specific to women and sexualized violence, Scheppele has a more general point rooted in post-modern epistemology. Invoking Wittgenstein, she notes that accounts of events ("stories") are narratives that are influenced by interpretive

²⁷⁴ Mahoney, *supra* note 248, 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").

²⁷⁵ Scheppele, *Just the Facts, Ma'am*, *supra* note 247.

²⁷⁶ *Id.* at 126–28.

²⁷⁷ *Id.* at 138–39.

²⁷⁸ *Id.* at 142–43.

²⁷⁹ *Id.* at 138–41.

frameworks.²⁸⁰ 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.²⁸¹ But the difference in accounts that the feminist lawyer and the battered woman would give is not, Scheppele writes, a difference "between truth and falsehood," but a difference in interpretive frameworks.²⁸² Consciousness raising may cause the same person to see the same event in different ways.²⁸³ The second 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."²⁸⁴

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.²⁸⁵ The revised accounts are not excluded from evidence, of course, but merely subject to impeachment by prior inconsistent statements.²⁸⁶ Although 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

²⁸⁰ Scheppele, *Just the Facts, Ma'am*, *supra* note 247, at 167–68.

²⁸¹ *Id.* at 168.

²⁸² *Id.*

²⁸³ *Id.* at 168–70.

²⁸⁴ *Id.* at 172.

²⁸⁵ "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." Scheppele, *Just the Facts, Ma'am*, *supra* note 247, at 169–70.

²⁸⁶ Scheppele recognizes this point in a later article on the same subject, in which she says:

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 247, at 330.

at the time of the events that led to the litigation.²⁸⁷ And even 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. Nevertheless, 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 citation to Elizabeth Loftus.²⁸⁸ Some of the psychology scholarship on eyewitnesses would be helpful to her argument—studies suggesting, for example, that consistency in description of suspects is not a strong predictor of accuracy.²⁸⁹ 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.²⁹⁰

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 that "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.²⁹¹ This attitude may explain why discussions of battered woman's syndrome sometimes seem result-oriented, as if the portrait that the theory paints of women is more

²⁸⁷ *Id.* at 334.

²⁸⁸ Scheppele, *Just the Facts, Ma'am*, *supra* note 247, at 167 n.174.

²⁸⁹ See CUTLER & PENROD, *supra* note 40, at 93–95.

²⁹⁰ In her more recent 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 247, at 325–26, she notes psychological experiments on the distortion 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.

²⁹¹ Taslitz, *What Feminism Has to Offer Evidence Law*, *supra* note 248, at 211.

important in judging its acceptability than the validity of the research methods that produced it.

Of course, much legal thinking and law rely on anecdotes and fireside inductions. In that, feminist scholarship is entirely mainstream. Moreover, many (and probably most) of the hypotheses developed by feminist scholars are amenable to empirical testing. Feminist scholarship generates ideas that are unlikely to have been thought of without the theoretical discipline that gave rise to them. The empirical soundness of many of those new hypotheses deserves to be studied and tested. Much the same can be said for evidence law and economics scholarship, to which we turn next.

VI. ECONOMICS AND EVIDENCE

Until quite recently, it could be said that law and economics scholars had virtually nothing to say about evidence law.²⁹² This is not exactly surprising. Evidence scholarship focuses on understanding, explaining, evaluating, and suggesting improvements for rules that are concerned principally with the goal of maximizing the ability of trials to discover the truth of a matter in dispute.²⁹³ It is easy to see how various fields might contribute to this endeavor, such as logic,

²⁹² See Park, *supra* note 8, at 849 n.2. 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. BENTHAM, *supra* note 3. 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 twentieth century. See TWINING, *supra* note 3, at ix.

²⁹³ See, e.g., William Twining, *Evidence and Legal Theory*, 47 MOD. L. REV. 261, 272 (1984). Twining states:

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.

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

psychology and other cognitive sciences, philosophy, statistics, feminism, and so on. But economics?

The first great law and economics movement, which arose in the late nineteenth century, involved the macroeconomics of law.²⁹⁴ It was concerned with political economy, the behavior of markets, and economic systems, and was reflected in areas of law such as 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—involves the microeconomics of law, the pursuit of efficiency and wealth maximization.²⁹⁵ 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-gauged attempt to apply microeconomics to evidence law. To be sure, there were occasional articles on incentives and disincentives for gathering evidence,²⁹⁶ how those incentives affect the evidence offered to courts,²⁹⁷ the resulting outcomes of trials and the impact of those outcomes on behavior (especially economic activity) outside of court.²⁹⁸ These studies focused mainly on the problems of assembling evidence for trials and the effects of verdicts. Recently, we have seen broader attempts to apply economics to evidence law and the philosophy of evidence. We focus here on three authors whose work illustrates this new contribution to evidence scholarship.

Perhaps fittingly, the first broad major law and economics treatment of evidence law, tackling a wide range of evidence topics, is by Richard Posner.²⁹⁹ In his article, Judge Posner's assumptions about rational planning often lead him to inferences about 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 cus-

²⁹⁴ See Herbert Hovenkamp, *The First Great Law & Economics Movement*, 42 STAN. L. REV. 993, 994–95 (1990).

²⁹⁵ See generally RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* (6th ed. 2003).

²⁹⁶ Andrew F. Daughety & Jennifer F. Reinganum, *On the Economics of Trials: Adversarial Process, Evidence, and Equilibrium Bias*, 16 J.L. ECON. & ORG. 365 (2000).

²⁹⁷ Jeffrey S. Parker, *Daubert's Debut: The Supreme Court, the Economics of Scientific Evidence, and the Adversarial System*, 4 SUPREME CT. ECON. REV. 1 (1995).

²⁹⁸ 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).

²⁹⁹ Posner, *supra* note 179.

tomary ways of doing things in order to obtain an advantage if they ever wind up in litigation, assumptions which sometimes seem unrealistic.³⁰⁰ As might be expected, Judge Posner also brings to his study of evidence law a sensitivity to costs, trade-offs, and substitutions.³⁰¹

Posner's rational choice, ex ante perspective is not the best starting point when drawing inferences about the issues 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 aiming to control reasoning at trial rather than future primary conduct.³⁰² They seek to protect against mistaken, unreasonable, or lawless interpretations of evidence by factfinders. But Posner is fearless in applying his perspective, even to seemingly unpromising topics such as character evidence.³⁰³ We describe and discuss some of those ideas. First, we present 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 damages remedies for illegal searches.³⁰⁴ 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.³⁰⁵ Therefore, he suggests, those who oppose *Mapp v. Ohio*³⁰⁶ ought to be arguing about the definition of illegal search rather than about the sanction.³⁰⁷ 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.³⁰⁸ 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

³⁰⁰ See *id.* at 1529, 1531.

³⁰¹ *Id.* at 1542-46.

³⁰² But see the discussion of Sanchirico's work, *infra* notes 320-332 and accompanying text.

³⁰³ For more thorough critiques of Posner's article, see Allen & Leiter, *supra* note 2, at 1510-27; Lempert, *supra* note 2, at 1639-700. See generally Park, *supra* note 2.

³⁰⁴ Posner, *supra* note 179, at 1533.

³⁰⁵ *Id.*

³⁰⁶ 367 U.S. 643 (1961).

³⁰⁷ Posner, *supra* note 179, at 1533.

³⁰⁸ For commentary on the blue bus problem, see generally Allen, *supra* note 179; Callen, *supra* note 179; Shaviro, *A Response to Professor Allen*, *supra* note 179; Shaviro, *A Response to Professor Callen*, *supra* note 179; Shaviro, *Statistical-Probability Evidence and the Appearance of Justice*, *supra* note 179.

“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.³⁰⁹

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.³¹⁰

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

³⁰⁹ Posner, *supra* note 179, at 1510.

³¹⁰ 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.

busses, because it would be less careful; moreover, a monopoly would eventually be created because A, burdened by higher liability costs, would withdraw from the route.³¹¹

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.³¹² At any rate, the consequence predicted is speculative and 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.³¹³ 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, 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 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 the following regarding prior-crime evidence:

It 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 subse-

³¹¹ Posner, *supra* note 179, at 1510.

³¹² 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* note 179, at 1510 (Posner's present comment that allowing naked statistical evidence to sustain "blue bus" verdict might lead to bus monopoly); *id.* at 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 Lempert, *supra* note 2, at 1671, 1690. 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.

³¹³ Allen & Leiter, *supra* note 2, at 1526.

quent offense may be reduced to the same level as the propensity to commit a first offense.³¹⁴

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.³¹⁵ 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,³¹⁶ can be positively misleading to policymakers unless the scholar is willing to check his assumptions against potentially disconfirming data.³¹⁷

Marital privilege is another area where Posner's analysis sometimes gets out of control, requiring more human foresight, knowledge, and flexibility than is plausible. It is doubtful 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.³¹⁸

We agree with Allen and Leiter that Judge Posner's article is prone to speculative theorizing.³¹⁹ We cannot say for certain that

³¹⁴ Posner, *supra* note 179, at 1525.

³¹⁵ See Park, *supra* note 62, at 758–63. Prediction systems using actuarial methods commonly use prior crimes as a predictor. See MICHAEL R. GOTTFREDSON & TRAVIS HIRSCHI, A GENERAL THEORY OF CRIME, *passim* and 107 (1990) (stating that “research 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 & HUM. SERVS., 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.”). See generally PETER HOFFMAN, PREDICTING CRIMINALITY, U.S. DEPT OF JUSTICE STUDY GUIDE (1988); VERNON L. QUINSEY, GRANT T. HARRIS, MARNIE E. RICE & CATHERINE A. CORMIER, VIOLENT OFFENDERS: APPRAISING AND MANAGING RISK (1998).

³¹⁶ 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 GOTTFREDSON & HIRSCHI, *supra* note 315, *passim* and 107.

³¹⁷ Park, *supra* note 2, at 2057–58.

³¹⁸ Posner, *supra* note 179, at 1530–31.

³¹⁹ Allen & Leiter, *supra* note 2, at 1521–27. However, for a similar economic analysis in a case that reaches a much more plausible legal conclusion, see Judge Posner's opinion in *Am. Hosp. Supply Corp. v. Hosp. Prods. Ltd.*, 780 F.2d 589, 593 (7th Cir. 1986) (opining that decision rules in preliminary motions should minimize the cost of error).

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, many of Posner's economic ideas about evidence law appear unpromising.

In contrast to Posner's shotgun approach to applying economic ideas to evidence law, the scholars whose work we examine next use a laser.

The first work 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 the rules on (mostly) crime deterrence.³²⁰ More specifically, Professor Sanchirico argues that the rule prohibiting the use of character evidence for propensity reasons cannot be explained coherently or convincingly in terms of enhancing the accuracy of trial fact finding.³²¹ He reviews each of the major extant truth-focused explanations for the rule—namely, the limited probative value of character evidence, the strong tendency of the jury to overweigh such evidence, the temptation of the jury to impose 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 biased distribution of persons selected for prosecution that further reduces the inferential value of character evidence)—and argues that they are unconvincing.³²²

³²⁰ See generally Sanchirico, *supra* note 293.

³²¹ *Id.* at 1231–32.

³²² *Id.* at 1239–59. We need not delay with a critique of his critiques, except to say that the traditional explanations do not strike us to be as weak as Sanchirico asserts. Sanchirico, like most if not all 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.

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. 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.³²³

The essential argument of Sanchirico's theory about the character evidence ban is this: character evidence has predictive and therefore probative value.³²⁴ But it has no incentive value—its presence or absence creates no incentive to refrain from proscribed acts.³²⁵ "Trace evidence,"³²⁶ 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.³²⁷ 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.³²⁸ Without a character evidence ban, a person with a "bad" character is in a "damned if you do and damned if you

³²³ *Id.* at 1231–32.

³²⁴ 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.

³²⁵ *Id.* at 1260–63.

³²⁶ Evidence left by the commission of an act, including the memories of eyewitnesses.

³²⁷ Sanchirico, *supra* note 293, at 1262.

³²⁸ *Id.* at 1266.

don't" situation.³²⁹ If he refrains from the proscribed 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 (because 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.³³⁰

The posited effect is plausible, but not intuitively compelling. It is possible, for example, that admission of character evidence would have the opposite effect. Persons tainted by provable prior offenses would realize that their chance of conviction would be high if they were arrested, and make a special effort to avoid situations that might lead to arrest. Sexual predators would not share beds with children; assaultive personalities would avoid dangerous barrooms. That effect could cause prior offenders to avoid situations and associations that might tempt them to crime, thus deterring it. To our intuitions, it seems unlikely that the rule prohibiting use of character evidence against the accused (but allowing use of bad acts for other purposes) has any substantial effect on the conduct of potential violators. In the exceptional situations where potential perpetrators do take the rule into account, the question whether it deters or promotes crime is highly uncertain and speculative.

Even though our intuitions differ from his, Sanchirico's theory is plausible in accounting for the existing law. What is less plausible is that judges, facing an issue of whether evidence should be admitted at trial, presented with arguments from counsel about accuracy versus fairness, and despite writing opinions that consider the problem in those terms, are nevertheless solving the dilemma with entirely different goals in mind, and that the most influential (if not most) judges are collectively engaged in the unexpressed enterprise of talking about in-court accuracy versus fairness but adopting rules calculated for out-of-court deterrence. It's not impossible, but surely improbable.

Sanchirico uses this basic incentivist notion to explain the wisdom of an array of related rules: the admissibility of character evidence for impeachment, at sentencing, and for punitive damages; exceptions to the character evidence rule that make sense from a

³²⁹ *Id.*

³³⁰ *Id.*

deterrence perspective; and exceptions to the character evidence rule that do not make sense from a deterrence perspective.³³¹

The article illustrates an important value of the "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. He was led to look away from seeing the rule as part of a backward-looking search for truth to seeing it instead as a forward-looking tool of social control.

If Sanchirico were arguing that evidence rules are primarily aimed at truth seeking or fair fighting in trials, but that they inevitably have evidence-generating or evidence-suppressing effects in the world outside of trials, his argument would be both hard to disagree with and less interesting. We refer to situations where potential defendants in sexual harassment suits act to ensure that their conduct with potential accusers takes place in the presence of witnesses; document retention (and destruction) policies at corporations; and careful management of hazardous waste in anticipation of the burdens of proof in environmental lawsuits. Or, if he argued that specific rules evolved to provide incentives for certain relationships (notably, the privilege rules, which promote candid communication among clients and patients and spouses) and behavior (notably, the exclusion of evidence of repairs following accidents so that lawsuits do not become disincentives to accident prevention), he would, again, be making an argument that would be both hard to disagree with and nothing new. But Sanchirico goes much further.

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."³³² That is a bold departure from (and challenge to) the heretofore nearly unanimous evidence scholarship of the past several hundred years. Whether it can be shown to be true remains to be seen. But who cannot be excited by the debates promised by so grand a claim, however those debates might turn out?

³³¹ The most prominent example is the relatively recent exception admitting evidence of an accused's propensity to engage in proscribed sexual conduct. FED. R. EVID. 412. According to Sanchirico's economic analysis, these rules will reduce the disincentive to commit such crimes. *Id.* at 1301.

³³² Sanchirico, *supra* note 293, at 1259.

While much of Posner's analysis and the central core of San-chirico's work look at evidence law less as a set of rules designed to improve truth finding and more as devices concerned with regulating behavior outside of trial by dispensing contingent incentives and disincentives, Alex Stein and his colleagues bring economic analysis back to evidence scholarship's traditional concern with the effect of the rules on decision making.

In a recent article, Stein succeeds in combining traditional evidence scholarship's core concern of accurate factfinding with a "consequentialist game-theoretic perspective."³³³ 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 Bentham³³⁴ 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.³³⁵ 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.³³⁶

Seidmann and Stein offer an economic argument supporting 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;³³⁷ 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 to

³³³ Daniel J. Seidmann & Alex Stein, *The Right to Silence Helps the Innocent: A Game-Theoretic Analysis of the Fifth Amendment Privilege*, 114 HARV. L. REV. 430, 431 (2000).

³³⁴ In fairness to Bentham, his views were at least sometimes compatible with the best evidence principle in regard to rules excluding evidence. See Dale A. Nance, *The Best Evidence Principle*, 73 IOWA L. REV. 227, 274-75 (1988).

³³⁵ Seidmann & Stein, *supra* note 333, at 451-52.

³³⁶ *Id.* at 452-55.

³³⁷ *Id.* at 455-58.

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.³³⁸ Consequently, innocent and guilty suspects and defendants become more distinguishable.³³⁹ Thus, the right to silence helps the innocent to be found not 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 works only when the standard of proof is "beyond a reasonable doubt" (rather than some lower threshold).³⁴⁰ Indeed, the authors argue that a reduction in the standard of proof would lead not only to more erroneous convictions, "but also [to] serious indeterminacy in suspect identification and selection."³⁴¹ 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.³⁴²

The game theory analysis consists largely of thinking through the strategies of innocent and guilty suspects and defendants under varying conditions of evidence, standards of proof, and several other variables.³⁴³ These are plausible, reasonable arguments about what suspects and defendants and factfinders would do. But they rarely are informed by empirical data about such behavior. They could be in-

³³⁸ *Id.* at 457.

³³⁹ *Id.*

³⁴⁰ Seidmann & Stein, *supra* note 333, at 470.

³⁴¹ *Id.* at 449.

³⁴² For examples, see *Mitchell v. United States*, 526 U.S. 314, 330 (1999) (applying the right to silence in sentencing proceedings); *Fletcher v. Weir*, 455 U.S. 603, 607 (1982) (permitting adverse inferences from pre-arrest and post-arrest silence for impeachment purposes if a defendant testifies); *Jenkins v. Anderson*, 447 U.S. 231, 238-40 (1980) (same); *Schmerber v. California*, 384 U.S. 757, 761 (1966) (holding that the right to silence protects only against compelled testimony, not against compelled production of physical evidence); *Griffin v. California*, 380 U.S. 609, 615 (1965) (prohibiting adverse inferences from refusal to testify).

³⁴³ Seidmann & Stein, *supra* note 333, at 466-74.

correct.³⁴⁴ The absence of empirical testing is, of course, typical of economic analysis of law.³⁴⁵

Seidmann and Stein do, however, offer a brief section that purports to test the implications of their theory against empirical data.³⁴⁶ 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 tenuous³⁴⁷ and presented in a way that is confusing if not contradictory.

³⁴⁴ For example, 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, *id.* at 450, 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 interrogation methods are designed to do)? Regarding the assumption that "a typical suspect confesses to a crime only when confronted with evidence that he believes to be irrefutable," *id.* at 450-51, to what extent is this true? Will guilty suspects 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 IND. L. REV. 925, 956-60 (2002). Is it plausible to believe simultaneously that factfinders 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 factfinders 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 Professor Robert Axelrod. ROBERT AXELROD, *THE EVOLUTION OF COOPERATION* (1984). Axelrod conducted a game theory contest in which entrants submitted computer 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 doubtless knew from empirical experiments on game theory by research psychologists (for an example, see ANATOL RAPOPORT, *EXPERIMENTAL GAMES AND THEIR USES IN PSYCHOLOGY* 19, 25-28 (1973)) that the most successful strategy for eliciting cooperation from an opponent is tit for tat, and who wrote a simple program to play that strategy. AXELROD, *supra*, at 31.

³⁴⁵ Posner has urged, in another context, that theoretical assumptions and hypotheses be empirically and rigorously tested with data. Richard A. Posner, *The Summary Jury Trial and Other Methods of Alternative Dispute Resolution: Some Cautionary Observations*, 53 U. CHI. L. REV. 366, 367 (1986).

³⁴⁶ Seidmann & Stein, *supra* note 333, at 498-502.

³⁴⁷ Seidmann and Stein themselves declare half of the data to be useless—in their words, "too contaminated with measurement error (including inconsistent classification schemes) to draw any meaningful conclusions." *Id.* at 500.

From the data mish-mash they conclude that the two predictions they derive from their model are confirmed.³⁴⁸

Perhaps the most telling finding (the import of which is not noted by Seidmann and Stein) is that very few suspects refuse to answer police questions. Siedmann and Stein cite two American studies that found that only 9.5% and 20.9% of suspects invoked their right to silence.³⁴⁹ British studies found about 10%, and a law allowing prosecutors to argue adverse inferences reduced that number by only a few percent.³⁵⁰ 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 that falsely tend to exonerate. Does this not create 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 rather far off the mark.³⁵¹

³⁴⁸ *Id.* at 499, 502.

³⁴⁹ *Id.* at 448 n.60.

³⁵⁰ *Id.* at 501.

³⁵¹ Which 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. 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. 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 ability to produce determinate predictions when applied to human actions and decisions. There is no good reason to believe that uniformly ob-

Other analyses by Stein take the view that evidence law is less a method of maximizing the chances of finding the truth in a disputed matter than a means of apportioning the risk of error when decisions are made under uncertainty. In his recent book, *Foundations of Evidence Law*, Stein proposes a theory built on economic analysis combined with probability theory, epistemology, and moral philosophy.³⁵² In Stein's analysis, evidence rules exist principally to ensure the just allocation of errors, not to prevent errors by finding truth.³⁵³ By taking the view that evidence law seeks the just allocation of the risk of error, Stein seeks to resolve a number of paradoxes that seem to result from a purely truth-seeking perspective.

Stein proposes a number of principles that both explain and justify the rules found in evidence law, and these principles work to serve the error allocation function. The "principle of maximal individualization" prevents factfinders from making a decision against a party when the evidence they have to work with is not subject to individualized testing.³⁵⁴ This principle is said to have broad application, although, like most principles, it can be trumped by other values. Three

served 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.

Seidmann & Stein, *supra* note 333, at 436-37 (citations omitted). Of course this is no different from most legal scholarship, and it would be unfair to lay any special criticism on the doorstep of Seidmann and Stein.

³⁵² See generally ALEX STEIN, *FOUNDATIONS OF EVIDENCE LAW* (2005). Stein, not surprisingly, has developed these ideas over time. See generally Alex Stein, *Of Two Wrongs That Make a Right: Two Paradoxes of the Evidence Law and Their Combined Economic Justification*, 79 TEX. L. REV. 1199 (2001); Alex Stein, *The Refoundation of Evidence Law*, 9 CANADIAN J. L. & JURISPRUDENCE 279 (1996).

³⁵³ STEIN, *supra* note 352, at 118-40.

³⁵⁴ *Id.* at 100. Readers will recognize in this a justification for holding "naked statistical evidence" to be inadmissible or insufficient. See *supra* notes 178-179 and accompanying text.

other principles guide the construction of rules that allocate the risk of error. The "cost-efficiency principle," which applies in all litigation, requires that factfinders minimize the total cost of errors and error-avoidance.³⁵⁵ The "equality principle" applies in civil litigation and posits that fact-finding procedures and decisions must not produce unequal apportionment of the risk of error between parties.³⁵⁶ The "equal best principle," which applies in criminal trials, requires that to justifiably convict a defendant the state must make its best efforts to protect the defendant from the risk of erroneous conviction and must not provide better protection to other individuals.³⁵⁷

Although Stein presents these ideas as a break from evidence law and the ostensible goals of evidence doctrine, it seems to us that his ideas are more a complement than an alternative to most existing theory. Rules that allocate the risk of making erroneous decisions are closer cousins of rules designed to enhance truth finding than Stein realizes. Indeed, it is hard to imagine any sophisticated system for seeking correct answers that is not at the same time concerned with allocation of the risk of error. Perhaps the best and most obvious analogy is to hypothesis testing in science: the rules for determining whether to reject a null hypothesis (in the pursuit of empirical truth) take fully into account the risk of erroneous rejection of true null hypotheses and erroneous failure to reject false null hypotheses. The risks of the two types of error are balanced in a manner that reflects the costs and harms associated with one type versus the other.³⁵⁸ In science, and, we believe, in trials, fact finding under conditions of uncertainty and the allocation of the risk of error work happily together hand in glove.

What we've seen so far in the law and economics contributions to evidence scholarship is usually a strong version of rational choice theory. But since evidence scholars are usually alert to the nonrational limitations of human reasoning, why not explore a weaker version of rational choice? Why not something like the bounded rationality of behavioral law and economics?³⁵⁹ Perhaps that will be the next turn.

³⁵⁵ STEIN, *supra* note 352, at 136–37.

³⁵⁶ *Id.* at 217–18.

³⁵⁷ *Id.* at 172–78.

³⁵⁸ See David H. Kaye & David A. Freedman, *Statistical Proof*, in 1 MODERN SCIENTIFIC EVIDENCE, *supra* note 102, at ch. 5. One can alternatively refer to any textbook on inferential statistics in any field that uses statistics.

³⁵⁹ See Russell B. Korobkin & Thomas S. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*, 88 CAL. L. REV. 1051, 1055–59 (2000).

CONCLUSION

In large part, the factors that influence the mix of doctrinal and interdisciplinary scholarship on evidence are similar to those that affect that mix in other fields. As in other legal fields, a variety of obstacles to interdisciplinary work exists. Consider first the category of quantitative empirical research, which in the case of evidence law often means research on law and psychology. Law professors have no training in empirical research and analysis. Even if they did, they do not have students who are in a position to assist in carrying out such research. Even if those problems can be surmounted, there are daunting funding problems, aggravated by the fact that the federal government has not seen funding of research for the improvement of law as a high priority.³⁶⁰ The traditions of single authorship and of grand theory scholarship also militate against empirical work.³⁶¹ Empirical research does not suit the scholarly habits of many law professors who rarely venture beyond the law library, and now, with so many resources online, need not even leave their offices.³⁶²

Changes in law schools can help reduce these obstacles.³⁶³ One of the most notable developments is the seemingly steady increase in the number of scholars joining the ranks of the legal academy who are educated in disciplines in addition to law, as well as having training in empirical research. For others, law schools could do more to support training of law professors in the needed methodological skills, encouraging them to attend research institutes or take needed courses. Changes in the way manuscripts are reviewed by law journals, for ex-

³⁶⁰ Federal funding agencies such as the State Justice Institute and the National Institute of Justice seldom have the funding of research on evidence law on their agendas. The National Science Foundation, however, has been more supportive of such work.

³⁶¹ 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).

³⁶² See generally Peter H. Schuck, *Why Don't Law Professors Do More Empirical Research?*, 39 J. LEGAL EDUC. 323 (1989).

³⁶³ For suggestions along the lines described in this paragraph, see generally Lee Epstein & Gary King, *The Rules of Inference*, 69 U. CHI. L. REV. 1 (2002).

ample by adding an element of blind peer review by faculty members, might encourage careful empirical research. The tradition of single authorship needs to be changed, encouraging legal scholars to collaborate with colleagues in other departments. In the field of evidence, collaboration is eased by the fact that the evidence scholar's concern with the accuracy of witnesses and factfinders overlaps with the psychologist's interest in memory, perception, and human reasoning.

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 an 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."³⁶⁴ Whether the issue of generalizability is greater here than in other kinds of 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. It is difficult to see why laboratory studies of sequential and simultaneous lineups would not be generalizable to legal situations. One also can do simulation studies about how, for example, people reason about evidence, such as Thompson and Schumann's studies of the prosecutor's fallacy and the defense attorney's fallacy,³⁶⁵ or Koehler's experiments on how people reason about probability.³⁶⁶ 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 Diamond and 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.³⁶⁷ On the other hand, as our assessment of the hearsay studies indicates, sometimes it is very difficult to mirror

³⁶⁴ Lempert, *supra* note 2, at 1709 (citation omitted).

³⁶⁵ See generally Thompson & Schumann, *supra* note 221.

³⁶⁶ See, e.g., Koehler & Shavero, *supra* note 164; Jonathan J. Koehler, *When Are People Persuaded by DNA Match Statistics?* 15 LAW & HUM. BEHAV. 493 (2001).

³⁶⁷ Shari Seidman Diamond & Neil Vidmar, *Jury Room Ruminations on Forbidden Topics*, 87 VA. L. REV. 1857, 1888-89 (2001).

the legal situation in laboratory studies. The use of qualitative information and fireside inductions always will be important in assessing evidence law. It is dangerous to generalize about generalizability; the devil is in the details, or in the particular subject being studied.

Interdisciplinary scholarship on forensic science is a particularly promising (if difficult) field. Work on forensic science topics is more challenging than much of the work on law and psychology, because of the absence of overlap of legally relevant issues with issues that are already being studied in other academic disciplines. In many areas, forensic science expertise has developed without input from the academy or attention to the scientific method.³⁶⁸ This creates obvious obstacles to research, but also an unmatched opportunity to contribute by doing something that is not being done elsewhere. Law professors should become involved, and try to interest experts in other fields as well. We hope that the development of institutes and centers will start to address this problem.³⁶⁹ If courts heed *Daubert's* call to screen expertise for scientific validity, the threat of exclusion of evidence should create an incentive for this type of work.

Scholarship on probability and proof, and on formal aids to drawing accurate inferences, will continue to be an important component of 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, one can 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. At the same time, we are unlikely to see Bayes' Theorem in trials themselves, where many items of evidence are involved and an expert would have to explain to the factfinder how to apply the theory.

³⁶⁸ See *supra* notes 146–163 and accompanying text.

³⁶⁹ For example, one development in progress is the creation of an "Institute for Studies in Science and Law," which would bring together ideas for needed research, researchers who could carry out the research, and funding for the research. See Inst. for Studies in Sci. and Law, Purpose Statement (Sept. 11, 2005) (on file with the authors).

Interest in evidence and feminism 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 moving beyond those limited areas will be useful to both the scholars and to the law of evidence. 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 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.

Law and economics scholarship on evidence will continue to be a presence in leading law reviews. It is possible that evidence law will not attract as much funding as in other areas of law and economics legal scholarship because it is not as obvious a fit with the political agenda of funders,³⁷⁰ but theoretical law and economics scholarship does not require any more funding than doctrinal scholarship, and elite journals are fond of it. Moreover, it need not be as time-consuming as empirical research, since the theorist can take a body of ideas and apply them to one legal topic after another.

Although we welcome interdisciplinary scholarship in general, we also see some unsettling elements. First, it sometimes exalts political and substantive concerns over the goal of accuracy. Second, it sometimes seeks or postulates implausible extrinsic effects.

The first of these concerns applies to feminist scholarship that, for example, argues for admission of battered woman's syndrome expert testimony. Some of the writing in the field seems to regard the results

³⁷⁰ Expressing this view, Richard Lempert writes:

[M]ajor funders of law and economics seem to have a pro-business, anti-regulation, 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.

Lempert, *supra* note 2, at 1637-38 (citations omitted).

that this admission would bring—favoring women defendants in murder cases, for example—as trumping concerns about the accuracy of the expertise admitted. And, to the extent that feminist scholars are influenced by post-modern epistemology, this might cause them to be fact-skeptics or to privilege anecdotal evidence—narratives—over more thorough, complete, critical, and systematic (in a word, scientific) evidence.

There is nothing illogical about preferring a substantive agenda to a procedural agenda. In the field of evidence, however, this posture is likely to be dangerous and self-defeating. First, there is no way to ensure that a politicized approach can be confined to feminist issues.³⁷¹ Taking such an approach would require either openly subordinating the goal of accuracy or secretly implementing another goal while adhering to the rhetoric of pursuing the truth. Both approaches have disadvantages. It is unlikely that establishing a practice of secretly distorting evidence law in pursuit of substantive agendas will work in the long run to protect those whom society otherwise victimizes; after all, judges come from the dominant group and share its prejudices and interests. And, of course, there is a great danger of getting caught. The open pursuit of substantive goals also invites imitation in other areas, and sacrifices one important way in which evidence law can promote equality. An avowed and honest pursuit of accuracy has the advantage of cutting both ways, helping the powerful in one case and the disadvantaged in another.³⁷²

Our concern about pursuing substantive effects also applies to law and economics scholarship, with the added twist that some of the substantive effects that have been posited seem implausible. With its strong interest in incentives, law and economics scholarship may be particu-

³⁷¹ For a reluctant endorsement of the idea of recognizing that the acceptance of BWS is frankly political, accompanied by expression of wishes that this approach can be cabined, see Robert P. Mosteller, *Syndromes and Politics in Criminal Trials and Evidence Law*, 46 DUKE L.J. 461, 509–16 (1996). Professor Mosteller's reservations about his own proposal are a good summary of our reasons for disagreeing with him. *Id.*

³⁷² For a prominent example: *Daubert* as doctrine advances the interests of rich civil defendants as well as indigent criminal defendants in avoiding having junky science used against them in court. In applying *Daubert*, judges may not be so evenhanded but instead scrutinize the offerings of civil plaintiffs more studiously than they scrutinize the offerings of government prosecutors; such judges are subject to serious criticism for their willful disregard of the law. See D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 ALB. L. REV. 99, 104–12 (2000). Also, compare the reviews of case law relating to expert evidence usually associated with civil cases to that usually associated with criminal cases, in the various chapters of MODERN SCIENTIFIC EVIDENCE, *supra* note 102.

larly vulnerable to this flaw. When your tool is a hammer, everything looks like a nail. As long as law and economics scholarship looks for the ways in which evidence rules shape everyday societal behavior, and disregards the truth-finding aspects of evidence rules, it risks overlooking core concerns in favor of chasing epiphenomena. While we do not agree that evidence-related law and economics is necessarily "common sense on stilts,"³⁷³ we do think that evidence scholars already are quite adept at imagining the possible effects of evidence rules without any training in economics. What is needed is empirical validation, not increasingly ingenious and speculative hypotheses about what the extrinsic effects of evidence rules might be. There are plenty of good hypotheses that are in need of testing, and inventing clever new ones that no one has thought of before should, all else equal, have lower priority than trying to validate plausible old ones.

We recognize that one cannot predict with assurance the future path of law and economics evidence scholarship based on what has appeared so far, and we have no desire to "strangle the infant in its crib."³⁷⁴ Moreover, we recognize that, in other fields, predictions that law and economics scholarship would level off or die out have proven to be wrong, or at least premature.³⁷⁵ Nonetheless, we think that the scholarship so far too often has shown a proclivity toward having fun by tracing out the implications of models, regardless of what other sources of knowledge might say about the plausibility of the models or the deduced effects.

Interdisciplinary scholarship that has the objective of improving fact finding is an obvious boon for evidence scholars. In this category we place scholarship that deals with probability and proof, with human psychology and human reasoning, with the scientific method and forensic science. Interdisciplinary scholarship that pursues other objectives is a less obvious match. To the extent that feminist scholarship pursues substantive objectives, it would do better to direct those efforts at changing the substantive law, rather than at addressing problems indirectly by changing (and perhaps distorting) evidence

³⁷³ Lempert, *supra* note 2, at 1619. Also see Posner's response, *supra* note 2.

³⁷⁴ See Posner, *supra* note 2, at 1721.

³⁷⁵ See *id.* at 1714 (citing Owen M. Fiss, *The Law Regained*, 74 CORNELL L. REV. 245, 245 (1989) ("[L]aw and economics . . . seems to have peaked."); Morton J. Horwitz, *Law and Economics: Science or Politics?*, 8 HOFSTRA L. REV. 905, 905 (1980) ("I have the strong feeling that the economic analysis of law has 'peaked out' as the latest fad in legal scholarship.")).

law.³⁷⁶ To the extent that law and economics scholarship posits a strong incentive effect of rules that purport merely to seek accuracy, few will find such arguments plausible. To the extent that it seeks this effect through the reform of evidence law, we worry that its influence will be harmful, leading evidence law away from the important mission of producing accurate verdicts.

Whether accuracy should be the primary objective of evidence law is a topic that could support an article twice as long as this one. Other general goals, such as satisfaction of the parties or catharsis, are at least plausible, and in particular contexts goals such as the protection of privacy or encouraging beneficial out-of-court conduct are paramount. But, as a general matter, it is hard to understand how a society can follow the rule of law in the absence of accurate fact finding. In pursuit of these purposes, one needs to have in mind goals for evidence law and evidence scholarship. We argue that the main, though certainly not the only, goal for evidence law is to promote accuracy in fact finding. Accuracy is essential to accomplishing the goals of substantive law. For the substantive law to work, the fact-finding mechanism must be accurate enough to enforce its prohibitions and dispense its rewards. While one should not be unduly optimistic about the goal of accuracy, one should also avoid extreme cynicism. There is no good reason to believe that the goal of accuracy ultimately will be undermined by the self-serving conduct of actors in the system. After all, those powerful enough to create the substantive law are motivated to make sure that it is enforced effectively, so accuracy in fact finding should in general have a powerful constituency. Any goal of accomplishing something else (such as pretending to be accurate while not actually being so) is likely to be self-defeating because the deception inevitably will be discovered.³⁷⁷

It is a heady task to make prescriptive statements about the direction of evidence scholarship, because there is so little general agreement about what constitutes good scholarship. The dispute about the use of fiction in scholarship in other fields is only one example of the

³⁷⁶ If the goal is to provide an escape from criminal liability for women who kill their abusers, the better way to accomplish that is to change the substantive law of self defense to insulate such persons from liability in appropriately defined circumstances, not to squeeze questionable expertise in through the evidence rules. Doing that weakens the evidence rules, invites backfire (as similarly questionable expertise is used against women), and might offer only temporary protection (as the experts change their own views over time). See DAVID L. FAIGMAN, *LEGAL ALCHEMY: THE USE AND MISUSE OF SCIENCE IN THE LAW* 74 (1999); Faigman, *supra* note 268, at 643-47.

³⁷⁷ See generally Nesson, *supra* note 233.

divergence of views.³⁷⁸ There is little agreement upon basic premises, such as whether scholarship should be useful.³⁷⁹ Evidence scholarship has many purposes, but surely one worthwhile purpose is to improve the accuracy of verdicts. Traditional doctrinal scholarship aimed at this goal by improving evidence law, for example by eliminating anomalies and obstacles to rational proof. The "New Evidence Scholarship" on probability and proof represented a turn from reasoning about evidence law to reasoning about evidence facts. With William Twining's influential approval,³⁸⁰ scholars became interested in how formal methods might be used as an aid in drawing inferences and finding facts.³⁸¹ We think that it is even more exciting to contemplate scholarship—like that done on eyewitness memory and on forensic science—that aims at improving the evidence facts themselves by influencing the way in which evidence is generated.

³⁷⁸ See DANIEL A. FARBER & SUZANNA SHERRY, *BEYOND ALL REASON: THE RADICAL ASSAULT ON TRUTH IN AMERICAN LAW* 95–117 (1997).

³⁷⁹ "[T]he scholar seeks knowledge for its own sake, not for some further purpose, although the knowledge he acquires may be instrumentally useful for other ends." Anthony T. Kronman, *Foreword: Legal Scholarship and Moral Education*, 90 *YALE L.J.* 955, 967–68 (1981).

³⁸⁰ See TWINING, *RETHINKING EVIDENCE: EXPLORATORY ESSAYS*, *supra* note 1, at 12–31, 341–72.

³⁸¹ See generally C.G.G. AITKEN, *STATISTICS AND THE EVALUATION OF EVIDENCE FOR FORENSIC SCIENTISTS* (1995); ROBERTSON & VIGNAUX, *supra* note 206; DAVID A. SCHUM, *THE EVIDENTIAL FOUNDATIONS OF PROBABILISTIC REASONING* (1994).