5-16-2012

The Nature and Purpose of Evidence Theory

Michael S. Pardo
University of Alabama - School of Law, mpardo@law.ua.edu

The Nature and Purpose of Evidence Theory

*Michael S. Pardo**

* Henry Upson Sims Professor of Law, University of Alabama School of Law. I was fortunate to have the opportunity to present some of the ideas in this Article in a number of different venues. These include the Association of American Law Schools ("AALS") Section on Evidence Panel on "The Future of Statistics in Evidence Law"; the conference on "Standards of Proof and Scientific Evidence," held at the University of Girona, Spain; the conference on "Legal Epistemology" held at the Institute for Philosophical Research, National Autonomous University of Mexico; the Program on Understanding Law, Science, and Evidence ("PULSE") Workshop at UCLA Law School; and faculty workshops at Cardozo, Fordham, Michigan State, San Diego, Vanderbilt, and William & Mary. My thanks to the participants at each of these events for their helpful questions and comments. I was also fortunate to receive many helpful comments on previous drafts—my thanks to Larry Alexander, Ron Allen, Amalia Amaya, Craig Callen, Ed Cheng, Shahar Dillbary, James Dwyer, Jordi Ferrer Beltrán, Michael Steven Green, Susan Haack, Jerry Kang, Jay Koehler, Larry Laudan, Youngjae Lee, Ethan Leib, Erik Lillquist, Jenny McEwan, Jennifer Mnookin, Erica Beecher Monas, Michael Risinger, Paul Roberts, David Schwartz, Jason Solomon, Lawrence Solum, Alex Stein, Bill Thompson, Peter Tillers, Fred Vars, and Ekow Yankah. My thanks also to Dean Ken Randall and the Alabama Law School Foundation for generous research support. Finally, my thanks to the editors of the VANDERBILT LAW REVIEW, in particular Deanna Foster, for their excellent work. This Article is dedicated to the memory of Craig Callen.

INTRODUCTION

Approximately twenty-five years ago, Professor Richard Lempert, reflecting on the then-current state of evidence scholarship, noted a dramatic shift underway.[1] He described what had become a largely "moribund" field giving way to a burgeoning "new evidence scholarship."[2] The scholarship in the moribund phase employed "a timid kind of deconstructionism with no overarching critical theory," was "seldom interesting," and any "potential utility" was "rarely realized"; Lempert proposed the following mock article title as a model representing the genre: "What's Wrong with the Twenty-Ninth Exception to the Hearsay Rule and How the Addition of Three Words Can Correct the Problem."[3] By contrast, the "new evidence scholarship" was moving from merely interpreting rules to "analyzing the process of proof" and drawing insights from "mathematics,

1. Richard Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof*, 66 B.U. L. REV. 439, 439–77 (1986).

2. The "moribund" nature of the field developed after the "great systematizers of the common law" (Wigmore, Maguire, McCormick, and Morgan). *Id.* at 439. In describing the then-current work as the "new evidence scholarship," Lempert also coined the phrase that would come to define this collective body of work. *See* Roger C. Park & Michael J. Saks, *Evidence Scholarship Reconsidered: Results of the Interdisciplinary Turn*, 47 B.C. L. REV. 949, 984–85 (2006) ("The term 'New Evidence Scholarship' coined by Richard Lempert 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 scholarship that applies formal tools of probability theory, such as Bayes' Theorem."). The new evidence scholarship revived theoretical projects anticipated by the late greats Bentham, Thayer, and Wigmore in exploring the philosophical foundations of evidence law. *See generally* JEREMY BENTHAM, AN INTRODUCTORY VIEW OF THE RATIONALE OF JUDICIAL EVIDENCE, WORKS VI, at 7–20 (John Bowring ed., 2002) (1843); JAMES BRADLEY THAYER, A PRELIMINARY TREATISE ON EVIDENCE AT THE COMMON LAW vi–viii (1898); WILLIAM TWINING, THEORIES OF EVIDENCE: BENTHAM AND WIGMORE 1–4 (1985) (explaining how Bentham and others laid the foundation for evidence scholarship); JOHN H. WIGMORE, THE PRINCIPLES OF JUDICIAL PROOF 1–4 (1913).

3. Lempert, *supra* note 1, at 439; *see also id.* at n.1 ("This title not only suggests the focus of the type of scholarship I am thinking about but also typifies it in length and lack of grace.").

psychology and philosophy."[4] To be clear, the new evidence scholarship still focused to a large extent on *rules*, but it provided more robust analysis in light of the legal proof process and its underlying goals—most significantly, the goals of fostering accurate outcomes, avoiding factual errors, and allocating the risk of error in a fair and justified manner.

The "new evidence scholarship" is no longer new. The burgeoning field described in Lempert's classic article produced an explosion of diverse theoretical and empirical work on the evidentiary proof process.[5] On the theoretical side, many scholars continue to utilize "mathematical models as modes of proof or as a means of understanding trial processes."[6] Here is an illustrative sample of issues on which current scholarship focuses:

(1) the desirability of basing legal judgments on explicitly statistical evidence;[7]

---

4. *Id.* at 439–40.

5. The symposium in which Lempert's article appeared marks an important phase in the literature, 66 B.U. L. REV. 377–952 (1986) ("Symposium: Probability and Inference in the Law of Evidence"), as does a subsequent symposium in 13 CARDOZO L. REV. 253–1075 (1991) ("Symposium: Decision and Inference in Litigation"). Earlier related articles analyzing the process of proof include: Vaughn Ball, *The Moment of Truth: Probability Theory and Standards of Proof*, 14 VAND. L. REV. 807, 807–08 (1961); Alan D. Cullison, *Probability Analysis of Judicial Fact-Finding: A Preliminary Outline of the Subjective Approach*, 1 U. TOL. L. REV. 538, 538–40 (1969) (examining probability analysis in the context of proof); Michael O. Finkelstein & William B. Fairley, *A Bayesian Approach to Identification Evidence*, 83 HARV. L. REV. 489, 490 (1970) (same); John Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065, 1065–70 (1968); Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357, 1357–60 (1985); Lawrence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1329–31 (1971) (analyzing proof in the context of mathematics). In addition to the projects by Bentham, Thayer, and Wigmore, *supra* note 2, another early precursor is J.R. GULSON, THE PHILOSOPHY OF PROOF: IN ITS RELATION TO THE ENGLISH LAW OF JUDICIAL PROOF v–ix (1905).

6. Lempert, *supra* note 1, at 440. For a brief overview of the literature, see D. Michael Risinger, *Introduction*: *Bayes Wars Redivivus—An Exchange*, 8 INT'L COMMENT. ON EVIDENCE 1, 1–41 (2010).

7. *Compare* Tribe, *supra* note 5, at 1350 (noting that the mere fact that mathematics "taken alone can rarely, if ever, establish [guilt] . . . does not imply that such evidence—*when properly combined with other, more conventional evidence in the same case*—cannot supply a useful link in the process of proof" (emphasis in original)), *with* Finkelstein & Fairley, *supra* note 5, at 490, 517 (arguing that mathematics should never be used alone to determine guilt or innocence, but that Bayes' Theorem used in conjunction with other incriminating evidence can be used to "lead to a fairer evaluation of identification evidence"). This issue has taken on more significance with the possibility of "cold hit" DNA cases, in which the "random match" probability of a DNA match (i.e., the likelihood that a DNA sample came from another random member of the population) is the only incriminating evidence against a defendant. *See* David H. Kaye, *Rounding up the Usual Suspects: A Legal and Logical Analysis of DNA Trawling Cases*, 87 N.C. L. REV. 425, 425–27 (2009) (noting that the issue of "DNA trawling" has become a more "urgent" question with the development of DNA databases containing "more than six million profiles from

(2) the feasibility of introducing statistical evidence or formal techniques into the proof process;[8]

(3) defining fundamental concepts such as relevance and probative value in probabilistic terms;[9]

(4) quantifying the probative value of particular items of evidence;[10]

(5) explaining, justifying, or critiquing particular rules of evidence;[11]

(6) explaining, justifying, or critiquing standards of proof;[12] and

---

convicted offenders and suspects"); Andrea Roth, *Safety in Numbers? Deciding When DNA Alone is Enough to Convict*, 85 N.Y.U. L. REV. 1130, 1130–35 (2010) (arguing that courts should adopt a numerical threshold in "cold hit" cases); Montgomery Slatkin, Yun S. Song & Erin Murphy, *The Probability that a "Cold Hit" in a DNA Database Search Results in an Erroneous Attribution*, 54 J. FORENSIC SCI. 22, 22–25 (2009).

8. *See* Craig R. Callen, *Notes on a Grand Illusion: Some Limits on the Use of Bayesian Theory in Evidence Law*, 57 IND. L.J. 1, 10–24 (1982) (discussing problems of computational complexity); 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, 395–98 (2005) (arguing that Bayesian presentations may improve juror understanding of DNA evidence).

9. *See* David H. Kaye, *The Relevance of "Matching" DNA: Is the Window Half Open or Half Shut?*, 85 J. CRIM. L. & CRIMINOLOGY 676, 683–84 (1995) ("The best developed and most plausible theory of probative value builds on a statistical concept known as the likelihood ratio."); Richard O. Lempert, *Modeling Relevance*, 75 MICH. L. REV. 1021, 1025 (1977) ("Where the likelihood ratio for an item of evidence differs from one, that evidence is *logically relevant*. This is the mathematical equivalent of . . . 'relevant evidence' . . . .").

10. *See generally* Ronald J. Allen & Michael S. Pardo, *The Problematic Value of Mathematical Models of Evidence*, 36 J. LEGAL STUD. 107, 108–09 (2007) (surveying attempts to quantify probative value); *supra* note 9.

11. *See* Richard D. Friedman, *Character Impeachment Evidence: Psycho-Bayesian Analysis and a Proposed Overhaul*, 38 UCLA L. REV. 637, 655–66 (1991) (critiquing character rules based on probabilistic analysis); Richard D. Friedman, *Route Analysis of Credibility and Hearsay*, 96 YALE L.J. 667, 690–729 (1987) (analyzing hearsay in probabilistic terms).

12. *See* Richard S. Bell, *Decision Theory and Due Process: A Critique of the Supreme Court's Lawmaking for Burdens of Proof*, 78 J. CRIM. L. & CRIMINOLOGY 557, 557–59 (1987) (arguing that "the Court's reasoning about both [elements underlying its logic on burden of proof jurisprudence] has been deficient"); Edward K. Cheng, *Reconceptualizing the Burden of Proof*, YALE L.J. (forthcoming), *available at* http://ssrn.com/abstract=2087254 (reconceptualizing the "preponderance standard as a probability ratio"); Cullison, *supra* note 5; Michael L. DeKay, *The Difference Between Blackstone-Like Error Ratios and Probabilistic Standards of Proof*, 21 LAW & SOC. INQUIRY 125, 125–26 (1996); David Hamer, *Probabilistic Standards of Proof, Their Complements and the Errors that Are Expected to Flow from Them*, 1 U. NEW ENG. L.J. 71, 72–74 (2007); Kaplan, *supra* note 5; D.H. Kaye, *Clarifying the Burden of Persuasion: What Bayesian Decision Rules Do and Do Not Do*, 3 INT'L J. EVIDENCE & PROOF 1, 1–4 (1999) (explaining Bayesian rules in the context of proof); Neil Orloff & Jery Stedinger, *A Framework for Evaluating the Preponderance-of-the-Evidence Standard*, 131 U. PA. L. REV. 1159 (1983); Fredrick E. Vars, *Toward a General Theory of Standards of Proof*, 60 CATH. U. L. REV. 1, 2–4 (2010).

(7) analyzing the various "proof paradoxes" (for example, "Blue Bus," "Gatecrashers," and "Prisoners in the Yard"), in which outcomes implied by a probabilistic conception of the proof process appear to clash with judgments about what the proper outcomes should be.[13]

This scholarship informs philosophical work exploring the epistemological and moral foundations of evidence law and legal proof,[14] as well as law-and-economics-inspired work examining the possible effects of evidence, evidentiary rules, and legal proof on primary (i.e., nonlitigation) behavior.[15]

---

13. For an overview of the literature, see Mike Redmayne, *Exploring the Proof Paradoxes*, 14 LEGAL THEORY 281 (2008). For a recent philosophical discussion, see David Enoch, Levi Spectre & Talia Fisher, *Statistical Evidence, Sensitivity, and the Legal Value of Knowledge*, 30 PHIL. & PUB. AFF. 197 (2012).

14. *See* ROBERT P. BURNS, A THEORY OF THE TRIAL 3–9 (2001); LARRY LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW: AN ESSAY IN LEGAL EPISTEMOLOGY 1–9 (2006) (arguing that, "whatever else it is, a criminal trial is first and foremost an *epistemic* engine"); ALEX STEIN, FOUNDATIONS OF EVIDENCE LAW 1–11 (2005); Ronald J. Allen & Brian Leiter, *Naturalized Epistemology and the Law of Evidence*, 87 VA. L. REV. 1491, 1491–93 (2001); Amalia Amaya, *Justification, Coherence, and Epistemic Responsibility in Fact-Finding*, 5 EPISTEME 306–09 (2008) (arguing for a greater role for epistemology in evidence law); Scott Brewer, *Scientific Expert Testimony and Intellectual Due Process*, 107 YALE L.J. 1535, 1538–42 (1998) (exploring the role of scientific experts in evidence law from an epistemological perspective); Alvin I. Goldman, *Quasi-Objective Bayesianism and Legal Evidence*, 42 JURIMETRICS J. 237, 237–40 (2002) (exploring the intersection of Bayesianism and evidence law); Michael S. Pardo, *The Field of Evidence and the Field of Knowledge*, 24 LAW & PHIL. 321, 321–24 (2005); Michael S. Pardo & Ronald J. Allen, *Juridical Proof and the Best Explanation*, 27 LAW & PHIL. 223, 224–25 (2008) (arguing that "the relative neglect of explanation-based reasoning," relative to a focus on probability theories, "has been a mistake"); Frederick Schauer, *In Defense of Rule-Based Evidence Law—And Epistemology Too*, 5 EPISTEME 295, 295–96 (2008); *see also* Paul Roberts, *Groundwork for a Jurisprudence of Criminal Procedure*, *in* PHILOSOPHICAL FOUNDATIONS OF CRIMINAL LAW 379, 387–94 (R.A. Duff & Stuart P. Green eds., 2011) (discussing "epistemological perspectives on evidence and proof" and concluding "this is one of the most vibrant and productive chapters of Evidence scholarship in recent years").

15. *See* Dan M. Kahan, *The Economics—Conventional, Behavioral, and Political—of "Subsequent Remedial Measure" Evidence*, 110 COLUM. L. REV. 1616, 1616–19 (2010) (criticizing the economic justifications for the "ban on proof of 'subsequent remedial measures' "); Louis Kaplow, *Burden of Proof*, 121 YALE L.J. 738, 741–52 (2012) (exploring the burden of proof from a social-welfare perspective); Gideon Parchomovsky & Alex Stein, *The Distortionary Effect of Evidence on Primary Behavior*, 124 HARV. L. REV. 518, 519–24 (2010) (arguing that "evidentiary motivations will often lead actors to engage in socially suboptimal behavior . . . to generate evidence that can later be presented in court and will increase their chances of winning the case"); Chris William Sanchirico, *A Primary-Activity Approach to Proof Burdens*, 37 J. LEGAL STUD. 273, 273–77 (2008) (exploring evidence law as a tool to influence behavior rather than simply find truth); Chris William Sanchirico, *Character Evidence and the Object of Trial*, 101 COLUM. L. REV. 1228, 1229–39 (2001) (arguing that character evidence rules can only be analyzed in a context where trials are just one part of a state's efforts to regulate the behavior of individuals); Robert E. Scott & George G. Triantis, *Anticipating Litigation in Contract Design*, 115 YALE L.J. 814, 816–22 (2006) (examining economic efficiency in contract law via the lens of standards of proof). The scholarship outlined above has also inspired nascent work on artificial intelligence ("AI") and law. *See* DOUGLAS WALTON, WITNESS TESTIMONY EVIDENCE:

On the empirical side, a robust literature provides a powerful psychological model of the behavior of legal fact finders in general and evidence on a host of specific evidentiary issues.[16] The well-confirmed model of jury behavior—the Story Model—posits that legal fact finders assimilate evidence into competing narratives of the events and select the most plausible or satisfying of the available accounts.[17] In addition to the general model, the literature presents a variety of claims about the behavior of juries and judges regarding specific types of evidence, instructions, scenarios, or influences,[18] and it sheds some light on the effects that rules regulating evidence such as hearsay, expert testimony, and prior convictions appear to have in actual cases.[19] The

---

ARGUMENTATION, ARTIFICIAL INTELLIGENCE, AND LAW 1–11 (2008) (surveying AI models of legal evidence).

16. For overviews of the literature on juries, see NEIL VIDMAR & VALERIE P. HANS, AMERICAN JURIES: THE VERDICT 15–21 (2007) (describing the effectiveness of the American jury system); Shari Seidman Diamond, *Beyond Fantasy and Nightmare: A Portrait of the Jury*, 54 BUFF. L. REV. 717, 717–22 (2006) (examining myths and realities of jury trials). On judges as fact finders, see Chris Guthrie, Jeffrey J. Rachlinski & Andrew J. Wistrich, *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 821 (2001) (exploring the decisionmaking processes judges employ); Andrew J. Wistrich, Chris Guthrie & Jeffrey J. Rachlinski, *Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding*, 153 U. PA. L. REV. 1251, 1330–31 (2005) (arguing that a judge's ability to disregard evidence deemed inadmissible is limited); *see also* Frederick Schauer, *On the Supposed Jury-Dependence Evidence Law*, 155 U. PA. L. REV. 165, 165–86 (2006) (discussing the need for evidentiary rules to guide and constrain judicial decisionmaking on factual issues).

17. *See* VIDMAR & HANS, *supra* note 16, at 135 ("Many subsequent studies . . . have lent support to the basic assumptions of the story model."); Nancy Pennington & Reid Hastie, *A Cognitive Model of Juror Decision Making: The Story Model*, 13 CARDOZO L. REV. 519, 519–20 (1991) (analyzing the Story Model). In presenting the Story Model, Pennington and Hastie contrast it with probability models and argue convincingly that it provides a better descriptive account of the cognitive processes of jurors. Pennington & Hastie, *supra*, at 519–20. A distinction between stories and probability theory continues to dominate scholarly discussions of evidence and trials; however, for the reasons discussed below, this distinction is problematic and based on a false theoretical dichotomy. *See infra* Part I.C.

18. *See* Jonathan J. Koehler, *The Psychology of Numbers in the Courtroom: How to Make DNA-Match Statistics Seem Impressive or Insufficient*, 74 S. CAL. L. REV. 1275, 1300 (2001) (detailing the behavior of attorneys and juries regarding DNA evidence); Michael J. Saks & Robert F. Kidd, *Human Information Processing and Adjudication: Trial by Heuristics*, 15 L. & SOC'Y REV. 123, 123–26 (1980-81) (discussing cognitive psychology and legal fact-finding); David Alan Sklansky, *Evidentiary Instructions and the Jury as Other*, 65 STAN. L. REV. (forthcoming 2013), *available at* http://ssrn.com/abstract=2029106 (discussing the empirical literature on jury instructions); *see also* VIDMAR & HANS, *supra* note 16 (surveying the literature).

19. *See* Edward K. Cheng & Albert H. Yoon, *Does* Frye *or* Daubert *Matter? A Study of Scientific Admissibility Standards*, 91 VA. L. REV. 471, 471–75 (2005) (arguing that the difference between the *Daubert* and *Frye* scientific admissibility tests is negligible in the course of a trial); Larry Laudan & Ronald J. Allen, *The Devastating Impact of Prior Crimes Evidence and Other Myths of the Criminal Justice Process*, 101 J. CRIM. L. & CRIMINOLOGY 493, 493–97 (2011) (asserting that "prior criminal convictions weigh heavily in jurors' decisions about acquittal and conviction" but that "jurors' learning . . . that defendant has been convicted of prior

relationships between the theoretical literature and the issues explored in the empirical literature raise additional theoretical, practical, and empirical questions.[20]

This Article takes up the theoretical project writ large. Exploring the landscape of evidence scholarship, the Article examines a number of methodological and metatheoretical questions: What would a successful evidentiary theory look like? By what criteria ought we assess such a theory? What is the purpose of such theorizing? What is the relationship between the theoretical and empirical projects? In exploring these questions, the Article identifies theoretical criteria by which *any* theory of the evidentiary proof process (or aspects of the process) may be evaluated.[21]

Although the discussion will be theoretical, its practical significance is real and far reaching. The skeptic might ask, "Why, from a practical perspective, do we need evidence theories in the first place?" We need them for three distinct reasons.[22] First, any application by judges of evidentiary concepts—and the rules and standards that contain these concepts—will presuppose *some* conception of what the concepts mean and what is required by the rules and standards.[23] For example, any determination or evaluation

---

crimes makes very little difference in conviction rates"); Eleanor Swift, *The Hearsay Rule at Work: Has it Been Abolished De Facto by Judicial Decision?*, 76 MINN. L. REV. 473, 473–77 (1992).

20. For example, what is the relationship between the Story Model and burdens and standards of proof? What normative implications follow from their interaction? *See* Kevin M. Clermont, *Standards of Proof Revisited*, 33 VT. L. REV. 469, 469–71 (2009) (discussing possible relationships between the Story Model and probabilistic standards of proof); Lisa Kern Griffin, *Narrative, Truth, and Trial*, 101 GEO. L.J. 281 (2013) (discussing potential ways to improve accuracy in criminal trials in light of the Story Model); Eleanor Swift, *Narrative Theory, FRE 803(3), and Criminal Defendants' Post-Crime State of Mind Hearsay*, 38 SETON HALL L. REV. 975, 975–76 (2008) (analyzing hearsay in light of the Story Model).

21. By "evidentiary proof process," I mean the legal procedures by which parties use evidence to prove or disprove material facts. These procedures regulate the process of proof at trial as well as at various pre- and posttrial proceedings (for example, at a pretrial hearing or at sentencing) in which parties present evidence and legal decisionmakers conclude whether a particular standard of proof has been met.

22. In addition to these three reasons, Peter Tillers explores a number of other ways in which evidentiary theories may contribute to our understanding of legal proof. Peter Tillers, *Trial by Mathematics—Reconsidered*, 10 L., PROBABILITY & RISK 167, 172 (2011).

23. *See* W.B. Gallie, *Essentially Contested Concepts*, 56 PROC. ARISTOTELIAN SOC'Y 167, 180–87 (1956) (discussing the concept-conception distinction). General concepts may be conceived of in different ways (i.e., different conceptions), and disagreements may emerge as to the best way to characterize the general concepts. For example, epistemologists offer different conceptions of the general concept of "knowledge," and moral, legal, and political philosophers offer different conceptions of the general concept of "justice." For a general overview, see Lawrence B. Solum, *Legal Theory Lexicon 028: Concepts and Conceptions*, LEGAL THEORY BLOG (Apr. 29, 2012), http://lsolum.typepad.com/legal_theory_lexicon/2004/03/legal_theory_le_1.html. Evidence theory

of whether an item of evidence is *relevant* (the evidentiary concept) will depend on some conception or understanding of what it means for evidence to be "relevant" in the first place. Without *some* conception, the concepts (and the rules and standards that employ the concepts) could not guide particular applications, nor could particular applications be evaluated in light of what is required by the rules and standards. Moreover, there may be more than one conception for many evidentiary concepts.

Second, the proof process provides a forum through which the law enforces the rights, duties, and obligations flowing from substantive areas of law. Applications in this process—including both rulings by judges on the admissibility of particular items of evidence and judgments by fact finders on whether the evidence as a whole proves a particular fact—may be principled, coherent, and justified, or they may be left to the subjective whims of individual decisionmakers. If the latter is the case, then is it not clear why we even have *law* on these issues at all.[24] Of course, even "untheorized" trial practices may ultimately turn out to be good ones, but evidence theory nevertheless makes explicit what is implicit in these practices, so that we can better examine, evaluate, critique, and perhaps improve them.

Third, the practical significance of evidence theory extends well beyond trials. The issues at stake affect doctrinal issues in criminal and civil litigation more generally in critical, and underappreciated, ways. The evidentiary rules and standards also determine important issues such as who gets to trial in the first place, which verdicts will be allowed to stand, and which convictions will be overturned. For example, summary judgment and judgment as a matter of law (civil cases) and challenges by defendants to sufficiency of the evidence (criminal cases) each turn on what fact finders could reasonably infer

---

also involves differing conceptions of key concepts such as "relevance," "probative value," and "evidential sufficiency." *See infra* note 116.

24. This is not to suggest that subjective decisions cannot be principled, coherent, justified, or correct. The key point is that purely subjective standards would cease to be *standards* in any meaningful sense, and the enforcement of substantive law under such a regime would depend on the subjective beliefs of decisionmakers and not on the principled application of generally applicable legal standards. In analyzing the disarray in English courts on standards of proof, Mike Redmayne expresses similar sentiments:

> Whatever the nature of the *process* of proof in legal fact finding, we need some basis for understanding why we have *standards* of proof and why we may wish to vary them in different adjudicative contexts. The price English law has paid for ignoring this is clear: a confused and inconsistent case law.

Mike Redmayne, *Standards of Proof in Civil Litigation*, 62 MOD. L. REV. 167, 194–95 (1999).

from the evidence in light of the applicable evidentiary standards.[25] In an era of "vanishing trials" on the civil side[26] and the attention to and concern with false convictions on the criminal side,[27] these practical doctrinal issues may be as important as any facing the law today. They depend, at root, on a satisfactory account of evidence and proof—in other words, on evidence theory.[28]

The Article proceeds in three main Parts. Part I provides general criteria for evaluating evidence theory, and Parts II and III apply the criteria to current theoretical accounts in evidence scholarship. The metatheoretical discussion in Part I contributes to the theoretical literature regardless of whether one accepts the specific applications and conclusions reached in Parts II and III. For this reason, it may be the most important Part of the Article.

Part I explores the nature of theoretical accounts of evidence and proof and the purposes of this theorizing. The discussion begins by isolating an "epistemological core" of procedural considerations at the

---

25. FED. R. CIV. P. 56 (summary judgment), FED. R. CIV. P. 50 (judgment as a matter of law); Reeves v. Sanderson Plumbing Prods., Inc., 530 U.S. 133, 149–50 (2000) (explaining that the standard for judgment as a matter of law "mirrors" the standard for summary judgment); Anderson v. Liberty Lobby, Inc., 477 U.S. 242, 248, 252 (1986) (explaining the summary judgment standard as whether "a reasonable jury could return a verdict for the nonmoving party" and that this determination depends on the "evidentiary standard of proof" at trial); Jackson v. Virginia, 443 U.S. 307, 318 (1979) (articulating sufficiency standard in criminal cases as whether "any rational trier of fact could find guilt beyond a reasonable doubt"). The evidentiary rules also set directions for discovery.

26. *See* Marc Galantar, *The Vanishing Trial: An Examination of Trials and Related Matters in Federal and State Courts*, 1 J. EMPIRICAL LEGAL STUD. 459, 460 (2004) (documenting decreases in rates of civil trials). Diminishing trials are also a concern on the criminal side. *See* Lawrence M. Friedman, *The Day Before Trials Vanished*, 1 J. EMPIRICAL LEGAL STUD. 689, 691–93 (2004).

27. *See* Brandon Garrett, *Judging Innocence*, 108 COLUM. L. REV. 55, 56–62 (2008) (examining the cases of the first two hundred individuals exonerated by postconviction DNA testing). Along with regulating trial outcomes and sufficiency review in criminal cases, the evidentiary rules also influence plea bargaining (and thus affect false convictions that result from pleas), although so do several other factors. *See* Stephanos Bibas, *Plea Bargaining Outside the Shadow of Trial*, 117 HARV. L. REV. 2463, 2470–96 (2004) (discussing the nonevidentiary factors that affect plea bargaining).

28. For detailed discussions of how issues in civil and criminal procedure depend on evidence law, see Michael S. Pardo, *Pleadings, Proof, and Judgment: A Unified Theory of Civil Litigation*, 51 B.C. L. REV. 1451, 1454–1508 (2010) [hereinafter Pardo, *Pleadings*] (describing how pleadings, summary judgment, and judgment as a matter of law relate to the civil litigation system and the evidentiary proof process); Michael S. Pardo, *Second-Order Proof Rules*, 61 FLA. L. REV. 1083, 1085–1113 (2009) [hereinafter Pardo, *Second-Order*] (arguing that proof rules fail to serve their purported goal in minimizing errors in allocation and proposing "second-order" proof rules). For a more recent effort to link pleadings, summary judgment, and burdens of proof in terms of the costs and benefits of deterring harmful conduct and chilling benign conduct, see Louis Kaplow, *Multistage Adjudication* (Harvard Pub. Law Working Paper 12-41), *available at* http://ssrn.com/abstract=2154683.

foundation of the law of evidence: (1) factual accuracy and (2) allocating the risk of factual errors.[29] This Part argues that evidence theory must account for these foundational considerations and further argues that a failure to do so is a defect for any theory.[30] Thus, whatever its other features, the *nature* of evidence theory includes this epistemological core.[31] The *purposes* of theoretical accounts of evidence and proof fall into two main categories.[32] Theories may be descriptive or explanatory (positive theories),[33] on the one hand, or they may be normative theories, on the other. Theories in the second category purport to provide accounts for correct applications and outcomes.[34] Whichever of these purposes an evidentiary theory is meant to serve, the theory must provide or rely upon a satisfactory account of the epistemological core. A descriptive or explanatory theory that fails to account for these foundational considerations fails as a description or explanation of the proof process; a normative theory that provides an implausible account of this core is not one by which we should judge our current practices.

Part I then identifies three theoretical criteria that a successful evidentiary theory ought to meet. These criteria form the heart of the Article and drive the subsequent analysis. First, the theory must explain, or otherwise rely upon, a satisfactory account of the relevance and probative value of individual items of evidence. This constraint arises at the "micro-level" of proof and underlies the admissibility and

---

29. These considerations are recognized as the primary criteria for formulating proof doctrine and for determining constitutional due process requirements regarding standards of proof. *See infra* note 47.

30. The law in any particular doctrinal area is the result of many different actors pursuing different goals and purposes. It is, of course, not the result of a single mind with focused intentions. Positing particular goals or functions for an area of law is thus open to the potential objections of oversimplifying and perhaps anthropomorphizing legal doctrine. Aware of these pitfalls, the discussion in Part I explains why these considerations are nevertheless foundational in the law of evidence.

31. Sometimes these considerations will be at the forefront of theorizing, and sometimes they will be presupposed by it.

32. *Cf.* Lempert, *supra* note 1, at 441–48 (describing the purposes of mathematical models of evidence to include "prescriptions for action," "normative models," and "descriptive models").

33. Theories in this category may describe how aspects of the process (e.g., a rule) operate in practice, or they may explain how aspects of the process operate in terms of their goals, in terms of other considerations (e.g., efficiency or deterrence), or in terms of underlying causal mechanisms (e.g., psychological or sociological).

34. They may serve a variety of distinct normative roles: a regulative function by providing guidance, a prescriptive function by providing reasons to change current practices, or an evaluative function by providing a standard or criteria for assessment. *See generally* JUDITH JARVIS THOMSON, NORMATIVITY 2 (2009) (broadly discussing normative values and dividing them into "directives" and "evaluatives").

exclusion of evidence at trial.[35] Second, the theory must either explain or rely upon a satisfactory account of the various standards of proof. This constraint arises at the "macro-level" of proof[36] and underlies both sufficiency of the evidence at trial and the various criminal and civil procedure doctrines that depend on the standards of proof.[37] Third, the theory must integrate its accounts at the micro- and macro-levels or otherwise rely upon on a plausible account of how these two levels relate. Inconsistent accounts of the two levels, or a failure to explain how they relate, generate theoretical and practical problems.[38] The criteria are referred to as the (1) "micro-level constraint," (2) "macro-level constraint," and (3) "integration constraint," respectively.

Part II applies the criteria to a *probabilistic* conception of evidence and proof. According to this conception, contested propositions are assessed according to their probability, given the evidence, with such assessments conforming to the axioms or rules of probability theory.[39] This Part argues that theoretical accounts of evidence and proof that rely on a probabilistic conception fail in light of all three criteria. First, the probabilistic conception fails descriptively and normatively at the *micro*-level. In support of this conclusion, this Part argues that evidence may be relevant even when it does not distinguish probabilistically between the parties' claims, and that a change in probability is neither necessary nor sufficient for evidence to be relevant. Second, at the *macro*-level, probabilistic standards of proof fail to fit with their underlying goals regarding

35. FED. R. EVID. 401–03.

36. The micro-macro distinction is meant simply as shorthand for proof issues that pertain to individual items of evidence (admissibility questions), on one hand, and those that pertain to the strength of evidence as a whole on particular issues (sufficiency questions), on the other. Nothing more elaborate is implied by the labels.

37. *See supra* note 25.

38. On the theoretical side, an inconsistent theory cannot be true, and a theory that does not explain the relationship is incomplete. On the practical side, both inconsistent and incomplete theories fail to provide guidance in applying evidentiary rules. *See infra* notes 85–87 and accompanying text.

39. "Probability" in the text refers to the conventionally understood notions associated with mathematical (or "Pascalian") probability—*see infra* notes 109–12 and accompanying text—not to alternative, nonquantifiable notions of probability (i.e., ordinal, inductive, or causative) associated with Francis Bacon, John Stuart Mill, and Jonathan Cohen. *See* L. JONATHAN COHEN, THE PROBABLE AND THE PROVABLE 9–12 (1977) (discussing the distinctions between these theories of probability); Alex Stein, *The Flawed Probabilistic Foundation of Law and Economics*, 105 NW. U. L. REV. 199, 202–07 (2011) (same). It also does not refer to an alternative conception based on "fuzzy logic" that relaxes some of the assumptions of conventional probability theory. *See* Kevin M. Clermont, *Death of Paradox: The Killer Logic Beneath the Standards of Proof*, 88 NOTRE DAME L. REV. (forthcoming 2013), *available at* http://ssrn.com/abstract=1986346 (defending a "fuzzy logic" interpretation of proof standards).

accuracy and the risk of error. This Part explains why probabilistic standards shift too much of the risk of error and fail to foster accuracy or provide criteria for evaluating applications. Third, regarding the integration constraint, this Part explains why the only plausible probabilistic accounts at the micro-level and the macro-level are inconsistent with each other.[40] When combined with tensions at the micro-level, the macro-level and integration constraints render the probabilistic conception implausible.

Part III discusses an *explanatory* conception of evidence and proof. According to this conception, contested propositions are assessed according to how well, if true, they would explain the evidence and events when compared with competing, contrastive propositions.[41] This Part argues that this conception fits with each of the three theoretical criteria, explains how it avoids the conceptual problems that inhere in the probabilistic conception, and responds to several counterarguments. First, at the micro-level, relevance and probative value depend on whether evidence forms part of, supports, or challenges one of the explanations put forward by the parties.[42] Second, at the macro-level, standards of proof are expressed in terms of explanatory criteria. The quality of an explanation needed to satisfy a standard of proof varies depending on the applicable standard—for example, the quality of an explanation needed to satisfy the preponderance standard is lower than that needed to satisfy the clear-and-convincing-evidence and beyond-a-reasonable-doubt standards.[43] Third, the explanatory conception satisfies the integration constraint because the micro- and macro-level accounts cohere in a straightforward way. The same explanatory considerations that elucidate relevance and probative value at the micro-level also elucidate the standards of proof and evidential sufficiency at the macro-level.[44]

---

40. Discussions in the scholarly literature tend to focus on either micro-level or macro-level proof issues, but not both, and thus the integration problem escapes notice. So far as I am aware, this Article provides the first detailed discussion of the problem.

41. The explanatory conception is discussed in detail in Pardo & Allen, *supra* note 14. The focus on explanatory criteria (or "inference to the best explanation," *see infra* notes 211–16) provides an epistemological foundation for the "relative plausibility" theory developed previously by Professor Allen. *See* Ronald J. Allen, *Factual Ambiguity and a Theory of Evidence*, 88 NW. U. L. REV. 604, 606 (1994).

42. This conception accounts for relevant evidence that does not distinguish between the parties' cases. *See infra* Part III.A.

43. This Part explains how explanatory standards of proof (1) fit with the goals of the standards regarding accuracy and the risk of error and (2) provide guidance for evaluating applications. *See infra* Part III.B.

44. *See infra* Part III.C.

Finally, a brief Conclusion outlines some general lessons for the projects of evidence theory.

I. EVIDENCE THEORY

Evidence theories provide conceptions or accounts of the structure and nature of the evidentiary proof process.[45] Their scope may include the proof process as a whole, or they may focus on a particular aspect of the process (for example, a particular rule, type of evidence, or type of reasoning). An evidence theory may aim to serve a *descriptive* or an *explanatory* function, in which case its success depends on how well it captures the underlying phenomena. An evidence theory may also serve *normative* functions, including an *evaluative* function by providing criteria for justifying or critiquing particular judgments, and a *regulative* function by providing considerations for guiding and constraining particular judgments. The success of such theories depends on whether the outcomes they specify or imply are justified in light of the goals of the proof process. Moreover, to serve these normative roles, a theory must be one that decisionmakers are capable of applying in practice.[46]

The evidentiary proof process may be evaluated based on two considerations: (1) factual accuracy and (2) allocation among the parties of the risk of factual errors. Any evidence theory—whether its aims are descriptive, explanatory, or normative—will provide, rely upon, or otherwise presuppose some account of how these considerations relate to the process of proof. The considerations, which are epistemological in nature, form a foundational core in the domain of evidence theory.[47] For the reasons discussed below, this

---

45. *See supra* note 21.

46. Theories in this context must be able to describe, explain, guide, constrain, or evaluate actual judgments and possible judgments that decisionmakers are capable of making. *See* Allen & Leiter, *supra* note 14 (discussing "ought implies can" in epistemological theory). Although there is theoretical value in theories of imaginative things or unrealizable ideals, the utility of such theorizing is diminished in practical settings like legal proof where actual applications are the proper object of theorizing. Compare, for example, a theory of balls and strikes in baseball that could never be implemented or could never be used to tell whether an actual pitch was a ball or strike.

47. This epistemological core forms the foundation of what William Twining has referred to as "the Rationalist Tradition" in evidence scholarship. WILLIAM TWINING, RETHINKING EVIDENCE: EXPLORATORY ESSAYS 33–91 (1st ed. 1990). The United States Supreme Court has recognized both of these goals as foundational. *See* Grogan v. Garner, 498 U.S. 279, 286 (1991) (risk of error); United States v. Havens, 446 U.S. 620, 626–27 (1980) (factual accuracy); Addington v. Texas, 441 U.S. 418, 423–34 (1979) (risk of error); Tehan v. United States, 382 U.S. 406, 415–16 (1966) (factual accuracy).

epistemological core provides a necessary (although not sufficient) condition for a successful evidence theory. To be clear, this is not to deny the significance of a host of other theoretical and practical issues (moral, political, and economic) within the law of evidence.[48] Rather, the discussion below focuses on this epistemological core because it provides a necessary step in constructing and evaluating an evidence theory.

Why, exactly, is this epistemological core necessary? To answer this question, we need to step back and look at procedural theory more generally and the role of evidence law within litigation. The civil and criminal procedural systems are animated by an array of procedural values. These values include: accuracy, efficiency, participation, respect for substantive rights, notice, predictability, fairness, equality, and political legitimacy.[49] Each of these values provides a criterion by which procedural theories may be evaluated.

The evidentiary process is the primary procedural location where policies regarding accuracy and the risk-of-error allocation are implemented. An evidentiary process that fails to implement these policies fails to serve its procedural function, and a theory that fails to account for how the law succeeds or fails at implementing these policies likewise fails as a theory.[50] This illustrates that this epistemological core is a *necessary* part of the domain of evidence theory, but why, exactly, is it *foundational*?

It is foundational because the other procedural values depend on this core. An evidentiary process that fails in light of this core will also fail to serve the other values, and a theory that cannot account for this core will likewise fail in light of the other values. Consider, first, factual accuracy. A theory of evidence that does not account for a sufficient level of accuracy also cannot explain or account for other procedural values. Specifically, an evidentiary process that delivered systematically inaccurate outcomes would (1) be terribly inefficient; (2) impose enormous costs on parties and society; (3) undermine notice and participation rights; (4) create unpredictability; (5) fail to guide

48. For an overview of recent scholarship on these other issues, see Park & Saks, *supra* note 2.

49. Although these values underlie both civil and criminal litigation, they are instantiated differently given the interests at stake. Civil litigation is often animated by considerations of equality, while criminal litigation is more often animated by respect for the rights of defendants.

50. A theory that succeeds in light of these goals may nevertheless fail in light of other procedural values.

and deter primary conduct; (6) violate substantive rights; and (7) very quickly become politically illegitimate.[51]

Now, consider allocating the risk of error. In civil cases, this consideration is aligned with the procedural values of fairness and equality.[52] The evidentiary proof process purports to treat parties equally with regard to the risk of error. The preponderance-of-the-evidence standard of proof is employed for this purpose,[53] and the rules of evidence apply symmetrically among litigants. Thus, a theory that fails to account for how the system allocates the risk of error in this manner also fails in light of the procedural values of fairness and equality. The criminal system, by contrast, purports to skew the risk of error to a substantial degree away from criminal defendants. This goal is manifested most obviously in the constitutionally required beyond-a-reasonable-doubt standard of proof,[54] but it also manifests itself in evidence rules that apply asymmetrically among the parties.[55]

---

51. Imagine a world in which litigation results in mostly erroneous judgments, and consider what would follow. Citizens would have an incentive to file lawsuits regardless of the underlying merits (indeed, particularly when they were not wronged). By contrast, citizens who commit crimes and wrongs would win at trial. The result—at least initially—would be a great deal more litigation in which costs are imposed on the legal system, innocent defendants who lose, and injured plaintiffs who could not recover. Similar costs would arise in criminal cases, with innocent defendants convicted and guilty ones acquitted. Rights to notice and participation would become largely meaningless, because these rights are related to (although not reducible to) accuracy considerations—parties need notice and the right to participate so that they can prove or defend their cases on the merits. Moreover, without the realistic possibility of accurate trial outcomes the substantive law would also become largely meaningless—conforming to the dictates of the law would be no protection against losing at trial, and violating the dictates could be done largely with impunity. Parties could not predict when they would be sued or prosecuted, nor could citizens predict when their rights would be violated. It is not hard to imagine that, in this dystopian world, legal judgments would soon fail to be legitimate, both descriptively (i.e., people would not perceive them as genuinely authoritative or as creating obligations) and normatively (i.e., they would not satisfy minimal standards for acceptable or justified exercises of political power that create genuine obligations). For a general overview of these two senses of political legitimacy, see Fabienne Peter, *Political Legitimacy*, STANFORD ENCYCLOPEDIA PHIL. (Apr. 29, 2010), http://plato.stanford.edu/entries/legitimacy/.

52. *See* Ronald J. Allen, *The Error of Expected Loss Minimization*, 2 L., PROBABILITY & RISK 1, 4 (2003) ("[E]qual treatment is incontrovertibly one critical component of fairness."); Lawrence B. Solum, *Procedural Justice*, 78 S. CAL. L. REV. 181, 237–73 (2004) (discussing the values that contribute to fair procedure).

53. Grogan v. Garner, 498 U.S. 279, 286 (1991) (explaining that the preponderance standard "results in a roughly equal allocation of the risk of error"); *accord* Herman & Maclean v. Huddleston, 459 U.S. 375, 390 (1983) (stating a preponderance standard "shares the risk of error in roughly equal fashion"); *see also* Redmayne, *supra* note 24, at 171–74 (discussing the principle of equality in civil cases).

54. *In re* Winship, 397 U.S. 358, 363 (1970).

55. *See, e.g.*, FED. R. EVID. 404 (giving defendants the option to introduce character evidence); FED. R. EVID. 410 (excluding evidence of statements made during failed plea negotiations, unless defendants first introduce related statements).

A theory that failed to explain or account for this risk-of-error allocation would also fail in light of other procedural values, including respect for the constitutional rights of criminal defendants.

Accuracy and allocating the risk of error thus provide foundational considerations in elucidating a successful theory of evidence and proof. In what follows, this Part articulates theoretical criteria in light of these considerations. The criteria focus on the two primary aspects of proof—admissibility and sufficiency—and the relationship between these aspects. The epistemological considerations aid in determining whether an evidence theory satisfies a theoretical criterion. In other words, a theory aiming to describe or explain an aspect of the process (e.g., the sufficiency of evidence to meet a decision standard) should be able to describe or explain how that aspect meets (or fails to meet) the system's goals regarding accuracy and the risk of error. Similarly, a normative theory should be able to explain why the theory would improve the system with regard to accuracy or the risk of error, or why, if based on other normative considerations (e.g., other procedural values), it would not lead to undesirable epistemic consequences. After articulating the theoretical criteria, this Part explains the relationships between the criteria and the array of evidentiary rules, and it closes with a number of clarifying assumptions.

A. Theoretical Criteria

A successful theory of evidence and proof must satisfy three constraints.[56] First, it must provide, or rely upon, a plausible account of the relevance and probative value of individual items of evidence. This is the micro-level constraint. Second, it must provide, or rely upon, a plausible account of the standards of proof. This is the macro-level constraint. Third, it must provide, or rely upon, a plausible account of the relationship between the accounts at these two levels. This is the integration constraint.

1. The Micro-Level Constraint

A theoretical account of evidence and proof should be able to explain the relevance and probative value of evidence.[57] The two

56. Although necessary, the constraints are not sufficient (individually or jointly). *See supra* notes 50–52 and accompanying text.

57. FED. R. EVID. 401 (defining relevance); FED. R. EVID. 402 (excluding irrelevant evidence).

fundamental tenets of modern evidence law are that (1) irrelevant evidence ought to be excluded, and (2) relevant evidence ought to be admitted, unless there is some good reason to exclude it.[58] The presuppositions underlying these tenets are that irrelevant evidence will not contribute to a rational assessment of disputed factual issues and that, all other things being equal, the addition of more relevant evidence (properly understood and interpreted) will lead to a more rational assessment of the disputed issues and thus more accurate decisions.[59] A satisfactory theory of evidence and proof ought to be able to explain whether evidence is relevant, along with an explanation *why*.[60]

A theory of evidence and proof will be potentially problematic if the answers it provides or implies about the relevance of evidence fail to accord with our considered judgments about the relevance of evidence. Neither the theory nor particular judgments ought to necessarily have priority, however, in every case.[61] If there is a mismatch, then either (1) the theory is false or fails as a theoretical account, or (2) the underlying judgments are false and ought to change in light of the theory. In this context, the judgments ought to give way, for example, if the outcomes provided or implied by a satisfactory theory would better foster accuracy. Although it is thus possible for a

---

58. THAYER, *supra* note 2, at 264–66. Justifications for excluding relevant evidence are discussed in Part I.B *infra*.

59. "Accuracy" refers to whether the outcomes match *what actually happened*, not merely whether they are supported by the evidence. *See* Pardo, *Pleadings*, *supra* note 28, at 1470–71 (distinguishing the former as "material accuracy" and the latter as "procedural accuracy"). This relationship between relevance and accuracy depends on proper interpretation by the decisionmaker. Michael Risinger refers to this as the "god perspective" view of relevance, in which the "inference maker knows all that is necessary to make as much accurate inference as possible from any given item proposed for consideration, and has no time or processing capacity constraints." D. Michael Risinger, *Inquiry, Relevance, Rules of Exclusion, and Evidentiary Reform*, 75 BROOK. L. REV. 1349, 1353 (2010).

60. *See supra* note 46.

61. This relationship between particular judgments and general theories is consistent with the idea of "reflective equilibrium" introduced by Nelson Goodman in discussing inferential rules, NELSON GOODMAN, FACT, FICTION, AND FORECAST 59–124 (4th ed. 1983), and made famous by John Rawls in JOHN RAWLS, A THEORY OF JUSTICE 18–22, 46–53 (1971). Judgments about particular cases, on one hand, and general rules, principles, or theories, on the other, are gradually brought into agreement with each other by refining each in light of the other. For an instructive overview of this process for Rawls, see John Mikhail, *Rawls' Concept of Reflective Equilibrium and its Original Function in* A Theory of Justice, 3 WASH. U. JURIS. REV. 1 (2010). *See also* Mitchell N. Berman, *Reflective Equilibrium and Constitutional Method: Lessons from John McCain and the Natural-Born Citizen Clause*, *in* THE CHALLENGE OF ORIGINALISM 246 (Huscroft & Miller eds., 2011) (essay "adapts the Rawlsian method of reflective equilibrium to the constitutional domain").

theory's implications to override particular judgments,[62] the initial burden is on the proponent of a theory to explain why a considered concrete judgment on an issue ought to give way to the dictates of an abstract theory. Judgments about particular cases, in other words, provide the starting point for theorizing; these judgments, however, are defeasible and subject to revision in light of a compelling theory.[63]

Similar theoretical and methodological considerations apply to probative value. The probative value of evidence refers to the strength of that evidence in proving a disputed factual proposition. Judgments on probative value are highly contextual and depend not only on the logical or empirical relationships between evidence and disputed propositions, but also on the importance of the evidence; the party's need for the evidence; the availability of other similar evidence; the extent to which the issue is contested; and whether the other side has introduced contrary evidence.[64] Probative value is a fundamental consideration for the admissibility of evidence—judges may exclude virtually any item of evidence if its probative value is substantially outweighed by the potential dangers of unfair prejudice, confusing the issues, or misleading the jury, or for efficiency considerations (undue delay, waste of time, or needlessly cumulative).[65] Thus, any satisfactory theory must contain a plausible explanation of the probative value of individual items of evidence. Similarly, if there is a mismatch between judgments about probative value and the answers provided by a theory, the initial burden is on the proponent of the

62. This possibility is necessary for a theory to play any sort of normative role. In other words, if the theory provides a standard for correct and incorrect judgments, then it must be possible that some judgments are incorrect.

63. Why start with particular judgments? The short answer is that legal proof is a complex social activity and thus is not likely capable of being captured fully by an abstract theory. Therefore, a theory needs to provide compelling reasons to override judgments made by experienced practitioners engaged in the practice, who are more likely (than theorists) to have practical knowledge of what is required in particular situations even if that knowledge cannot be articulated propositionally in a theory. For a defense of the primacy of particular judgments along these lines, see BURNS, *supra* note 14. This assumption, however, is *defeasible*.

64. *See* Old Chief v. United States, 519 U.S. 172, 182–83 (1997) (discussing the factors that affect probative value).

65. FED. R. EVID. 403 (providing that trial courts may exclude evidence if its probative value is "substantially outweighed" by these countervailing considerations). *But see* FED. R. EVID. 609(a)(2) (dictating that prior convictions involving dishonest acts or false statements shall be admitted for impeachment purposes). Probative value is also important at the macrolevel because determinations of whether evidence is sufficient to satisfy a decision standard are essentially determinations regarding what a reasonable jury could conclude about the probative value of the evidence as a whole.

theory to explain why evidence that appears to have low probative value in fact has high probative value, or vice versa.[66]

2. The Macro-Level Constraint

A theory of evidence and proof must provide, or otherwise rely upon, a plausible account of the nature and structure of proof at the macro-level. Rather than focusing on individual items of evidence, explanations at this level focus on whether micro-level evidence as a whole satisfies a standard of proof—in other words, whether or not a fact has indeed been "proven" for legal purposes.

At the macro-level, one party carries the burden of proof on each litigated issue that forms an element of a civil claim, crime, or affirmative defense. The burden of proof includes two components: a burden of production and a burden of persuasion. The burden of production is a function of the burden of persuasion—parties will have satisfied a *production* burden when they have presented evidence from which a reasonable fact finder could conclude that the burden of *persuasion* has been satisfied.[67] The burden of persuasion is determined by the various proof standards: preponderance of the evidence, clear and convincing evidence, or beyond a reasonable doubt.[68]

---

66. *See supra* note 63.

67. The classic discussion of this relationship is John McNaughton, *Burdens of Production of Evidence: A Function of the Burden of Persuasion*, 68 HARV. L. REV. 1382, 1383–85 (1955).

68. The standards apply to general categories of cases. *See* Santosky v. Kramer, 455 U.S. 745, 757 (1982) (rejecting the process of determining the applicable proof standard on a case-by-case basis). On the wisdom of having three decision standards, see Kevin M. Clermont, *Procedure's Magical Number Three: Psychological Bases for Standards of Decision*, 72 CORNELL L. REV. 1115, 1154–55 (1987) (explaining the cognitive limitations on more fine-grained judgments). On the wisdom of the general winner-take-all approach to burdens of persuasion, as opposed to proportional or fractional judgments, see David Kaye, *The Limits of the Preponderance of the Evidence Standard: Justifiably Naked Statistical Evidence and Multiple Causation*, 7 AM. B. FOUND. RES. J. 487, 514–16 (1982) (concluding if the most likely cause of injury is "a legally responsible defendant," that defendant should be fully liable). Additional complexities may arise when jurors accept the same outcome but disagree on the reasons; for an illuminating discussion, see Peter Westen & Eric Ow, *Reaching Agreement on When Jurors Must Agree*, 10 NEW CRIM. L. REV. 153, 155–202 (2007) (analyzing the difficulties that arise when jurors disagree on the means by which an offense was committed). I put aside this complexity in the discussion to follow. For comparative perspectives of proof standards, compare Kevin M. Clermont & Emily Sherwin, *A Comparative View of Standards of Proof*, 50 AM. J. COMP. L. 243, 244–74 (2002) (looking at differences between common law and civil law standards of proof and concluding the preponderance-of-the-evidence standard should apply), with Michele Taruffo, *Rethinking the Standards of Proof*, 51 AM. J. COMP. L. 659, 660–73 (2003) (critiquing the Clermont and Sherwin article).

The proof standards specify when evidence proves a particular disputed fact. Considerations of accuracy and allocating the risk of error animate the standards. Each consideration provides an independent criterion by which to assess the acceptability of the standards. For example, the preponderance-of-the-evidence standard functions (or aims to function) to minimize errors and maximize accuracy, other things being equal.[69] This follows from the fact that decisions made under this standard favor the most likely alternative, given the admitted evidence.[70] Assuming that evidence generally is a good indicator of truth,[71] decisions that appear more likely given the evidence will be more accurate than decisions that appear less likely.[72] The preponderance standard also allocates the risk of error roughly equally among the parties (either side may have the less likely case and the decision may be an error).[73] Although this allocation is acceptable in most civil cases because equality is a primary procedural

69. These functions also depend on other factors such as the distribution of truly deserving plaintiffs and defendants who go to trial, the quality of the evidence and the parties' access to it, and whether evidence is properly interpreted by fact finders.

70. *See* STEIN, *supra* note 14, at 143–44 (discussing the preponderance standard in light of these considerations).

71. This assumption underlies having an evidentiary proof process in the first place. If evidence does not bear this relation to truth, then dispute resolution should proceed in some other manner.

72. Another way to perceive how the preponderance standard may minimize total errors is to reflect on likely errors under higher or lower standards. *But see supra* note 69 and accompanying text. When the standard is higher than a preponderance, there will be a class of cases in which certain facts are more likely to be true than false, but they will be found to be "not proven" because they do not reach the threshold for the higher standard. Because these facts are more likely true than false, we would expect more errors than correct judgments in this class of cases in the long run. Vice versa for a lower standard, except now the errors will result from finding facts proven even when they are more likely false than true. Under the preponderance standard, by contrast, we would expect more correct judgments than errors for these classes of cases.

73. Either side may have the less likely case, and the decision may be an error, with cases of evidential "ties" going against the party with the proof burden. *See* Eyal Zamir, *Loss Aversion and the Law*, 65 VAND. L. REV. 829, 833 (2012) (defending the policy of ties going to the defendant on the ground that errors will seem more costly to defendants than to plaintiffs). *See generally* Adam Samaha, *Law's Tiebreakers*, 77 U. CHI. L. REV. 1661, 1675–84 (2010) (discussing the policies underlying tiebreaker rules).

value,[74] this is generally deemed an unacceptable risk of error for criminal defendants to bear for the elements of criminal offenses.[75]

A satisfactory theory ought to explain the proof standard in light of these considerations. Successful explanations will include how the standards allocate the risk of error. A satisfactory theory will also explain how the standards relate to accuracy.[76] Moreover, theories ought to provide guidance for decisionmakers.[77] Standards of proof are instructions to judges and juries on how to arrive at judgments. Correct application of the standards is thus the object of successful theorizing. A theory that can explain formally how the standards fit with their purported goals, but cannot explain how fact finders could ever implement the standards consistent with the goals, fails as a theory.[78]

A theory ought to also provide guidance in separating "reasonable" from "unreasonable" applications. This important doctrinal issue has, unfortunately, not been the subject of sustained dialogue among proceduralists.[79] It is critical not only for proof at trial; it arises throughout the civil and criminal litigation systems more generally. On the civil side, the standards for summary judgment before trial and judgment as a matter of law at or after trial depend on whether a "reasonable" fact finder could find for the nonmoving party.[80] What is reasonable is a function of the evidence,

---

74. *See* Grogan v. Garner, 498 U.S. 279, 286 (1991) ("[T]he preponderance-of-the-evidence standard results in a roughly equal allocation of the risk of error between litigants."); Herman & Maclean v. Huddleston, 459 U.S. 375, 391 (1983) (declining to depart from the preponderance-of-the-evidence standard); Redmayne, *supra* note 24, at 171–74 (discussing the equality principle and civil litigation).

75. *See In re* Winship, 397 U.S. 358, 364 (1970) (expressing the constitutional goal of reducing risk of errors where personal liberty is at stake in criminal cases); *see also* Addington v. Texas, 411 U.S. 418, 424 (1979) (explaining that the "clear and convincing" standard applies when there are asymmetrical interests at stake in civil cases).

76. These explanations will include how the preponderance standard may maximize accuracy, and how other standards may sacrifice accuracy. Total accuracy may be sacrificed because cases proven by a preponderance of the evidence (but not to the higher standard) will go against the party with the proof burden, even though that party's case is more likely to be accurate.

77. *See supra* note 46.

78. How the standards ultimately operate in practice is a complex empirical question that depends on a host of issues, including: the other evidentiary and procedural rules; the quality of the evidence; the ability of fact finders to properly assess the evidence; and the distribution of deserving parties on each side. It is not a defect of a theory that it cannot provide answers to these empirical questions; rather, theories clarify the issues in need of further empirical investigation.

79. *See supra* note 28.

80. *See supra* note 25.

the burden of proof, and the decision standard. For example, summary judgment may depend on whether a reasonable fact finder could find for the plaintiff on a particular issue (e.g., causation) by a preponderance of the evidence.[81] This implies some understanding or presupposition of what is required by the preponderance standard; otherwise we could not distinguish between reasonable and unreasonable applications. The same goes for judgment-as-a-matter-of-law determinations.[82] Thus, whether parties get to trial in the first place and whether particular verdicts are upheld depend on a conception of the proof standard.

In criminal cases, likewise, whether evidence is sufficient to support a conviction depends on whether a reasonable jury could find the issues proven beyond a reasonable doubt.[83] This determination depends on a conception of what is required by the beyond-a-reasonable-doubt standard.[84] Whether cases will be brought to trial, dismissed at trial, or reversed on appeal depend on a distinction between reasonable and unreasonable applications. A successful theory must explain this critical distinction and provide criteria by which to guide, evaluate, justify, or critique particular judgments.

3. The Integration Constraint

Finally, a theory of evidence and proof must provide, or rely upon, a plausible explanation of how the micro- and macro-levels fit together. A theory that provides plausible accounts at both levels may nevertheless fail if these accounts are inconsistent. For example, a theory that, at the *micro*-level, accounts for relevance and probative value in terms of the subjective beliefs of individual fact finders (i.e., that there is nothing more to relevance and probative value than what each fact finder thinks, perhaps like matters of taste) cannot then, at the macro-level, account for sufficiency of the evidence in terms that imply that the subjective beliefs of individual fact finders may be mistaken. Either there is more to the value of evidence than subjective

---

81. *See*, *e.g.*, Yeschick v. Mineta, 675 F.3d 622, 633 (6th Cir. 2012) (upholding summary judgment for defendant on issue of causation); *cf.* Zuchowitz v. United States, 140 F.3d 381, 387 (2d Cir. 1998) (analyzing whether finding of causation was clear error).

82. *See* St. Mary's Honor Ctr. v. Hicks, 509 U.S. 502, 509 (1993) (discussing the relationship between burdens of proof and judgment as a matter of law).

83. *See, e.g.*, Jackson v. Virginia, 443 U.S. 307, 318–19 (1979) (discussing the reasonable jury standard).

84. *See, e.g.*, Roth, *supra* note 7, at 1147–49 (assuming that DNA evidence requires a probabilistic interpretation of "beyond a reasonable doubt" to determine sufficiency).

beliefs or there is not. A theory trying to have it both ways must at least explain how this is possible and why it is plausible.[85]

A theory may also be problematic if the relationships between the two levels are left unexplained or are otherwise mysterious. For example, a theory that explains relevance in terms of one set of criteria and sufficiency in terms of a different set of criteria[86] must explain how the micro-level criteria translate at the macro-level, and vice versa. In sum, a successful theory will bridge the gap between the two levels.[87]

B. Additional Evidentiary Rules

The rules discussed thus far provide the doctrinal foundation for admissibility and sufficiency decisions. A matrix of additional evidentiary rules, however, also regulates specific aspects of the proof process. For purposes of completeness and clarification, this Section briefly explains the relationships among these additional rules; the three theoretical criteria (micro, macro, and integration); and the epistemic considerations (accuracy and the risk of error).

From an epistemic perspective, the addition of relevant evidence at the micro-level increases the evidentiary base on which decisions are made and thus should, other things being equal, improve accuracy.[88] Therefore, micro-level rules that exclude relevant evidence

---

85. Such an explanation may perhaps be constructed around the idea that there should be more judicial control at one level rather than the other. This explanation, however, would fail to explain current doctrine, which allows for substantial judicial control of the fact-finding process at both levels.

86. For example, one level may be explained in terms of subjective probability assessments and the other level in terms of objective relative frequencies.

87. The better the relationship between these levels is explained, the better the theory. To illustrate how this constraint may create problems for a theory, consider two nonlegal examples. Suppose one claimed that morality is a matter of subjective taste (there are no moral facts) but also claimed that a particular judgment on a moral issue was objectively true (a matter of fact). Or suppose one claimed that water molecules are made of hydrogen and oxygen but also claimed that water in sufficiently large quantities no longer contained hydrogen or oxygen. No doubt apparent inconsistencies can sometimes be explained or explained away (we can explain why water is wet but water molecules are not wet, for instance). In the examples, however, the tension between the two claims creates additional problems in need of explanation, potentially undermining each of the claims involved, in addition to any problems each claim may have on its own. An evidentiary theory is better to the extent it avoids problems along these lines.

88. *See* RUDOLF CARNAP, LOGICAL FOUNDATIONS OF PROBABILITY 211 (1950) ("[T]he total evidence available must be taken as a basis for determining the degree of confirmation."); TIMOTHY WILLIAMSON, KNOWLEDGE AND ITS LIMITS 189–90 (2002) (discussing the "total evidence" principle). When the relevance of evidence is itself in doubt, or depends on facts that are uncertain or unlikely, evidence may be admitted "conditionally," subject to proof of these conditional facts. FED. R. EVID. 104(b). It is important to note, however, that *all* evidence is

must be justified. The possible justifications fall into four categories: (1) economic, (2) other, nonepistemic policy goals, (3) jury control, or (4) party control. First, evidence may be justifiably excluded if it is not worth the costs of presenting and considering it. This may be the case if it has only minimal probative value, it is cumulative of other evidence, or its evidentiary value on the particular litigated issue is otherwise not worth the costs.[89] Second, relevant evidence may be justifiably excluded to serve a number of nonepistemic purposes internal and external to the litigation process.[90] "Internal" rules include, for example, the attorney-client privilege and the rule excluding evidence of compromise negotiations and related statements.[91] "External" rules include, for example, the marital privileges and the rule excluding evidence of subsequent remedial measures.[92] The third and fourth categories serve epistemic rationales. Third, evidence may be excluded if jurors are prone to "overvalue" the evidence (or otherwise be distracted by it) to such an extent that it will detract from rather than aid in rational assessment of the evidence and the disputed issue.[93] Fourth, evidence may be excluded if doing so is likely to induce parties to produce better evidence.[94]

Although the relationship between admissibility rules and the micro-level constraint is more readily apparent, the admissibility rules also interact with the *macro-level* and *integration* constraints. This is

---

conditionally relevant in the sense that its value is contingent on the other evidence in the case and what is known or not known about that evidence. *See* Ronald J. Allen, *The Myth of Conditional Relevancy*, 25 LOY. L.A. L. REV. 871 (1992) (extending the analysis of Vaughn Ball *infra*); Vaughn C. Ball, *The Myth of Conditional Relevancy*, 14 GA. L. REV. 435, 436 (1980) (discussing this contingency). For an earlier discussion examining this contingency in the context of authentication requirements, see Larry A. Alexander & Elaine A. Alexander, *The Authentication of Documents Requirement: Barrier to Falsehood or to Truth*, 10 SAN DIEGO L. REV. 266 (1973). For a recent defense of foundational requirements, including conditional relevance, see David A. Schwartz, *A Foundation Theory of Evidence*, 100 GEO. L.J. 95 (2011).

89. FED. R. EVID. 403.

90. "Internal" justifications include reasons for exclusion that are designed to aid or foster aspects of the litigation process; "external" justifications include reasons for exclusion that are designed to serve goals outside of litigation.

91. FED. R. EVID. 410.

92. FED. R. EVID. 407.

93. FED. R. EVID. 403.

94. FED. R. EVID. 602, 801–07, 901, 1001–08. *See generally* Dale A. Nance, *The Best Evidence Principle*, 73 IOWA L. REV. 227 (1998). Note, however, that there is no *necessary* connection between probative value and these requirements. In other words, sometimes evidence that fails to meet them may be more probative than evidence that does. The rules regulating expert testimony, *see* FED. R. EVID. 702–06, are arguably justified based on some combination of the first, third, and fourth rationales.

so because—in addition to altering the evidentiary base and thus the sufficiency of evidence at the macro-level—admissibility rules may also shift the risk of error. For example, the exclusion of a document on authentication grounds[95] may cause a party with the burden of proof (who would otherwise win) to lose if the evidence without the document is insufficient (but would have been sufficient with the document).[96] Moreover, a rule that admits or excludes evidence that is typically proffered by one side rather than the other systematically shifts a risk of error at the macro-level.[97]

C. Clarifying Evidence Theory: Assumptions, Distractions, Confusions

Given the explosion of theoretical work in evidence law, metatheoretical discussions travel on well-worn and hotly contested ground.[98] Therefore, before proceeding to apply the theoretical criteria, this Section briefly clarifies what is and what is *not* at stake in the analysis to follow. In order to avoid possible misunderstanding and confusions, the discussion proceeds by isolating the theoretical issues examined in this Article from adjacent issues in the scholarly literature that are not necessarily implicated.

In current debates about mathematical models of evidence,[99] there are two different projects at the heart of the discussions. Although they are related, it helps to keep them distinct. The first project is the integration, as a practical matter, of statistical information and statistical techniques into the litigation process.[100]

---

95. FED. R. EVID. 901.

96. *See* Alexander & Alexander, *supra* note 88, at 269. Exclusion under these circumstances may also function to shift a burden of proof. *See id.* at 274.

97. *See, e.g.*, FED. R. EVID. 412–15 (admitting types of prosecution and plaintiff evidence, and excluding types of defense evidence, in sexual assault cases). Other evidentiary devices, including presumptions and judicial comment on the evidence, may also shift the risk of error. *See* RONALD J. ALLEN ET AL., EVIDENCE: TEXT, CASES & PROBLEMS 717–71 (5th ed. 2011) (discussing presumptions, comment, and burdens of proof). Criminal offenses may also operate in this manner when they penalize conduct not for its own sake but because of its probabilistic or evidentiary relationship with other (harder to prove) criminal conduct. *See* Frederick Schauer, *Bentham on Presumed Offences*, 23 UTILITAS 363 (2011) (analyzing several examples of such offenses).

98. There appears to be substantial disagreement among evidence scholars about the nature of relevance and probative value. *See* Roger Park et al., *Bayes Wars Redivivus—An Exchange*, 8 INT'L COMM. ON EVID. 1 (2010).

99. *See supra* notes 7–13.

100. *See, e.g.*, Edward K. Cheng, *When 10 Trials Are Better than 1000: An Evidentiary Perspective on Trial Sampling*, 160 U. PA. L. REV. 955 (2012); Edward K. Cheng, *A Practical Solution to the Reference Class Problem*, 109 COLUM. L. REV. 2081 (2009) [hereinafter Cheng, *Practical Solution*]; Alexandra D. Lahav, *The Case for 'Trial by Formula'*, 90 TEX. L. REV. 571,

The second is the integration, as a theoretical matter, of statistical and other evidence into a conception or theoretical account of the proof process that accords with the process's normative goals. These projects are, of course, related—the theoretical accounts provide guidance for practical integration as well as a standard by which to evaluate these practices. For example, we have no idea whether evidence (statistical or otherwise) is sufficient to satisfy the preponderance or beyond-a-reasonable-doubt standards without some prior conception or presupposition of what these standards require. The two projects, however, involve distinct methodological issues. The practical issues of integration involve difficult empirical questions that are distinct from the conceptual task of articulating an underlying theoretical account of the nature and structure of legal proof. This Article focuses on the theoretical, conceptual project.

To clarify this conceptual project, and to isolate it from adjacent issues, this Part closes by making explicit six assumptions presupposed by the analysis thus far and on which the applications to follow will rely:

(1) *Statistical evidence is distinct from a conception or account of the proof process.* In other words, whether statistical evidence is relevant, how probative it is, and whether it is sufficient to satisfy a proof standard are distinct issues from whether relevance, probative value, or proof standards are probabilistic or ought to be conceived in probabilistic terms. For example, consider the plaintiff's evidence in a negligence case that the defendant owns ninety-five percent of the buses in town in order to prove that one of the defendant's buses caused the accident that injured the plaintiff. The relevance, probative value, and sufficiency of the ninety-five percent datum as evidence in the case are distinct issues from whether the evidentiary concepts and rules are probabilistic or ought to be conceived as such. One cannot draw evidentiary conclusions about statistical evidence for purposes of legal proof without relying upon some conception of what is required by the evidentiary rules and concepts.[101]

593 (2012) (advocating the use of statistical methods to improve outcome equality in civil cases); Michael I. Meyerson & William Meyerson, *Significant Statistics: The Unwitting Policy Making of Mathematically Ignorant Judges*, 37 PEPP. L. REV. 771 (2010); *see also* COLIN AITKEN, PAUL ROBERTS & GRAHAM JACKSON, FUNDAMENTALS OF PROBABILITY AND STATISTICAL EVIDENCE IN CRIMINAL PROCEEDINGS (Royal Stat. Soc'y, 2011).

101. For example, one cannot determine whether statistical evidence is sufficient to satisfy the preponderance standard without some conception of what is required by the standard. "Greater than 0.5" is one interpretation, but it is not the only one. What is required by the preponderance standard is a conceptually distinct issue from the sufficiency of statistical evidence in a given case.

(2) *The practical issues surrounding whether a type of evidence (e.g., statistical evidence or statistical techniques) ought to be introduced in litigation are distinct issues from the theoretical issues regarding conceptions of evidence and proof.* The practical issues depend on whether it will improve or detract from accurate fact-finding, and they depend on complex empirical questions concerning how fact finders are likely to respond to the evidence under different conditions, compared with likely fact finder behavior in the absence of this evidence.[102]

(3) *Statistical evidence is neither necessarily inferior nor superior to nonstatistical evidence as an epistemological matter*. There is no abstract *epistemic* distinction, *a priori* or otherwise, between evidence expressed in statistical terms and evidence expressed in nonstatistical terms.[103]

(4) *Both statistical and nonquantified evidence are based on similar types of generalizations.* For example, the probative value of a witness's testimony depends on generalizations about the witness's propensity for accurate perception, memory, narration, and truthfulness, as well as generalizations about the behavior of other people in the witness's situation. These generalizations link the testimony to the facts of a particular case in ways that are substantially similar to the generalizations that apply to statistical evidence.[104]

(5) *Theoretical conceptions of evidence may be useful as teaching tools or for drawing attention to certain issues, even if they otherwise fail as a theory*. That they serve this heuristic value should not be underestimated or diminished. But serving heuristic or pedagogical purposes should also not be confused with a successful theory.

---

102. Even if jurors are not particularly adept at assessing statistical evidence, *see* Koehler, *supra* note 18, the introduction of this evidence may still do epistemic good if decisions are better than what they would be without the evidence. *Cf.* Frederick Schauer, *Can Bad Science Be Good Evidence? Neuroscience, Lie Detection, and Beyond*, 95 CORNELL L. REV. 1191, 1192 (2010) (arguing that the admissibility of scientific evidence should depend, in part, on the quality of the evidence that would be relied on in its absence).

103. *See* Amit Pundik, *What Is Wrong with Statistical Evidence? The Attempts to Establish an Epistemic Deficiency*, 27 CIV. JUST. Q. 461 (2008) (assuming that evidence scholars rely on this distinction). Even if statistical evidence is not epistemically inferior, some types of statistical and nonstatistical evidence may differ with regard to other epistemic features. *See* Enoch, Spectre & Fisher, *supra* note 13 (distinguishing examples of statistical and nonstatistical evidence based on the "sensitivity" of the inferences it supports). There may also be moral or political reasons to distinguish between statistical and nonstatistical evidence.

104. For an illustration of this inferential process, see ALLEN ET AL., *supra* note 97, at 122–30.

(6) *Finally, with regard to these theoretical issues, there is a false dichotomy between probabilistic conceptions of proof and the Story Model*.[105] The alternative to rejecting probabilistic conceptions, however, is not an evidentiary free-for-all in which whatever persuades, works, or in which all fact-finding reduces to intuitions about narratives. The Story Model is a descriptive psychological account of jury behavior, and a particularly good one at that.[106] Any theoretical conception would be wise to take account of empirical knowledge of fact finder behavior. But whether the account is true,[107] and whether lawyers present cases in the form of narratives, are distinct issues from theoretical accounts of the nature and structure of legal proof. There is conceptual space for a *nonprobabilistic* theoretical account of legal proof that also provides normative guidance and constraint.[108]

II. A PROBABILISTIC CONCEPTION OF EVIDENCE AND PROOF

The dominant theoretical account conceives of evidence and proof in probabilistic terms. This Part first outlines the basic contours of this conception and then evaluates the conception in detail based on the three theoretical criteria articulated in Part I.

Under the probabilistic conception, factual propositions may be assessed according to their probability given the evidence. Although the details of this conception may be characterized in a variety of ways, a few features are common. First, at the micro-level, relevance and probative value are expressed as an increase or decrease in probability based on evidence. Second, at the macro-level, standards of proof are expressed as probabilistic thresholds, and a factual issue is "proven" when its probability given the evidence exceeds the threshold. Finally, the probabilistic assessments conform to the axioms or rules of probability theory.[109] The following are the most prominent of these rules:

---

105. *See* Kenworthey Bilz, *We Don't Want to Hear It: Psychology, Literature and The Narrative Model of Judging*, 2010 U. ILL. L. REV. 429, 435 (discussing these two models). *See generally* LEGAL EVIDENCE AND PROOF: STATISTICS, STORIES, AND LOGIC (Henry Kaptein et al. eds., 2009).

106. *See supra* note 17.

107. *Id*

108. *See* Griffin, *supra* note 20 (arguing that changes in evidentiary and procedural rules may improve accuracy given current empirical knowledge of jury decisionmaking).

109. *See generally* ANDREI KOLMOGOROV, FOUNDATIONS OF THE THEORY OF PROBABILITY (1956). Individual probabilities may be expressed in various ways: (1) logical relations, (2) frequencies, (3) propensities, or (4) subjective degrees of belief. For a useful overview see Alan

(1) probability assessments fall somewhere on a scale between 1 (which equals certain truth) and 0 (which equals certain falsity);

(2) the probability of a proposition and its negation equals 1;[110]

(3) the probability of two independent propositions conforms to the product rule;[111] and

(4) the probability of a proposition in light of new evidence conforms to Bayes' Theorem[112]

This Part argues that the probabilistic conception fails each of the three theoretical constraints. To be clear, my focus is on theoretical accounts of legal proof that rely on probability theory. My analysis is *not* a critique of probability theory; it a critique of particular *uses* to which probability theory is put by legal theorists.[113] At the micro-level, this conception fails to adequately explain the relevance and probative value of evidence. At the macro-level, this conception fails to explain standards of proof. This conception also

---

Hajek, *Interpretations of Probability*, STANFORD ENCYCLOPEDIA PHIL. (Dec. 19, 2011), http://plato.stanford.edu/entries/probability-interpret/. In the context of legal proof, probability assessments depend primarily on relative frequencies when the relevant data are available and subjective degrees of belief when data are not available.

110. If the probability of a proposition P (e.g., "the defendant is guilty") is 0.6, then the probability of not-P ("the defendant is not guilty") equals 0.4.

111. If the probability of P1 is 0.5 and the probability of P2 is 0.5, and the two events expressed by the propositions are independent of one another, then the probability of P1 and P2 equals 0.25.

112. Thomas Bayes, *An Essay Towards Solving a Problem in the Doctrine of Chances*, 53 PHIL. TRANS. ROYAL SOC'Y LONDON 370 (1763), *available at* http://rstl.royalsocietypublishing.org/content/53/370 . Bayes' Theorem allows one to calculate the probability of proposition in light of new evidence by multiplying (1) the "prior probability" (before receiving the evidence) by (2) the "likelihood ratio" (the probability of the evidence, given the proposition, divided by the probability of the evidence, given the negation of the proposition). Bayes' Theorem may be expressed in two different forms: an "odds" form and a "probability" form.

Suppose we want to calculate the odds (O) or probability (P) of a defendant's guilt (G) based on some evidence (E) using Bayes' Theorem. In addition to the above letters, the symbol "|" will be used to indicate "given" (as in the O(G)|E means "the odds of guilt given the evidence") and "not-" will indicate negation.

The *odds* form of the theorem is the following:

O(G)|E =

(P(E)|G / P(E)|not-G) x O(G)

The slightly more complicated *probability* form of the theorem is the following:

P(G)|E =

(P(E)|G / (P(E)|G x P(G) + P(E)|not-G x P(not-G))) x P(G)

For clear overviews of Bayes' Theorem, and how it may be derived, see IAN HACKING, INTRODUCTION TO PROBABILITY AND INDUCTIVE LOGIC (2001), and AITKEN, ROBERTS & JACKSON, *supra* note 100.

113. *Cf.* Kaplow, *supra* note 15, at 859 (discussing conventional probabilistic conceptions of legal proof, noting conceptual difficulties, and concluding that we "need to start from scratch").

fails the integration constraint because its micro- and macro-level accounts are inconsistent with one another.

A. The Micro-Level Constraint

The micro-level constraint requires that a theory account for relevance and probative value. It must explain these concepts and provide criteria for evaluating applications.

1. Relevance

Evidence is "relevant" if it "has any tendency to make a fact more or less probable than it would be without the evidence" and the fact is "of consequence in determining the action."[114] Evidence that is not relevant is not admissible.[115] The phrase "more or less probable" in the definition naturally invites the probabilistic conception (perhaps even suggesting a tautology), but this interpretation does not necessarily follow.[116]

The use of "probable" in Rule 401 of the *Federal Rules of Evidence* should not necessarily be taken to codify the probabilistic conception. First, the particular theory of mathematical probability underlying this conception is just one of several different theories or types of "probability."[117] Therefore, "probable" in the rule is ambiguous; it does not necessarily mean "probable according to the dictates of conventional probability theory." Second, the *Advisory Committee Notes* to Rule 401 make clear that the relevance definition is referring to evidence that provides *epistemic support* for material propositions more generally—that is, evidence that provides some reason in support of (or "tends to prove") whether a particular fact is true or false—rather than a particular theory of probability.[118] The

114. FED. R. EVID. 401.

115. FED. R. EVID. 402.

116. *See* Risinger, *supra* note 59 ("[R]elevance is a much more problematic concept than most people realize."). My use of "necessarily" here and in the following paragraph in the text is intended to leave open the possibility that the probabilistic conception may, at the end of the day, turn out to provide the best theoretical conception of relevance (although the analysis that follows will reject this possibility). The discussion in the text is meant to foreclose declarations of early victory for the probabilistic conception based on the language (and history) of Rule 401. For a discussion of the relationship between the language of Rule 401 and the probabilistic conception, see Park et al., *supra* note 98.

117. *See supra* note 39.

118. *See* FED. R. EVID. 401 advisory Committee notes ("Relevancy . . . exists only as a relation between an item of evidence and a matter properly provable in the case. Does the item of evidence tend to prove the matter sought to be proved?") By "epistemic support" throughout this

committee explains that whether evidence "tends to prove" a matter in the case will depend on principles from "experience," "science," and "logic," and that "[a]ny more stringent requirement is unworkable and unrealistic."[119] Moreover, the note explains that the "probability" language in the Rule was used merely to avoid "confusion between questions of admissibility and questions of sufficiency."[120] Thus, any evidence that serves the function of providing epistemic support ("tends to prove") is relevant, regardless of whether or not it conforms to the dictates of a particular theory. The probabilistic conception explains some, but not all, cases of epistemic relevance, the notion expressed in Rule 401.[121]

According to the probabilistic conception, relevance depends on whether it changes the probability of a proposition at issue. Relevance is expressed either in terms of a "likelihood ratio"[122] or in terms of whether it increases (or decreases) the probability of a proposition from its prior probability.[123] This conception is offered both as an

---

Article, I mean evidence that provides some reason to believe that a particular proposition is true or false. This support need not provide a conclusive reason and it may or may not provide epistemic justification for beliefs about the proposition. On the latter, see Michael S. Pardo, *The Gettier Problem and Legal Proof*, 16 LEGAL THEORY 37 (2010).

119. *See* FED. R. EVID. 401 advisory Committee notes ("Whether the [relevance] relationship exists depends upon principles evolved by experience or science, applied logically to the situation at hand. . . . Any more stringent requirement is unworkable and unrealistic.").

120. *See id.* ("Dealing with probability in the language of the rule has the added virtue of avoiding confusion between questions of admissibility and questions of the sufficiency of the evidence.").

121. As a general matter, the term "relevance" has two different senses. The first, *conversational* relevance, concerns whether an assertion relates to the topic of a conversation. For discussions of this sense, see PAUL GRICE, *Logic and Conversation*, *in* STUDIES IN THE WAY OF WORDS 24–27 (1989) (discussing conversational implicatures and the maxim of relevance). *See also* DAN SPERBER & DEIRDRE WILSON, RELEVANCE: COMMUNICATION AND COGNITION (1986). The second, *epistemic* relevance, concerns whether evidence supports the truth of a proposition. For a discussion of this sense, see Peter Achinstein, THE BOOK OF EVIDENCE 95–112 (2001) (discussing ways in which epistemic justification diverges from different probabilistic conceptions of evidence). The second sense is the one expressed in FED. R. EVID. 401 (stating that evidence is relevant if "it has the tendency to make a fact more or less probable than it would be without the evidence"). *See also* L. Jonathan Cohen, *Some Steps Towards a General Theory of Relevance*, 101 SYNTHESE 171, 172–82 (1994) (discussing the similar inferential structure of these two senses).

122. *See* Lempert, *supra* note 9, at 1025–27 (describing the formula for the likelihood ratio and its relation to logical relevance).

123. *See* Deborah Davis & William C. Follette, *Rethinking the Probative Value of Evidence: Base Rates, Intuitive Profiling, and the "*Post*diction" of Behavior*, 26 L. & HUM. BEHAV. 133, 136–53 (2002). Davis and Follette attempt to depict how probability can be used to show that certain testimony is irrelevant. In this particular study, they argue that "the relative increase in likelihood that an unfaithful man will murder his wife, over the likelihood that a faithful man will murder his wife is so infinitesimal . . . as to be totally insignificant." *Id.*

explanation of the meaning of Rule 401's definition of relevance[124] and as a normative standard by which to assess possible relevance.[125]

The "likelihood ratio" measures the probability of receiving evidence if a proposition is true, divided by the probability of receiving the evidence if the proposition is false.[126] Consistent with Bayes' Theorem,[127] we can multiply the likelihood ratio by the prior probability to arrive at an updated probability. Consider the following example provided by Professor Lempert in his classic article:

> *Blood Type*: If we know that the perpetrator of a crime has Type A blood, and we have evidence that the defendant has Type A blood, the evidence is relevant to proving the defendant is the perpetrator if (1) the probability of this evidence (he is Type A), given that he is the perpetrator, is greater than (2) the probability of this evidence, given that he is not the perpetrator.[128]

Lempert assumes that fifty percent of the population have the same blood type and analyzes the evidence as follows: the probability of the evidence, given that he is the perpetrator, is 1; the probability, if he is not the perpetrator, is 0.5. Dividing the first number by the second yields 2:1 odds of guilt (or 0.67 probability).[129]

As a general matter, any likelihood ratio other than 1 yields a conclusion that the evidence is relevant.[130] If the odds are greater than 1:1, then the evidence is relevant for proving the proposition at issue. If the odds are *lower* than 1:1, then the evidence is relevant for *disproving* the proposition at issue.[131] Probative value is expressed under this conception based on how much greater (or lower) than 1 the

---

124. *See* Lempert, *supra* note 9, at 1025 (describing likelihood ratio as the "mathematical equivalent" of "relevant evidence" in FED. R. EVID. 401).

125. *Id.* at 1023 ("[T]his is a normative model.").

126. *Id.* at 1023–24; *see also* AITKEN, ROBERTS & JACKSON, *supra* note 100, at 38–39, 43–50 (discussing likelihood ratios).

127. *See* Lempert, *supra* note 9, at 1022–25 (describing the application of Bayes' Theorem to determining whether a new piece of evidence would influence a "rational decision-maker's odds that a defendant is guilty").

128. *See id.* at 1023–25 (providing the *Blood Type* example and describing the likelihood ratio and its effects on relevance).

129. *Id.* at 1024–25. The 2:1 ratio derives from dividing 1 by 0.5. The odds form is converted to a probability form by dividing the numerator (2) by the total number of possibilities (3). If we start with even odds (1:1 or 0.5)—or no reason to think guilt or non-guilt is more likely—then when we multiply the likelihood ratio by the prior odds we will get the same result: 2:1. We can then use this new number as the "prior odds" and multiply it by the likelihood ratio for the next item of evidence, and so on. *Id.* at 1023–24.

130. *Id.* at 1025.

131. The odds will be lower when the probability of the evidence, assuming the proposition at issue is true, is lower than the probability assuming the proposition is false.

likelihood ratio is.[132] A likelihood ratio of 1:1 means the evidence is irrelevant. Consider the following example of irrelevant evidence: *Liberal Candidates*: "In a criminal trial, there is evidence that the defendant supports liberal political candidates."[133]

As Lempert explains:

> Absent *some reason to believe* that liberals are more prone to commit the crime in question, the probability that the defendant could have been shown to be liberal were he guilty is . . . the same as the probability that he could have been shown to be a liberal were he not guilty.[134]

The odds of guilt remain what they were before considering the evidence. Thus, under this conception, the evidence is irrelevant.

Likelihood ratios, to be sure, capture something intuitive and important about relevance.[135] But likelihood ratios do not adequately explain the meaning of relevance[136] or provide the necessary guidance and constraint to serve as a normative theory.[137] This is so because evidence may be relevant even when there is no "reason to believe" anything at all about the likelihood ratio, whether it is 1:1 or not. Moreover, in some cases there will be good reasons to believe evidence is relevant when the likelihood ratio implies that it is irrelevant.

---

132. *See* Kaye, *supra* note 9, at 684 (stating that the greater the amount that the likelihood ratio departs "from unity, the greater the probative value of the data").

133. Lempert, *supra* note 9, at 1025.

134. *Id.* (emphasis added).

135. The likelihood ratio also draws attention to the important issue of how likely the evidence would exist even if the proposition for which it is offered were false. *See id.* at 1023–24 (providing the formula for the likelihood ratio).

136. Claims about the "meaning of relevance" purport to be analytically true claims. *See generally* H.P. Grice & P.F. Strawson, *In Defense of a Dogma*, 65 PHIL. REV. 141 (1956) (discussing analytic claims). Therefore, a single counterexample in which relevance and likelihood ratios diverge is sufficient to demonstrate that likelihood ratios do not capture the meaning of relevance. For doubts about analytic claims generally, the classic article is W.V. Quine, *Two Dogmas of Empiricism*, *in* FROM A LOGICAL POINT OF VIEW 20 (1953).

137. In pointing to examples of relevance that do not fit the likelihood-ratio theory, I do not mean to suggest *merely* that there are other aspects to relevance (e.g., that some illustrative or demonstrative evidence may be relevant only to the extent it helps jurors understand other evidence and has no independent relevance, general background details, or *res gestae*). Moreover, if evidence is not relevant, it should not be admitted over a relevance objection. *See* FED. R. EVID. 402 ("Irrelevant evidence is not admissible."). It is sometimes assumed that *Old Chief v. United States*, 519 U.S. 172 (1997), expanded the scope of relevance to include anything that makes a story persuasive. But this is not right. *Old Chief* was about probative value and the substitution of one piece of relevant evidence for another piece of relevant evidence because of narrative considerations. *See Old Chief*, 519 U.S. at 174, 184 (holding that in a trial for a crime that requires a previous conviction "a district court abuses its discretion if it spurns . . . an offer [to concede the fact of a prior conviction] and admits the full record of a prior judgment, when the name or nature of the prior offense raises the risk of a verdict tainted by improper considerations, and when the purpose of the evidence is solely to prove the element of prior conviction").

To see how relevance may diverge from the likelihood-ratio view, begin with a simple observation about actual trials. In any given trial, there may be evidence that each side claims supports its theory (or story) of what happened, but that evidence may not distinguish between the cases. In other words, the evidence will form an integral part of each side's proof, but there may be no reason to believe that the likelihood ratio for that evidence is anything other than 1. Like the *Liberal Candidates* example, any conclusion to the contrary would be sheer speculation. Unlike that example, however, the evidence may still be relevant *precisely because it forms an integral part of each side's explanation as to what happened*. If any such evidence exists, then it provides a counterexample to the likelihood-ratio view as either a theory of the meaning of relevance[138] or as a normative theory.[139]

Here is a detailed example of such evidence from a criminal trial.[140] The defendant, an inmate at a maximum-security prison, was charged with two counts of battery on prison guards.[141] The charges arose from an altercation between the defendant and guards after the defendant refused to return a food tray in his cell. The prosecution's theory was that the defendant battered the officers when they opened the cell door to retrieve the tray.[142] The defendant testified that one of the guards rushed in and began hitting him first, and his attorney argued that, even if the defendant made contact first with the officer, the defendant was acting in self-defense.[143]

In sorting out what happened, a reasonable juror might like to know what precipitated the altercation. Why did the defendant withhold the tray in the first place? Both sides provided a number of details on this point. The attorneys discussed during opening and closing arguments that the defendant had not received a package sent to him by his family, and that after several weeks and several attempts to speak with a sergeant about it, the defendant refused to

138. *See supra* note 136 (stating that divergence between relevance and likelihood ratios in a single counterexample would show that relevance is not captured fully by likelihood ratios).

139. If the evidence is relevant, and the likelihood ratio implies that it is irrelevant and ought to be excluded, then the likelihood ratio fails as a normative standard by which to gauge relevance.

140. The trial transcript is printed in ALLEN ET AL., *supra* note 97, at 1–77.

141. *Id.* at 2–3.

142. *Id.* at 5–7.

143. *Id.* at 39–40, 53.

return his food tray.[144] Multiple witnesses for each side testified to details about the package and the mail procedures at the prison.[145] Each side used this evidence to support its competing theory: (1) the defendant was frustrated and angry about not receiving the package, withheld his tray, and charged the guard,[146] and (2) the defendant was frustrated about not receiving the package, withheld the tray to get a sergeant's attention about the matter, and in response the guards attacked him (to retaliate or punish him for this behavior).[147] The evidence does not appear to distinguish between the two theories; as with the *Liberal Candidates* example above,[148] there is no reason to believe that this evidence supports one theory over the other. In other words, the likelihood ratio is 1:1. Under the likelihood-ratio theory, the evidence is irrelevant (and *a fortiori* has no probative value), and, thus, should have been excluded.[149] If this conception were true, the jury would have been forced to choose between who started the altercation without giving them any information about the events that precipitated it. The apparent relevance of this evidence—despite its

144. From the Prosecution's Opening Statement: "For whatever reason, inmate Johnson, the defendant in this case, was standing in his cell with the trays in his hand and he had some discussion with the officers about a package. He wanted some package." *Id.* at 6.

From the Defense's Opening Statement: "[T]he defendant was notified that a package had arrived—that he was aware of the fact somehow or other that a package had been sent to him by his family. A substantial period of time goes by. . . . Finally, on the 28th he withholds the tray, which I agree is disobedient." *Id.* at 39.

From the Prosecution's Closing Argument: "[F]or some reason inmate Johnson chose to draw the line that day over some package that he claimed he wasn't getting." *Id.* at 66.

From the Defense's Closing Argument: "Johnson chose to draw the line. I agree. He drew the line, 'I want something done; I want to see a sergeant . . .' And it's obvious from the testimony that if he just kept on drawing the line at some point he would have seen a sergeant. That doesn't mean Johnson made a choice to go out and start battering the officer or officers." *Id.* at 71.

145. *See, e.g.*, *id.* at 34 (testimony from prison guard), 42 (testimony from cellmate), 52 (testimony from defendant); *see also id.* at 42 (overruling relevance objection regarding package procedures).

146. *Id.* at 66.

147. *Id.* at 71.

148. *See supra* note 133 and accompanying text (describing the *Liberal Candidates* scenario). Lempert also notes that "the court is justified in excluding evidence on the ground of logical irrelevance" when "it would be unreasonable for a jury to find the likelihood ratio to be slightly different from one." Lempert, *supra* note 9, at 1026 n.21. In both *Liberal Candidates* and *Johnson*, the jury has no reasonable basis for concluding the likelihood ratio is slightly different from 1. Both examples would require the same types of unsupported speculation to conclude otherwise. This is not to deny, however, that the evidence in both examples might be "conditionally relevant," or relevant if other facts are known or are supported by evidence (e.g., the relative frequency of batteries by inmates on guards versus by guards on inmates at the prison in *Johnson*). Conditional relevance is discussed *supra* note 88.

149. *See generally* FED. R. EVID. 402 ("Irrelevant evidence is not admissible.").

likelihood ratio of 1—provides a *reductio ad absurdum* for the theory. The evidence is relevant because it provides epistemic support for each side's explanation of what happened.[150] The fact that it supports *both* explanations does not make it irrelevant. The existence of this type of "overlapping" relevant evidence is problematic for the likelihood-ratio theory.[151]

The analytical tensions between relevance and the probabilistic conception can be brought into focus with a few more stylized examples.

> *Twins*: A witness will testify that someone matching the defendant's description was seen fleeing a crime scene. The defendant claims that it was his identical twin and introduces evidence establishing the twin's existence. Suppose there is no reason to believe the testimony distinguishes the defendant from his twin.

If we are comparing the likelihood of the defendant's guilt versus his twin, then, as with *Liberal Candidates* and *Johnson*, there does not appear to be any reason to think the likelihood ratio is

---

150. *See supra* note 118 and accompanying text (discussing epistemic support for relevancy). One way it supports each explanation is by providing evidence that each side must explain in order for the explanations to be plausible; explanations that could not account for the overlapping evidence would fail. The evidence, in other words, provides a "test" for possible explanations. For a similar idea in the philosophy of science, see Deborah G. Mayo, *Evidence as Passing Severe Tests: Highly Probable vs. Highly Probed Hypotheses*, *in* SCIENTIFIC EVIDENCE: PHILOSOPHICAL THEORIES AND APPLICATIONS 95–127 (Peter Achinstein ed., 2005). A similar phenomenon arises with questions of law: if two sides to a legal dispute can each explain why a previous precedent supports their answer, this does not render the prior precedent irrelevant. Legal questions have ineluctable factual aspects and so the issue of "overlapping" evidence arises in this context just as it does in the context of legal proof. *See* Ronald J. Allen & Michael S. Pardo, *The Myth of the Law-Fact Distinction*, 97 NW. U. L. REV. 1769 (2003). Keeping "overlapping" evidence from juries is likely to detract from rather than foster accuracy. This is, of course, an empirical claim that might turn out to be false in certain cases. But the claim is consistent with the general principle in evidence law that more information about cases will lead to more accurate determinations, unless there is a compelling countervailing reason to keep that information from decisionmakers. *See supra* notes 88–94 and accompanying text. It is also intuitively plausible. Consider two analogies. Suppose trying to determine which of two pictures is more accurate by seeing only the differences between them rather than the parts they each share; or imagine building two chairs with only the parts that differ between them rather than the parts they are likely to share (legs, a seat, a back, etc.). Whether it is pictures, chairs, or explanations at trial, what is shared may be as important as what is distinct, even when the judgments that follow concern comparative assessments. If legal proof involves judgments about competing explanations as a whole, as Part III argues *infra*, then it is not surprising that evidence supporting parts of those explanations would be relevant even when the evidence supports more than one explanation and does not distinguish them. Two different chairs may be built with the exact same type of seat, but this does not mean that information about the seat is irrelevant in assessing which is a better chair.

151. The prevalence of this type of evidence is an empirical question, but most trials will have at least a few examples. The prevalence is not relevant for purposes of this Article; so long as at least one example exists, that is sufficient to falsify the theory as either the meaning of relevance or as a normative theory. *See supra* notes 136–39.

different from 1. Nevertheless, the evidence is relevant. It provides some epistemic support for both possibilities (i.e., the defendant did it or the twin did it),[152] and it "tends to prove" each side's explanation of what happened. The fact that it supports *both* explanations (as in *Johnson*) does not make it irrelevant. Of course, the probabilist has a rejoinder as to why the evidence is also relevant under a probabilistic interpretation: namely, it eliminates everybody except the defendant and his twin, and by eliminating everyone else it thereby increases the probability the defendant is guilty. The rejoinder is correct—but notice the tension between this conclusion and the implications of the likelihood-ratio view. Although the evidence is relevant because it eliminates all other suspects, it technically fails to fit the likelihood-ratio conception as soon as evidence about the twin is introduced. As soon as the twin evidence is introduced, the probability of the evidence, given the defendant's guilt, is exactly the same as the probability of the evidence, given the defendant's nonguilt (assuming this is equivalent to the probability of the twin's guilt). If that is so, then under this interpretation the likelihood ratio implies that the witness's testimony should be excluded as irrelevant.[153]

Moreover, evidence may be epistemically relevant for proving a proposition even when it does not appear to increase the probability of that proposition. Consider the following example:

> *Poison*: The Prosecution alleges that Victim died of poisoning, and Defendant contends that Victim died from some other cause. There is evidence that at 12:00 p.m. on the day he collapsed and died, Victim's lunch contained a poison that is fatal for ninety percent of the people who ingest it. Suppose there is also evidence that at 12:30 p.m., Victim ingested a second poison concealed in a drink that completely counteracts the first poison; however, it is fatal for eighty percent of the people who ingest it.[154]

Is evidence of the second poison relevant for proving that Victim died of poisoning? Yes, of course. Articulating exactly *why*, however, is critical for understanding the potential analytic gap between epistemic relevance and probability. The evidence supports the Prosecution's theory that Victim died of poisoning and is relevant *for that purpose*, even if the evidence about the first poison has been

---

152. By contrast, in *Liberal Candidates*, the evidence provides no epistemic support for the defendant's guilt.

153. *But see supra* notes 150–51 and accompanying text.

154. This example and the next one (*Lottery*) are based on similar examples provided by Peter Achinstein, writing within the philosophy of science. *See* ACHINSTEIN, *supra* note 121, at 69–71 (discussing the examples entitled "intervening cause counterexample" and "first lottery counterexample"). Despite some important differences between the scientific and legal contexts, Achinstein similarly concludes that his examples show that an increase in probability is neither necessary nor sufficient to constitute evidence for a proposition. *Id.* at 71.

admitted. Notice, however, that the evidence regarding the second poison actually *lowers* the probability that Victim died of poisoning from the prior probability without the evidence.

A defender of the probabilistic conception has two possible rejoinders to account for the relevance of this evidence. First, because the evidence lowers the probability, it is also relevant for *disproving* that Victim died of poisoning. Yes, perhaps[155]—but this, by hypothesis, was not the Prosecutor's theory of relevance for seeking to admit the evidence. A party's theory of relevance (when faced with a relevance objection) commits it to a particular reason why the evidence is relevant (i.e., what it tends to prove).[156] In this example, the Prosecution is seeking to admit the evidence for a particular purpose, not to support an alternative theory implied by the probabilistic conception. Alternatively, the probabilist defender may also attempt to recharacterize the example so that it supports the Prosecution's theory while also resulting in an increase in probability. For example, we might separate the two effects of the second poison as two distinct pieces of evidence: counteracting the first poison and causing death. Under this reinterpretation, the first piece of evidence lowers the probability to zero percent and, then, the second piece of evidence raises the probability to 0.8, thus making the evidence relevant and raising the probability. This type of ad hoc recharacterization suggests

155. Parties offering evidence will have to make an offer of proof in response to a relevance objection. See FED. R. EVID. 103; JACK B. WEINSTEIN & MARGARET BERGER, WEINSTEIN'S FEDERAL EVIDENCE § 130.20[4] (2d ed. 1997) (explaining that offers of proof under Rule 103 are "meant to give the judge contemporaneous knowledge about the proposed evidence, so that the judge can make a proper ruling on the evidence at the time it is proffered"). Accordingly, except in extraordinary circumstances, courts will not entertain theories of relevance not offered by the proffering party. *See* United States v. Boulware, 558 F.3d 971, 975–76 (9th Cir. 2009); Unites States v. Lara-Hernandez, 588 F.2d 272, 274 (9th Cir. 1978) ("The presentation of additional evidentiary theories [of relevance] for the first time after the jury has returned its verdict does not comport with the salutary purpose of the timeliness requirement to allow the trial judge to make an informed ruling based on the issues *as framed by the parties*." (emphasis added)) Parties may, of course, be mistaken about whether and why evidence is relevant. The example in the text illustrates, however, that relevance determinations will (1) involve a particular argument about *why* the evidence is relevant, and (2) the evaluation of that argument depends on more than the likelihood ratio, including whether it supports one of the explanations at trial. This theory of relevance reveals an additional disjunct between the probabilistic conception and relevance in actual criminal trials—it seems unlikely that a criminal defendant would introduce evidence of the second poison in order to prove Victim did not die of poisoning.

156. Parties may also introduce evidence that supports the other side's case to "take the sting out" of potentially damaging evidence. Nothing in the above analysis is inconsistent with this possibility. Evidence does not become irrelevant because it does not support the party offering it. Rather, parties are "taking the sting out" precisely because it supports the other side's case. If the evidence in *Poison* supports Defendant's theory, then the Prosecution could also offer the evidence for this reason.

that there may indeed be creative ways to make the probabilistic conception fit with epistemic relevance.[157] But this also reveals that judgments about epistemic relevance are doing the real work justifying particular inferences; the ad hoc probabilistic descriptions (which can be articulated in a variety of different ways[158]) are epiphenomenal.

Furthermore, evidence may not be epistemically relevant for proving a proposition even when it does appear to raise the probability of that proposition. Consider the following example:

> *Lottery*: Victim was murdered. The motive appears to be that Victim ran an illegal lottery and refused to pay the winner. It is unknown who actually won the lottery. The Prosecution claims it was Defendant, and Defendant claims it was Rival. Defendant purchased one of the one thousand total lottery tickets and Rival purchased ninety-nine tickets. Suppose there is also evidence that the other nine hundred tickets were never sold and have been accounted for.[159]

Is evidence about the other nine hundred tickets relevant for proving Defendant's guilt? Is it relevant for proving Defendant's innocence? Both? Neither? Here, the probability that Defendant was the winner of the lottery (and thus is guilty) went from 1 in 1000 to 1 in 100 with the introduction of the evidence about the nine hundred tickets. Despite this increase in probability, the evidence does not provide epistemic support for proving Defendant's guilt.[160] Indeed, the evidence better supports Defendant's theory that Rival is the culprit.[161] As with the other examples, a defender of the probabilistic conception may respond by either offering alternative theories of relevance (e.g., it is relevant for proving guilt because it eliminates nine hundred other possibilities) or offering a recharacterization consistent with the analysis (e.g., it increases the probability of Rival's guilt and, hence, Defendant's innocence), but the fact remains that epistemic support undergirds relevance and justifies evidentiary inferences. Probabilistic judgments may diverge from this support, or they may be adjusted in an ad hoc fashion to conform to this support.

---

157. *See generally supra* note 121 and accompanying text.

158. Is the second poisoning one event or two (or twenty or a hundred)? Is it one piece of evidence, or two, or twenty, or one hundred, and so on? Any piece of evidence can be described in more and more fine-grained ways such that it becomes several items of evidence.

159. *See* ACHINSTEIN, *supra* note 121, at 69–70 (setting forth the "first lottery counterexample"). Although the evidence in this example is statistical in form, nothing in the analysis necessarily turns on this feature. *See supra* note 103 and accompanying text.

160. *See supra* notes 118–19 and accompanying text.

161. Perhaps one might argue the evidence is irrelevant because it keeps the same relative ratio of tickets between Defendant and Rival. But either side would likely try to take advantage of evidence of the unknown nine hundred tickets, which may be another example of overlapping evidence. *See supra* notes 150–51 and accompanying text.

The inadequacy of the probabilistic conception to account fully for epistemic relevance derives from the fact that the category of epistemic support is a general one, which probabilistic inferences may or may not fit. Jonathan Cohen has analyzed this potential for divergence between relevance and probability in a previous article. As Cohen explains: "[Relevance] is a relation—typically appreciated or pointed out in the course of a reasoned enquiry, debate, conversation, meditation, explanation, or justification—between a true proposition and an askable question."[162]

And, "[a]nything that can sanction a reason, even if an incomplete or inconclusive reason, for accepting a particular type of proposition as a correct answer, or for rejecting it as an incorrect answer, to an askable type of question can count as a criterion of relevance."[163]

In the trial context, the askable question is typically "what happened?" (or other questions that relate to whether the elements of the cause of action, crime, or affirmative defense have been proven), and any evidence that provides a reason for accepting an explanation as a correct answer to that question is relevant. A change in probability is "just one criterion of relevance alongside the others,"[164] and the misconception is to assume it is the *sine qua non* or "lies at the heart of the matter."[165]

The examples and the analysis of relevance do not, to be sure, present a knockdown refutation of the probabilistic conception. Nevertheless, they do reveal the primacy of epistemic support as the foundation of relevance,[166] and they raise doubts about the ability of the probabilistic conception to provide an adequate theory of relevance. This is troubling for the probabilistic conception because relevance is the one evidentiary issue on which the probabilistic

---

162. Cohen, *supra* note 121, at 178. Cohen provides the following schematic outline for nonconversational, epistemic relevance:

> [A] true proposition R is nonconversationally relevant to an askable question Q if and only if there is a proposition A such that the truth of R is or would be some reason, though not necessarily a complete or conclusive reason, for anyone's accepting or rejecting A as an answer to Q, where this is warranted by some nonconversational principle.

Id.; *see also* Susan Haack, *The Embedded Epistemologist: Dispatches from the Legal Front*, 25 RATIO JURIS 206, 215–18 (2012) (distinguishing epistemic warrant from standard mathematic probabilities in the context of legal proof).

163. Cohen, *supra* note 121, at 181.

164. *Id.*

165. *Id.* at 180.

166. *See supra* note 118 and accompanying text (discussing relevance and epistemic support).

conception appears to be on the firmest ground.[167] As we will see, however, the troubles with relevance are just the beginning, and are by no means the most serious problems, for this conception. The problems with explaining probative value, and the macro-level and integration constraints even more so, will turn the doubts raised thus far into warranted conclusions about the implausibility of this conception as a general theory of evidence and proof.

2. Probative Value

Probative value refers to the strength of the evidence in proving or disproving a fact of consequence. This value depends on the logical or empirical relationships between evidence and the propositions for which it is offered, along with other contextual factors such as a party's need for evidence, the importance of the issue, and the extent to which it is contested by the other side.[168] Evidence may be excluded if its probative value is substantially outweighed by one or more countervailing considerations.[169] As with relevance, likelihood ratios are offered to explain probative value.[170]

Attempts to quantify probative value using likelihood ratios, however, run into the problem of reference classes.[171] Quantifying probative value involves placing the evidence in a particular class (for which data are available). Consider this example from the McCormick Treatise on Evidence concerning "a behavioral pattern said to be characteristic of abused children":

> *Abuse*: If research established that the behavior is equally common among abused and nonabused children, then its likelihood ratio would be 1, and evidence of that pattern would not be probative of abuse . . . And if it were a thousand times more common among abused children, its probative value would be far greater.[172]

---

167. This is because FED. R. EVID. 401 appears to provide some textual support for this conception of relevance, *but see supra* notes 116–20 and accompanying text, and because relevance is thought to be synonymous with a change in probability. *See* Lempert, *supra* note 9, at 1025–26.

168. *Supra* note 64 and accompanying text. *See generally* FED. R. EVID. 403.

169. FED. R. EVID. 403. The considerations include unfair prejudice, confusing or misleading the jury, and time or resource considerations. *Id.*

170. *See* Kaye, *supra* note 9, at 684 ("A convenient measure of probative value is the ratio of the likelihoods for the two hypotheses."); Jonathan J. Koehler, *When Do Courts Think Base Rate Statistics Are Relevant?*, 42 JURIMETRICS J. 373, 375 (2002) ("The ratio of these likelihoods (known as the likelihood ratio) captures the probative value of the evidence.").

171. *See* Allen & Pardo, *supra* note 10, at 113–14 (explaining that evidence has a multitude of possible likelihood ratios depending on which reference class is used).

172. MCCORMICK ON EVIDENCE 277 (Strong ed., 5th ed. 1999).

These conclusions depend on a number of further assumptions (which may or may not be true). Even if the behavior is equally common among both groups of children, it might nevertheless be highly probative in a given case if, for example, abused children exhibiting this behavior also possess, and nonabused children lack, an additional characteristic and the particular child at issue possesses (or lacks) this characteristic. Similarly, even if the behavior is one thousand times more likely in abused children, the probative value may nevertheless be minimal if the child possesses (or lacks) an additional characteristic that places the child in the group of nonabused children who exhibit the behavior. In these examples, the evidence and the likelihood ratios remain constant, but the probative value may vary dramatically. *A fortiori*, the probative value is not the likelihood ratio.

This is a general problem affecting any type of legal evidence. A previous article[173] demonstrates its prevalence with regard to such diverse issues as drug-smuggling data,[174] carpet fibers,[175] marital infidelity,[176] and DNA random-match probabilities.[177] For purposes of this Article, the reference-class problem provides another reason why the probabilistic conception fails the micro-level constraint.

This theoretical failure can be illustrated with another example that focuses on practical issues raised by statistical evidence—the famous *Blue Cab* example by Amos Tversky and Daniel Kahneman.[178] Although they were not purporting to offer a theory of probative value in law, their discussion is nevertheless instructive.

> *Cab*: A cab was involved in a hit-and-run accident, and two cab companies, Green and Blue, operate in the city. The evidence: eighty-five percent of the cabs in the city are

173. Allen & Pardo, *supra* note 10, at 116–23.

174. *See* United States v. Shonubi, 103 F.3d 1085, 1090–92 (2d Cir. 1997) (holding that statistical and economic analysis relating to drug trafficking generally did not qualify as "specific evidence" of drug quantities that the defendant allegedly smuggled); United States v. Shonubi, 895 F. Supp. 460, 501–24 (E.D.N.Y. 1995) (evaluating the admissibility of probabilistic evidence at sentencing and holding that a combination of statistical and nonstatistical evidence was sufficient to satisfy the government's burden of proof).

175. *See* Michael O. Finkelstein & Bruce Levin, *On the Probative Value of Evidence from a Screening Search*, 43 JURIMETRICS. J. 265, 266–70 (2003) (discussing the evidentiary challenges of a community screening search based on a hypothetical carpet fiber found at a crime scene).

176. *See* Davis & Follette, *supra* note 123, at 137–39 (using base rates to demonstrate that the fact of infidelity is not probative of whether a man murdered his wife).

177. *See* Nance & Morris, *supra* note 8, at 409, 419 (noting that jurors may not appropriately evaluate the probative value of DNA evidence when given random-match probabilities).

178. Amos Tversky & Daniel Kahneman, *Evidential Impact of Base Rates*, *in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 153, 156–60 (Daniel Kahneman et al. eds., 1982).

> Green and fifteen percent are Blue. What is the probability the cab involved in the accident was Blue?[179]

The example is famous for illustrating "base-rate neglect."[180] Decisionmakers typically ignore background base-rate statistics in estimating probability, and this may lead to cognitive errors. For example, if a witness says the cab was Blue, test subjects will credit this testimony and ignore the distribution statistics (even when subjects are told that the witness is only eighty percent likely to be accurate).[181] Tversky and Kahneman's path-breaking work drew attention to this general cognitive issue. Base-rate neglect is no doubt an important issue in the psychology of decisionmaking.[182] It is also a mistake, however, to try to read off or quantify probative value from any given reference class for which data are available—in other words, to assume that the probative value just *is* the relevant statistics (0.85 and 0.15, in the example).[183] Why? To quantify probative value based on these statistics would be a mistake because the distribution of cabs *in the city* is just one class to which this event (this accident and this cab) is a member. The event and the evidence are members of a virtually infinite number of sets, each with differing rates. For example, suppose that eighty-five percent of the cabs *on the street* are Blue; eighty-five percent of the cabs running *at that time of day* are Green; eighty-five percent of the cabs *in the service area* (which includes several towns) are Blue, and so on.[184] Inferences from a particular reference class involve a host of background assumptions about the appropriateness and homogeneity of that particular class, the likely distribution of other classes, and several other possible issues. In sum, the data from one reference class cannot by themselves serve as either a normative standard by which to quantify probative value or as a theoretical explanation of the concept.[185]

---

179. *Id.* at 156–57.

180. *Id.* at 157–58.

181. *Id.* at 157–59.

182. For a recent discussion, see DANIEL KAHNEMAN, THINKING, FAST AND SLOW 166–74 (2011) (using the *Blue Cab* example and other experiments to explain base-rate neglect).

183. *See* Koehler, *supra* note 18, at 1300 n.69 ("Many scholars agree that the probative value of statistical evidence can be represented in terms of a likelihood ratio.").

184. It is sometime suggested that we should rely on the smallest reference class available. Notice, however, the smallest class is the event itself, which will have a probability of 1 or 0. Beyond that, what matters is homogeneity among the class, not its size. In other words, a larger class may be more reliable than a smaller class; without more information about the items in the class, we cannot know one way or the other. These second-order doubts affect the probative value of evidence.

185. This is not to suggest that we cannot make judgments about the value of evidence from different reference classes. In a recent article, for example, Professor Edward Cheng

B. The Macro-Level Constraint

The macro-level constraint requires that a theory ought to explain when evidence is sufficient to satisfy standards of proof. A successful theory will be able to explain these outcomes in light of accuracy and allocation of the risk of error. It will also provide criteria for guiding and evaluating applications of the standards.

A probabilistic conception of standards of proof is commonplace. Scholars typically posit or presuppose that preponderance of the evidence equals proof greater than 0.5; that clear and convincing evidence equals greater than 0.75 or so; and that beyond a reasonable doubt equals greater than 0.9 or higher.[186] Under this conception, evidence is sufficient when the probability of the proposition at issue exceeds the decision threshold. This conception is relied on by both descriptive and normative theories, and, although it captures some important aspects of proof standards,[187] it is deeply problematic. The problems for this conception concern considerations of both accuracy and allocating the risk of error.

---

instructively explains how selection criteria for choosing among statistical models may provide a practical way to evaluate legal evidence when the parties in a case each present statistical evidence from different reference classes. Cheng, *Practical Solution*, *supra* note 100, at 2093–2101. The model selection criteria—he discusses primarily Akaike's Information Criterion ("AIC"), Hirotugu Akaike, *A New Look at Statistical Model Identification*, 19 IEEE TRANSACTIONS ON AUTOMATIC CONTROL 716 (1974)—provide a method for choosing among dimensions of fit and complexity. *See also* Malcolm Forster & Elliott Sober, *How to Tell When Simpler, More Unified, or Less Ad Hoc Theories Will Provide More Accurate Predictions*, 45 BRIT. J. FOR PHIL. SCI. 1, 2–11 (1994) (discussing the basic premises of Akaike's theory); James Franklin, *Feature Selection Methods for Solving the Reference Class Problem: Comment on Edward K. Cheng, "A Practical Solution to the Reference Class Problem,"* 110 COLUM. L. REV. SIDEBAR 12, 18–22 (2010) (providing a brief overview of statistical modeling and suggesting that feature selection methods provide a better solution to the reference-class problem). Cheng's interesting and sophisticated proposal, however, applies only to a limited domain of cases where: (1) both sides offer statistical evidence; (2) the evidence depends on different reference classes; and (3) the evidence is the primary evidence for each side on a contested material issue in the case. More important for purposes of this Article, Cheng does not purport to offer a theory of, or purport to quantify, probative value in this limited domain of cases, or more generally. As discussed in Part III *infra*, the value of statistical evidence depends on explanatory criteria. The reference-class problem is avoided under the explanatory conception because explanatory criteria determine the value of evidence from different classes and probative value does not need to be quantified.

186. *See supra* note 12.

187. Most importantly, it can be used to illustrate formally, given a number of assumptions, the distribution of factual errors likely to follow from different standards of proof. *See, e.g.*, Hamer, *supra* note 12, at 87–96 (proposing a mathematical function to estimate the expected conviction error rate for a given standard of proof); Kaplow, *supra* note 15, at 757–72, 793, 805, 821–22 (discussing the interaction between standards of proof, conviction errors, and levels of sanctions).

1. Accuracy

Probabilistic theories fail to provide guidance in applying the standards to achieve accuracy goals. They also lack criteria by which to determine which applications are reasonable. To apply a probabilistic standard, one must also quantify the evidence.[188] Quantification could be done based on either "objective" data[189] or "subjective" beliefs.[190] Objective data (relative frequencies) do not exist for most items of evidence in any given trial.[191] Thus, subjective probability assessments provide the only plausible probabilistic option. Unfortunately, subjective standards provide little guidance on how to proceed. Suppose you are on a jury in a civil case applying the preponderance standard, and you are trying to decide, after hearing a number of witnesses, whether the probability of a proposition exceeds 0.5. How do you do this? Close your eyes and think of a number? Suppose you conclude it is 0.6 likely and your fellow juror says it is 0.2 likely. Are both reasonable? Is 0.99? Is 0.01? How can we tell? Subjective assessments could be anything at all, and there is simply no reason to think they will be truth conducive. [192]

A subjective conception also fails to provide a basis for distinguishing reasonable from unreasonable applications. This limitation affects not only applications by jurors at trial. It has profound doctrinal implications. In civil cases, the standards for both summary judgment and judgment as a matter of law depend on whether a reasonable jury could find for the nonmoving party.[193] What is reasonable or not depends on the burden of proof and the applicable decision standard. Thus, decisions to grant or deny either motion

188. This could be done precisely or with vague judgments about whether evidence surpasses a threshold.

189. But note that objective data would still be subject to the reference-class problem. *See supra* notes 171–85 and accompanying text.

190. *See* Hajek, *supra* note 109, at 3.3 (explaining probability when conceived of as a rational agent's degree of belief).

191. *See* Kaplow, *supra* note 15, at 773–76 (discussing the practical difficulty of obtaining relative-frequency evidence). Even if the data existed, the computational complexity may be too great for most trials. *See* Ronald J. Allen, *Rationality and the Taming of Complexity*, 62 ALA. L. REV. 1047, 1054–55 (2011) (discussing near-infinite amount of complex factors in juridical fact-finding); Callen, *supra* note 8, at 4.

192. The point is not just that jurors may disagree about exact numbers or thresholds; it is that there could be any relationship at all (including a random one) between subjective beliefs and truth. *See* LAUDAN, *supra* note 14, at 51–62 (arguing that standards of proof should instruct jurors on *what to believe based on the evidence*, not instruct them to consult their existing subjective beliefs after having heard the evidence).

193. *See supra* note 25.

presuppose some conception of the proof standards and what they require. Whether parties will get to trial in the first place, whether the case will go to a jury, and whether a jury verdict will be overturned, all depend on principled applications of standards of proof. But subjective probabilistic standards provide no guidance. Suppose you are a judge deciding a summary judgment motion and trying to determine whether a reasonable jury could find for the plaintiff under a decision rule of "greater than 0.5 according to their subjective beliefs." First, there is simply no way to know what any actual jury would do without letting it decide. More importantly, because the criteria are subjective, there is no independent basis for concluding which applications are unreasonable.

This same limitation arises in criminal cases under the beyond-a-reasonable-doubt standard. Challenges by defendants to the sufficiency of the evidence before, during, and after trial depend on whether a reasonable jury could find guilt beyond a reasonable doubt.[194] Thus, whether criminal defendants will go to trial, whether cases will go to a jury, and whether convictions will be overturned depend on principled applications of the beyond-a-reasonable-doubt standard. Subjective standards, however, provide no guidance or criteria for principled applications.

2. Risk of Error

Probabilistic conceptions suffer from additional limitations regarding the risk of error. These standards are expressed as *noncomparative* in a way that would frustrate, rather than explain, standards of proof in light of their goals in allocating the risk of error.[195] Conventional probabilistic standards compare the probability of a proposition *with its negation*, but they should be comparing the competing propositions put forward by each side. Suppose the preponderance rule is, as is commonly supposed, trying to minimize errors and allocate the risk of error roughly evenly among plaintiffs and defendants.[196] Now, suppose a plaintiff offers a story that a reasonable jury concludes is 0.4 likely and the defendant offers a story

194. *See, e.g.*, Langston v. Smith, 630 F.3d 310, 316–20 (2d Cir. 2011) (holding that evidence was constitutionally insufficient to support conviction for felony assault); O'Laughlin v. O'Brien, 568 F.3d 287, 302–04 (1st Cir. 2009) (holding that circumstantial evidence identifying defendant as perpetrator of burglary and assault was insufficient to support conviction); *supra* note 25.

195. A similar problem would arise for noncomparative standards designed to achieve other goals. *See, e.g.*, Kaplow, *supra* note 15, at 752–72 (arguing that proof standards should create an optimal deterrence/chilling effect).

196. *See supra* notes 74, 76.

that the jury concludes is 0.2 likely.[197] The "greater than 0.5" standard implies that the plaintiff should lose—even though the plaintiff's account is twice as likely to be true as the defendant's alternative account.[198] This frustrates the goal of equalizing the risk of error; plaintiffs should not bear the risk of error for all of the unknown probability space.[199] The mistake is to assume that any unknown possibilities favor the defendant (or the party without the burden of proof). This is inconsistent with *equalizing* the risk of error.

This noncomparative conception may be compounded by two additional issues, one empirical and the other theoretical. First, this noncomparative aspect may be exacerbated if, as an empirical matter, jurors employ a probabilistic threshold higher than the conventional ones. In a recent study, for example, Eyal Zamir and Ilana Ritov report that subjects (lawyers and students) employed a probabilistic threshold of approximately 0.7 in applying the preponderance standard.[200] Employing this threshold shifts more of the risk of error onto plaintiffs (or parties with the burden of proof) and is inconsistent with equalizing the risk of error. The second issue concerns theoretical puzzles generated by the so-called "conjunction problem."[201] Because the standard of proof applies to individual elements rather than cases as a whole, this creates a variety of formal problems when conceived

197. There are other problems with this conception. For example, to truly prove a case beyond 0.5 requires a plaintiff to disprove all the other ways the world could have been from what the plaintiff alleges. Disproving all other possibilities, however, is an impossible task. *See* DOUGLAS WALTON, ABDUCTIVE REASONING 122–28 (2004) (applying abductive reasoning to evidence law).

198. Two recent probabilistic interpretations improve on this problem with the conventional account. Professor Cheng rejects the "0.5" requirement and conceives of the standard as comparing the probability of the plaintiff's case with the probability of the defense's alternative account. Cheng, *supra* note 12, at 4–9. By contrast, Professor Clermont's recent proposal based on fuzzy logic conceptualizes the standard as measuring belief in the truth of a proposition compared with belief in the negation of the proposition (or that it is false), with a third category for "unknown" belief. Clermont, *supra* note 39, at 54–59. These conceptions better fit with the goal of equalizing the risk of error, but they also depend on the subjective beliefs of fact finders and thus raise similar accuracy problems as with the conventional probabilistic conception. They also do not appear to provide criteria for evaluating which applications are reasonable.

199. This also frustrates the goal of minimizing error; to achieve this goal the law should go with the more likely account.

200. *See* Eyal Zamir & Ilana Ritov, *Loss Aversion, Omission Bias, and the Burden of Proof in Civil Litigation*, 41 J. LEGAL STUD. 165, 186–91 (2012) (determining that the actual standard of persuasion required in civil disputes is nearly twenty points higher than .51).

201. *See* Ronald J. Allen & Sarah A. Jehl, *Burdens of Persuasion in Civil Cases: Algorithms v. Explanations*, 2003 MICH. ST. L. REV. 893, 897–904 (exploring jury instructions and jury verdicts to determine if a conjunction problem actually exists); Saul Levmore, *Conjunction and Aggregation*, 99 MICH. L. REV. 723, 724–33 (2001) (discussing the "math-law divide" and responses to the law's rejection of the product rule).

probabilistically. For example, the plaintiff wins in a claim with two elements by proving each to 0.6, even though the likelihood of the case as a whole is 0.36 (assuming the elements are independent of each other). If employed in such a manner, too much of the risk of error would fall on defendants. If, by contrast, the probabilistic threshold of 0.5 applied to the case as a whole, then too much of the risk of error may be shifted onto plaintiffs and may arbitrarily depend on the number of formal elements used to specify a claim.[202] The extent to which this issue may create any problems in actual cases is not clear.[203] It arises as a theoretical matter as an implication of the probabilistic conception of the proof process and is thus another strike against this particular conception.

C. The Integration Constraint

The integration constraint requires that a theory's accounts of the micro- and macro-levels cohere. This criterion is a problem for the probabilistic conception for the following reason: *macro*-level proof standards based on subjective probabilities (which are the only plausible kind) are inconsistent with *micro*-level conceptions of relevance and probative value in terms of objective probabilities (which are the only plausible kind). This tension requires some unpacking to make the inconsistency plain.

Begin with the micro-level. Relevance and probative value depend on more than the subjective assessments of individual fact finders.[204] This point accords with conventional probabilistic explanations of evidence at the micro-level: some evidence may be relevant even if jurors believe it is irrelevant, and evidence may have probative value that deviates from what particular jurors believe. For example, in *Liberal Candidates* the evidence is irrelevant even if particular jurors think it is incriminating evidence that supports the prosecution. Indeed, this is why judges have authority to exclude evidence.[205] And, at the macro-level, applications of the proof standards are essentially assessments of the probative value of the

202. For example, in a claim with four independent elements, the plaintiff would have to prove each element beyond 0.841.

203. This is so because jurors do not typically reason probabilistically in an element-by-element fashion. *See supra* note 17 and accompanying text.

204. If it did not, judges would have no basis to exclude evidence on grounds of relevance or under FED. R. EVID. 403.

205. FED. R. EVID. 401–03. Note, however, that once evidence has been admitted it will generally be up to individual jurors to decide for themselves whether evidence is relevant and its probative value.

evidence as a whole, which presupposes the possibility of a reasonable or unreasonable application. For example, a plaintiff's evidence in a civil case may be so weak that no reasonable jury could find in its favor at trial or so weak that an appellate court should reverse a judgment in the plaintiff's favor. Likewise, a criminal case may be so weak that it should be dismissed before trial or a conviction overturned on appeal following a guilty verdict by a particular jury. Indeed, this is why judges have authority to issue summary judgments and judgments as a matter of law in civil cases and to dismiss criminal cases based on insufficient evidence.[206] So far, so good.

Although the foregoing implies that there is more to the probative value of evidence than subjective beliefs, subjective probability standards at the macro-level deny this fact and are thus inconsistent with the micro-level account. At the macro-level, whether the evidence in a case is sufficient to satisfy a probabilistic proof standard will depend simply on whether the particular fact finders think that the evidence surpasses the threshold. The perceived strength of the evidence may be quantified based on each juror's "degree of belief" or "degree of confidence" in the truth of the disputed facts,[207] but the inherent subjectivity remains. Moreover, the lack of objective data for most items of evidence means that a switch to "objective" probabilistic conceptions at the macro-level is not a feasible alternative and cannot fill the gap.[208] Therefore, as a theoretical matter, we have a contradiction. Something more—and more plausible—is needed.

We can illustrate the problem from the opposite direction as well. Begin now at the macro-level with proof standards based on subjective assessments. According to this conception, there is nothing more to the probative value of evidence beyond subjective assessments.[209] But this is inconsistent with the idea that jurors could

206. *See supra* note 25.

207. *See* Hajek, *supra* note 109, at 3.3 (explaining that probabilities can be calculated based on degrees of confidence, which leads to several interpretations of probability that vary by agent and time).

208. Furthermore, if objective, relative-frequency data were available for each item of evidence (and we knew the dependence relationships among these items), then there would be only one reasonable answer, the objective one.

209. The elegance of subjective Bayesian "convergence" to agreement or to truth is lost in typical litigation settings. *See* Goldman, *supra* note 14, at 240 n.10 ("It is doubtful . . . that any of these convergence results would apply to examples in the law because they only concern getting truth 'in the limit,' and legal adjudication is presumably interested in the here and now, not 'the limit.' "). The various convergence theorems depend on the notion that posterior probabilities

be mistaken about relevance and probative value, and thus judges would have no basis to exclude evidence on these grounds. And the micro-level account denies these claims. To avoid a contradiction, a subjectivist at the macro-level must concede that relevance and probative value also depend solely on subjective assessments.[210] With this concession, however, the theory fails as an explanation of relevance and probative value as they currently operate, and it fails as a normative account. Either way, the theory is in a bind: *the only plausible probabilistic conception for satisfying the macro-level constraint is inconsistent with what is required to meet the micro-level constraint.*

III. AN EXPLANATORY CONCEPTION OF EVIDENCE AND PROOF

An alternative conception based on explanatory considerations provides a superior theory of evidence and proof. This Part first outlines the contours of this conception and then examines it in light of the three theoretical constraints.

The alternative conception focuses on potential *explanations* of the evidence and the disputed events.[211] Rather than examining probabilistic relationships—either the probability of *propositions* given evidence, or the probability of *evidence* given propositions—the explanatory conception examines whether particular explanations, if true, would be better or worse at explaining the evidence and the

based on large amounts of data will converge and "wash out" subjective priors. *See* JOHN EARMAN, BAYES OR BUST? 144–49 (1992) (discussing Bayesian convergence); *cf.* Kaplow, *supra* note 15, at 773 n.62 (noting that consistency constraints on subjective beliefs will eventually cause the beliefs to converge on a proof threshold).

210. Or this apparent inconsistency must be explained away. In other words, what is needed is an explanation of why different probabilistic conceptions apply or ought to apply at each level.

211. *See* Pardo & Allen, *supra* note 14, at 225, 229–42 (discussing inference to the best explanation). The explanatory conception relies on both the *process* of explanatory ("abductive") reasoning and the *products* of that reasoning (i.e., explanations). For an overview of the psychology of explanatory reasoning, see Tania Lombrozo, *Explanation and Abductive Inference*, *in* THE OXFORD HANDBOOK OF THINKING AND REASONING 260, 261, 263, 265–68, 270 (Keith J. Holyoak & Robert G. Morrison eds., 2012) (discussing the role of explanation in cognition, everyday life, and learning). *See also* PHILIP N. JOHNSON-LAIRD, HOW WE REASON 186–88 (2006) (discussing the various roles of explanatory reasoning in daily life, law, and science). Explanations concern epistemic relationships between propositions (the explanations) and what they are meant to explain (e.g., other propositions, events, actions, or evidence). For an illuminating philosophical analysis of these relationships, see PETER ACHINSTEIN, THE NATURE OF EXPLANATION 74–102 (1983) (distinguishing explanations from sentences, propositions, and arguments).

underlying events than alternative explanations.[212] The inferential process under this conception is "abductive" (or one of "inference to the best explanation").[213]

This inferential process involves two aspects. First, potential explanations are generated. The primary explanations at issue are those provided by the parties; fact finders, however, may (but are not required to) develop alternative explanations. Second, an explanation is selected as best, or at least acceptable, depending on the decision standard. Different standards of proof mean that the process is not always inference to the best explanation; the acceptability of an explanation for proof purposes shifts with the burden and standard of proof.[214] Several criteria make an explanation better or worse than its alternatives. These criteria include: consistency, coverage, completeness, simplicity, absence of gaps, coherence, consilience, and fit with background knowledge.[215] However, there is no agreed-upon

---

212. Like the probabilistic conception, the explanatory conception is also ultimately concerned with the likelihood or probability of contested propositions at trial. The primary difference, however, is that the explanatory conception posits explanatory rather than explicitly probabilistic criteria to guide inferences and to arrive at judgments about these propositions.

213. Abductive inferences or "inferences to the best explanation" feature prominently in scientific (as well as everyday) reasoning. *See* PETER LIPTON, INFERENCE TO THE BEST EXPLANATION 55–70 (2d ed. 2004) (discussing the "inferences to the best explanation" theory, its advantages and disadvantages, and the ways in which evidence may be relevant to hypotheses); WALTON, *supra* note 197, at 21–23 (reviewing the reasoning processes involved in an "inference to the best explanation," and examining "abductive inferences"); Gilbert Harman, *The Inference to the Best Explanation*, 74 PHIL. REV. 88, 88–89 (1965) (asserting that the truth of an explanation is inferred by rejecting the plausibility of alternative hypotheses). For numerous examples of the roles played by explanatory considerations in science, see the answers provided to *Edge*'s 2012 annual question: "What is Your Favorite Deep, Elegant, or Beautiful Explanation?" *2012: What Is Your Favorite Deep, Elegant, or Beautiful Explanation*, THE EDGE, http://edge.org/annual-question/what-is-your-favorite-deep-elegant-or-beautiful-explanation (last visited Oct. 20, 2012). For a discussion of why objections to "inference to the best explanation" in the philosophy of science do not carry over to law, see Pardo & Allen, *supra* note 14, at 242–45 (arguing that the scientific objections to "inference to the best explanation" don't apply to law because the law deals with different questions and has different epistemic goals). Legal reasoning and theorizing about doctrinal areas of law also involve "inference to the best explanation." *See* Amalia Amaya, *Legal Justification by Optimal Coherence*, 24 RATIO JURIS 304, 306–10 (2011) (outlining the existing approaches to evidential reasoning); W. Bradley Wendel, *Explanation in Legal Scholarship: The Inferential Structure of Doctrinal Legal Analysis*, 96 CORNELL L. REV. 1035, 1036–40, 1055–73 (2011) (providing an overview of how doctrinal analysis relies on inference to the best explanation).

214. *See infra* Part III.B.

215. These criteria provide norms for inference. *See* LIPTON, *supra* note 213, at 122–23 (discussing the relationship between inferential virtues and explanations). Although causal considerations play an important role in assessing explanations, abductive reasoning involves a distinct process from causal reasoning. Explanatory reasoning is broader and involves noncausal considerations; moreover, not all causal hypotheses are explanatory. *See* Lombrozo, *supra* note 211, at 270–71 (distinguishing explanatory and causal reasoning). For this reason, the

algorithm or method for combining or ranking the criteria as a general matter. Rather, the salience and importance of each will depend on details of the particular context and the inferential task at hand.[216] Once an explanation is selected, then verdicts proceed depending on whether the selected explanation includes the elements of the claim, crime, or affirmative defense.[217]

The explanatory process overlaps to a significant extent with the empirical Story Model of juror decisionmaking,[218] but there are important differences. These differences shed light on the distinct roles played by a *theoretical* project as compared with empirical projects. Consider first the overlap. Many of the criteria that make a story persuasive also make an explanation better, and vice versa.[219] This overlap is a theoretical virtue of the explanatory conception because the theory coheres with the actual inferential operations at trial in ways the classical probabilistic conception does not. The overlap makes prescriptive or evaluative recommendations more feasible because they arise from a conception that builds upon, rather than one that is at odds with or alien to, the existing reasoning processes of fact finders.[220]

When the Story Model and the explanatory conception diverge, however, the latter provides normative guidance. They diverge in three crucial ways. First, explanations, unlike stories, may be quite general.[221] Rather than the specific stories posited by the Story

---

explanatory conception is distinct from alternative probabilistic models that rely on causal reasoning (most notably, Mill's). *See* LIPTON, *supra* note 213, at 124–28 (contrasting "inference to the best explanation" and Mill's "Methods"); Stein, *supra* note 39, at 235–46 (articulating a causative probability theory).

216. *See* Paul Thagard, *Evaluating Explanations in Law, Science, and Everyday Life*, 15 CURRENT DIRECTIONS PSYCHOL. SCI. 141, 144 (2006) (discussing the impact of emotional bias and other circumstances on inferential criteria). Although the explanatory criteria must be applied on a case-by-case basis to determine the quality of an explanation, *that* explanatory criteria determine outcomes and the quality of explanations needed to satisfy evidentiary standards remains constant among types of cases. The context dependence and lack of an algorithm for assessing explanations may be a theoretical downside for the explanatory conception, but the success of the conception in light of the three theoretical criteria nevertheless makes it superior to the probabilistic conception (which fails all three). Moreover, as discussed below, the explanatory conception allows for more objective assessments on a case-by-case basis than the inherently subjective ones required by the probabilistic conception.

217. As an empirical matter, verdicts are determined by deducing whether chosen accounts of what occurred include the elements or not, not by deciding serially (and inductively) among the elements. Pennington & Hastie, *supra* note 17, at 520–29.

218. *Id.* at 520–21.

219. *See id.* at 527–29 (listing coverage, coherence, and uniqueness as primary considerations for story persuasiveness and juror confidence).

220. *See supra* note 46.

221. For example, in a *res ipsa loquitur* tort case a general explanation may be sufficient.

Model,[222] the generality of alternative explanations at trial will depend on the substantive law and the ways in which the parties attempt to prove their cases. Second, unlike integrated stories, explanations may be disjunctive.[223] General or disjunctive explanations may still be the best available explanation, and, therefore, the results implied by the explanatory conception may diverge from the Story Model or the decisionmaking of an actual jury. Finally, although decisionmaking under the Story Model fits reasonably well with the preponderance standard, it does not fit well with higher proof standards. With higher standards, the explanatory conception provides guidance and constraint in ways that may diverge from the Story Model and the decisionmaking of actual jurors—and that converge with the goals of the proof standards. This convergence is a further theoretical virtue.

The discussion that follows articulates how the explanatory conception fits with each constraint; explains how it overcomes problems that beleaguer the probabilistic conception; and responds to potential counterarguments.

---

222. *See* Pennington & Hastie, *supra* note 17, at 552.

223. Parties may offer multiple stories consistent with the Story Model, but these stories will typically work together as an explanation. *See* Reid Hastie, *What's the Story?: Explanations and Narratives in Civil Jury Decisions*, *in* CIVIL JURIES AND CIVIL JUSTICE 31–32 (B.H. Bornstein et al. eds., 2008) (explaining that defendants typically offer two stories: "The story of the defendant's activities and a second story to account for the events that led to the lawsuit."). By contrast, the disjunctive explanations referred to above may also be alternatives that are inconsistent with each other, rather than explanations of different aspects of the evidence and events. For example, an explanation that "X, Y, or Z happened" may be sufficient if X, Y, and Z all support the party offering this explanation, and this disjunctive explanation is itself better than the alternative explanation (which may or may not itself be disjunctive). For cases that fit this pattern of proof, see Zuchowitz v. United States, 140 F.3d 381 (2d Cir. 1998), and Rhesa Shipping Co. v. Edmunds, (1985) W.L.R. 948 (H.L. (E.)), *reprinted in* EVIDENCE, PROOF, AND FACTS: A BOOK OF SOURCES 304 (Peter Murphy ed., 2003). *See also* Alex Stein, *An Essay on Uncertainty and Fact-Finding in Civil Litigation, With Special Reference to Contract Cases*, 48 U. TORONTO L.J. 299 (1998) (analyzing *Rhesa* and defending a disjunctive explanation). Similarly, a unified explanation may be better than a disjunctive series of known and unknown explanations. *See* Stubbs v. City of Rochester, 124 N.E. 137 (N.Y. 1919) (holding that a reasonable jury could find defendant's negligence caused plaintiff's illness without plaintiff needing to eliminate all other causes). The disjunctive explanations referred to above and in the text arise within a particular legal claim, not the aggregation of multiple legal claims. Whether parties should be able to aggregate the likelihood of their multiple legal claims is outside the scope of this Article. For an illuminating recent discussion of this issue, see Ariel Porat & Eric A. Posner, *Aggregation and Law*, 122 YALE L.J. 2 (2012).

A. The Micro-Level Constraint

Relevance and probative value depend on whether evidence supports a party's explanation or is a challenge to the other side's explanation. This conception avoids three problems with a probabilistic conception of relevance and probative value: (1) "overlapping" evidence that is relevant but does not distinguish between the cases probabilistically; (2) relevant evidence that does not coincide with increases (or decreases) in probability; and (3) the reference-class problem.

First, the explanatory conception accounts for evidence that is relevant but does not distinguish between cases. Evidence is relevant if it is part of or supports one side's explanation (or challenges one side's explanation),[224] regardless of whether it is also part of or supports the other side's explanation. The probative value of evidence depends on the role it plays in each side's explanation or in distinguishing between them.[225] This conception accounts for the examples presented in Part II of overlapping evidence—the case of *People v. Johnson* and *Twins*:

> *People v. Johnson*[226]: The evidence about the package forms a critical part of the explanations put forward by each side. For the prosecution, the evidence supports a motive for the defendant to attack; for the defense, it supports an (understandable) motive for the defendant to withhold the tray, which then provides a motive for the guards to attack. The evidence provides a reason to believe each explanation. It is relevant.
>
> *Twins*[227]: The eyewitness testimony is relevant because it forms part of each side's explanation of who committed the crime: the defendant or his twin. The testimony provides information about the relevant events. It supports, or provides a reason to believe, each explanation. It is relevant.

Second, the explanatory conception accounts for examples of relevance that are not explained by the conventional probabilistic conception. Evidence may support one side's explanation even when it decreases the probability of that explanation, and evidence may not be

---

224. Evidence may provide a challenge to one side's explanation in a variety of ways. Impeachment evidence is one example. Other examples include evidence that contradicts part of the explanation, implies that part of the explanation is false, or suggests that for the explanation to be true a number of extraordinary and implausible assumptions would need to be made.

225. Indisputable evidence on a material issue that is consistent with one side's explanation and that could not possibly exist if the other side's explanation were true would be highly probative. For example, indisputable evidence that a defendant was in another country at the time of the crime will be highly probative if the prosecution's explanation of what occurred requires the defendant's presence in the country.

226. *See supra* notes 140–51 and accompanying text.

227. *See supra* notes 151–52 and accompanying text.

relevant to prove a proposition even when it increases its probability. Part II presented the examples of *Poison* and *Lottery* to illustrate that an increase in probability may be neither necessary nor sufficient for epistemic relevance. The explanatory conception accounts for these examples:

> *Poison*[228]: The second poisoning is part of the prosecution's explanation of what occurred. Even if the evidence lowers the probability of poisoning from the probability prior to its introduction, it nonetheless provides evidence that supports, or provides a reason to believe, the prosecution's explanation. It is relevant.
>
> *Lottery*[229]: The evidence about the nine hundred tickets supports Defendant's explanation that Rival was the culprit. Even though it raises the probability that Defendant is the culprit from the probability prior to its introduction, it supports, or provides a reason to believe, Defendant's explanation (Rival did it).

Finally, because the explanatory conception does not attempt to quantify the probative value of evidence, it avoids the reference-class problem. Probative value in every case depends on the explanatory connections between the evidence and the alternative explanations. Of course, data from different reference classes will have different explanatory power, and explanations will do better or worse at explaining data from different classes. But these considerations fit naturally into the explanatory approach. The critical question—what explains the evidence—isolates what is important about the evidence and how it contributes to individual cases without the need to quantify based on reference classes. In the *Abuse* example, the issue is what best or better explains the child's behavior, and the statistical evidence is relevant to answering this question.[230] Two other examples where statistical evidence from particular reference classes may be highly probative are (1) epidemiological evidence (which is statistical in nature) comparing the incidence of disease among two groups to prove causation,[231] and (2) statistics regarding an employer's hiring,

---

228. *See supra* note 154 and accompanying text.

229. *See supra* note 159 and accompanying text.

230. The distribution of cabs in the *Blue Cabs* example, *see supra* note 179 and accompanying text, does not have this explanatory connection and, although relevant, it is thus less probative than it would be were there such a connection. The primary conceptual difficulty generated by the reference-class problem is that it undermines all attempts to quantify probative value. This conceptual difficulty simply does not arise under the explanatory conception because there is no claim that any particular class may be used to quantify probative value. *See* Allen & Pardo, *supra* note 10; Ronald J. Allen & Michael S. Pardo, *Probability, Explanation, and Inference: A Reply*, 11 INT'L J. EVIDENCE & PROOF 307 (2007) (discussing the relationship between the reference-class problem and inference to the best explanation).

231. *See generally* Daubert v. Merrell Dow Pharm., 509 U.S. 579 (1993) (analyzing proper standard for admission of expert testimony). Even in this example, however, probative value cannot be quantified and must somehow be integrated with other evidence. *See* Jennifer L.

promotion, or firing practices to support a charge of discrimination.[232] There is nothing unique about statistical evidence—it can be comfortably integrated into the explanatory conception, without the need to quantify probative value and the conceptual difficulties generated by the problem of reference classes that follows as a result.[233]

In addition to avoiding the conceptual difficulties underlying the probabilistic conception, the explanatory conception also coheres with the general concept of epistemic relevance. On the latter, recall the previous quotations from Cohen: "[R]elevance . . . is a relation—typically appreciated or pointed out in the course of a reasoned enquiry, debate, conversation, meditation, explanation, or justification—between a true proposition and an askable question."[234]

And, "[a]nything that can sanction a reason, even if an incomplete or inconclusive reason, for accepting a particular type of proposition as a correct answer, or for rejecting it as an incorrect answer, to an askable type of question can count as a criterion of relevance."[235]

Evidence is relevant in the legal context if it provides a reason for thinking a particular explanation is true or false. The askable questions in this context include general questions like "what happened?" and whatever specific questions are necessary for deciding the elements of the claims, crimes, or defenses at issue. The alternative explanations provide potential answers to those questions, and evidence is relevant when it provides a reason to support or challenge one of the explanations.

One might object, however, that the explanatory conception is *overinclusive*. According to this potential counterargument, if evidence can be relevant even when it does not distinguish the cases

---

Mnookin, *Expert Evidence, Partisanship, and Epistemic Competence*, 73 BROOK. L. REV. 1009, 1022–23 (2008) (discussing the role of epidemiological and other evidence in the context of toxic torts).

232. *See* Int'l Bros. of Teamsters v. United States, 431 U.S. 324, 339 (1977); *see also* Rachel F. Moran, *What Counts as Knowledge? A Reflection on Race, Social Science, and the Law*, 44 LAW & SOC'Y REV. 515, 543 (2010) (noting increasing resistance among courts to statistical evidence in "disparate impact" cases).

233. This conception also explains the examples provided by Cheng. *See* Cheng, *Practical Solution*, *supra* note 100, at 2102–03. For example, he argues that in a case about the disputed value of a home, reference class data that included "age" and "number of bathrooms" would be better than data that included only one or the other of these variables. *Id.* The house's age and number of bathrooms together provide a better explanation of why it has the value that it does than does an explanation that omitted one of these.

234. *See* Cohen, *supra* note 121, at 178.

235. *See id.* at 181.

probabilistically, then does this mean that (all) irrelevant evidence will also be counted as relevant? No. Parties still have to provide reasons why the evidence supports their explanations (or reasons why it challenges the other side's explanation).[236] The failure to satisfy this requirement when faced with a relevance objection provides adequate grounds for excluding evidence as irrelevant.[237] Relevance objections ought to be sustained if either (1) the fact for which the evidence is offered is not material (i.e., is not a "fact . . . of consequence") or (2) the reasons given about the relationship between the evidence and the explanation are either false or too speculative.[238] For example, consider the *Liberal Candidates* example.[239] The evidence is irrelevant because it does not provide any reasons for thinking the prosecution's explanation (e.g., the defendant did it) provides a correct answer for the question at issue (e.g., what happened?). To overcome relevance objections, parties must provide plausible reasons as to why the evidence supports or challenges one of the explanations. Evidence supports or challenges an explanation when it provides a plausible reason, even if it is a partial or incomplete one, for thinking the explanation is true or false.

B. The Macro-Level Constraint

Standards of proof based on explanatory criteria avoid the conceptual problems that arise for probabilistic standards. These problems include failing to (1) fit the goals of the rules regarding accuracy and the risk-of-error allocation and (2) provide criteria for guiding and evaluating applications of the evidentiary rules consistent with the goals.

Under an explanatory conception of the preponderance standard, fact finders should infer the best-available explanation, and the party with the burden of proof ought to win if that explanation includes the formal elements. The party with the burden ought to lose if the best-available explanation does not include one or more of the elements. This conception of the standard instantiates the goals of minimizing errors—to the extent better explanations are more likely

236. Under the probabilistic conception, parties must likewise offer theories of relevance that explain the epistemic relationship between the evidence and the fact it is offered to prove.

237. FED. R. EVID. 402; *see supra* note 155.

238. FED. R. EVID. 401; *see* FED. R. EVID. 402. Sufficient uncertainty about whether a reasonable relationship exists between the evidence and the explanations may be dealt with through "conditional relevancy" rulings. *See* FED. R. EVID. 104(b); *supra* note 88.

239. *See* Lempert, *supra* note 9, at 1025.

to be true[240]—and equalizes the risk of error among the parties. Because the standard is comparative, neither side benefits from the unknown probability space. Each side bears the risk that the jury will select an explanation favoring the other side, with ties going to the party without the burden of proof.[241]

Under standards of proof higher than a preponderance of the evidence, the quality of an explanation needed to satisfy the standard rises accordingly. For the clear-and-convincing standard, the explanation must be substantially better than the alternatives. For beyond a reasonable doubt, the standard is whether there is a plausible explanation that includes the formal elements and no plausible explanation consistent with innocence (or that fails to include one or more element).[242] Raising the quality of an explanation needed to satisfy these standards will correspondingly shift the risk of error consistent with the goals of the rules.[243] This is so because whenever the standard is so raised, the party with the burden may provide the better explanation and still lose. Parties with the burden of proof thus bear an additional risk of error whenever their explanations are better than the other side's but still not good enough to satisfy the standard of proof.

---

240. Although better explanations are not guaranteed to be more likely true than worse ones, the world does tend to work this way. At least it has so far. *See infra* notes 266–67 and accompanying text.

241. *See supra* note 73 and accompanying text.

242. This expression tracks the two conventional ways in which defendants can succeed at trial: (1) demonstrating the prosecution's theory is implausible or (2) offering a reasonable alternative consistent with innocence. *See, e.g.*, O'Laughlin v. O'Brien, 568 F.3d 287, 304–08 (1st Cir. 2009) (considering alternative explanations in evaluating evidential sufficiency); United States v. Beard, 354 F.3d 691, 692–93 (7th Cir. 2004) (same); United States v. Newell, 239 F.3d 917, 920 (7th Cir. 2001) (same); *see also* Wilson v. Graetz, 608 F.3d 347, 352–53 (7th Cir. 2010) (finding ineffective assistance of counsel when defense counsel failed to offer an alternative explanation); United States v. Stever, 603 F.3d 747, 753 (9th Cir. 2010) (holding it was abuse of discretion by trial court in not granting defendant's request to compel discovery of evidence that would support his "alternative explanation").

243. Although the standards are not as precise as probability thresholds, they may be better understood by fact finders and may better fit with the reasoning processes of jurors. These features may make explanation-based jury instructions more effective than current instructions. For more on this issue, see Pardo, *Second-Order*, *supra* note 28. A recent empirical study suggests that although the difference between the preponderance and the clear-and-convincing standards has some effect on decisions, the distinction may not be well understood. *See* David L. Schwartz & Christopher B. Seaman, *Standards of Proof in Civil Litigation: An Experiment from Patent Law*, HARV. J.L. & TECH. (forthcoming 2013), *available at* http://ssrn.com/abstract=2110342. When the standards are characterized in explanatory terms, the difference is modest; the higher standard should change the outcomes only in the class of cases where the jury concludes that the choice between explanations is a close call or where they harbor sufficient second-order doubt that the selected explanation is better.

The explanatory conception also provides guidance for fact finders and reviewing courts. There are objective ways to examine explanatory relationships between evidence and competing propositions.[244] In some cases there will be disagreement about which explanation is better (or what makes an explanation better), but this disagreement *in some cases* does not render the explanatory framework as subjective or as problematic as one based on subjective probabilities. The subjectivity problems inherent in the latter are there *in every case*.[245] By contrast, sometimes the objectively better explanation will be clear, and sometimes not.[246] For example, suppose an explanation (E1) at trial could be true only if fact X were also true, and fact X is known to be false; a second explanation (E2) that is consistent with X being false and contains no similar problems would be better than E1. More generally, other things being equal, a consistent explanation is better than a contradictory one[247]; an explanation that accounts for more evidence and the most important items of evidence[248] is better than one that cannot; and an explanation that adheres with background knowledge and does not require extraordinary assumptions is better than one that is inconsistent with background knowledge and requires unrealistic assumptions.[249] These explanatory criteria connect with deeper doctrinal issues on the civil and criminal sides: they provide criteria for evaluating applications.[250]

---

244. *See* Clermont, *supra* note 20, at 482 ("[T]he law by its standard of proof seeks to force the fact-finder, in the final decisional step, to link its inside mental state to the outside real world.") Explanatory considerations force these mental states to connect with features of evidence in the "outside real world." Some areas of law—e.g., employment discrimination and antitrust—have developed doctrinal frameworks for evaluating competing explanations. *See* Reeves v. Sanderson Plumbing Prods., Inc., 530 U.S. 133, 149–50 (2000); Matsushita Elec. Indus. Co. v. Zenith Radio Corp., 475 U.S. 574, 588–95 (1986).

245. These problems are discussed in Part II.B *supra*.

246. This makes the explanatory conception better in some cases—and no worse in any other cases—than the probabilistic conception on this point.

247. When there is an inconsistency between evidence and an explanation, fact finders will need to reject either the evidence or the inconsistent explanation; the more credible the evidence, the less plausible the inconsistent explanation becomes.

248. What is "important" will vary from case to case and among jurors in individual cases. But, in general, the important evidence will include what the parties argue best supports their explanations; the nonoverlapping evidence that distinguishes the parties' explanations; and any evidence that, if true, shows that one of the proffered explanations is false.

249. Eliminating alternatives makes unlikely possibilities more plausible. *See* Anderson v. Griffin, 397 F.3d 515, 521 (7th Cir. 2005) ("[I]f in a particular case all the alternatives are ruled out, we can be confident that the case presents one of those instances in which [a] rare event did occur.").

250. They also provide a burden of explanation for parties and reviewing courts. *Cf.* Los Angeles v. Alameda Books, Inc., 535 U.S. 425, 437 (2002) ("Neither the Court of Appeals, nor respondents, nor the dissent provides any reason to question the city's theory. In particular, they

We can illustrate these abstract considerations with a few examples. In a case in which a plaintiff alleges age discrimination as the reason for not hiring him, for example, the case will depend on the plausibility of the age-discrimination explanation when compared with the alternative(s) presented by the defendant (or constructed by the fact finder).[251] The comparative assessment may sometimes be straightforward. The plaintiff's explanation will be worse—objectively worse—than the defendant's if the defendant can produce uncontested evidence that the defendant hired employees older and better qualified than the plaintiff.[252] Similarly, suppose a negligence case following a traffic accident turns on whether a seatbelt was defective and failed during the accident (plaintiff's explanation) or whether the driver failed to wear the seatbelt in the first place (defendant's explanation).[253] The plaintiff's explanation will be better than the defendant's if the plaintiff's injuries appear to be caused by wearing the seatbelt for at least part of the accident ("seat belt burn") and cannot be explained by other means.[254] In other cases, the comparative assessment will be more difficult because of conflicting evidence or because the determination depends on judgments of witness credibility. In neither of these two examples, however, could one

---

do not offer a competing theory . . . ."). A court granting summary judgment or judgment as a matter of law for a civil defendant, for example, ought to justify its decision by explaining why a reasonable jury could not find the plaintiff's explanation of the evidence and events to be more plausible than defendants. Similar considerations also structure the federal pleading regime. *See generally* Pardo, *Pleadings*, *supra* note 28 (discussing pleadings after Bell Atlantic Corp. v. Twombly, 550 U.S. 544 (2007), and Ashcroft v. Iqbal, 556 U.S. 662 (2009)).

251. *See* Reeves v. Sanderson Plumbing Prods., Inc., 530 U.S. 133, 141–49, 151 (2000) (comparing age-discrimination explanation with alternatives in upholding verdict for plaintiff).

252. *See*, *e.g.*, Yeschick v. Mineta, 675 F.3d 622, 627 (6th Cir. 2012) (upholding summary judgment for defendant when defendant presented evidence of hiring employees, for the same job, who were older than plaintiff).

253. *See* Bammerlin v. Navistar Int'l Transp. Corp., 30 F.3d 898, 902 (7th Cir. 1994) (concluding evidence of seatbelt defect was sufficient to support a jury verdict for plaintiff).

254. Here is how Judge Easterbrook explained the matter in *Bammerlin*:

> [Plaintiff] proceeded by eliminating the alternatives. We know that he wound up outside the cab. How did he get there? Navistar's theory is that he was not wearing a seat belt. Bammerlin countered with his say-so, plus a physician's testimony that some of his injuries are more consistent with his wearing a belt than with the hypothesis that he was not wearing one. ("Seat belt burn," the physician called it.) A biomechanics expert added that the injuries are more consistent with Bammerlin's being belted for *part* of the time during the crash . . . .
>
> Bammerlin produced evidence that could lead a rational jury to eliminate the hypotheses inconsistent with his favored theory, which in turn permits an inference that his hypothesis is true.

Id..

specify either the objective probabilities for each explanation or which subjective probabilities are reasonable and unreasonable.

Moreover, even explanations that seem unlikely (or have low probabilities in the abstract) will be better if the evidence eliminates the other plausible alternatives. For an example, consider the case of *Anderson v. Griffin*.[255] The case involved a traffic accident that resulted from the "driveline" breaking on a semi-tractor truck, severing the connection between the truck's brake pedal and the brakes.[256] The lawsuit involved a claim by a plaintiff who was driving a car and was subsequently injured in the crash against a dealership that performed service on the truck's driveline weeks before the accident.[257] The plaintiff's explanation, supported by expert testimony, was that the defendant negligently failed to repair the driveline.[258] The defendant's explanation, supported by expert testimony, was that "road junk," or debris on the highway, was pulled up against the driveline, causing it to break.[259] The defendant supported this explanation with evidence that (1) the driveline was working properly and appeared to be in good shape when it left the defendant's shop, and (2) no work was performed on the particular area (the "slip yoke") where the driveline broke during the accident.[260] *If* a fact finder believes the defendant's evidence about the appearance of the truck when it left the shop, then the road junk explanation is at least as good as, if not better than, the plaintiff's negligence explanation, even if the road junk explanation is, itself, unlikely.[261] As with the above explanations, neither objective nor subjective probabilities provide guidance or constraint in deciding which inferences are reasonable and unreasonable.[262]

---

255. 397 F.3d 515 (7th Cir. 2005).

256. *Id.* at 517. The driveline "transmits power from the engine to the real axle." *Id.*

257. *Id.* at 517–18.

258. *Id.* at 518.

259. *Id.*

260. *Id.* at 517–18, 521.

261. In upholding a jury verdict for the defendant, Judge Posner explained that if the jury believed the defendant's evidence, then:

> [T]he likeliest alternative explanations for the accident are either that some deeply hidden defect that [the defendant] could not have been expected to discover had caused the accident or some external force, such as road debris somehow thrown against the yolk by the motion of the truck; in neither event would [the defendant] be liable.

Id. at 521–22.

262. Similar considerations apply to criminal cases. The higher standard of proof means that the comparative assessment shifts—rather than a straight comparison, a conviction is warranted when the prosecution's explanation is plausible (and includes the elements of the

The normative guidance provided by the explanatory conception also responds to potential empirical and theoretical problems with current implementations of the rules. For example, if fact finders employ a conception of the preponderance standard that places too much of the risk of error on parties with the burden of proof,[263] then an instruction expressed in comparative, explanatory terms may better align decisionmaking with an equality-based, risk-of-error allocation.[264] Moreover, the explanatory standards resolve any potential "conjunction" problems that may shift an unjustified risk of error to one side or the other.[265] If fact finders first decide among competing explanations (as dictated by the applicable standard of proof), and then examine whether the selected explanation includes the substantive elements, this decision procedure removes any conjunction issues that may result as implications of an element-by-element decision process. The undesirable shifts implied by conjunction under the probabilistic conception are avoided because fact finders first select an explanation and then determine whether it

---

crime) and there is no plausible explanation consistent with innocence. The criteria for evaluating the quality of an explanation remain the same, however. For an example in which a conviction is not warranted under the explanatory conception, see O'Laughlin v. O'Brien, 568 F.3d 287 (1st Cir. 2009). The defendant was convicted of armed robbery and burglary; the victim was severely beaten in her apartment and did not remember the attack or her attacker. The prosecution's explanation was that the defendant, who was on the maintenance staff at the victim's apartment complex, committed the crime to get money for crack cocaine. In concluding that the evidence was insufficient to support a conviction, Judge Torruella noted that this explanation was not plausible given the following evidence: (1) nothing was taken from the victim's apartment, including jewelry in plain sight; (2) the injuries to the victim (an estimated fifteen to twenty blows to her face and skull) were not consistent with someone motivated only by money; and (3) "no physical or DNA evidence linked [the defendant] to the attack despite the copious amount of blood at the crime scene." *Id.* at 302–04. For an example of a close case in which a conviction is warranted under the explanatory conception, see United States v. Beard, 354 F.3d 691 (7th Cir. 2004). The prosecution's explanation that a gun carried during a drug offense belonged to the defendant was sufficient when (1) the police witnessed the drug sale take place, and (2) there was no other plausible alternative explanation as to how the loaded gun ended up in the car with the defendant. Even though the defendant did not own the car, once it was determined that the gun did not belong to the owner, the only plausible explanation was that it belonged to the defendant. *Id.* at 692 ("No one supposes that the derringer was the property of the car's owner . . . . It would mean that someone [else] who borrowed the car . . . placed a loaded gun in the console, covered it with papers to conceal it, and then—what? Forgot about it? That is possible, but it was not so lively a possibility to compel a reasonable jury to acquit . . . .").

263. This problematic allocation may occur either because jurors employ a noncomparative standard, *see supra* notes 195–99 and accompanying text, or because they employ a probabilistic threshold that is higher than conventionally required by the standard, *see* Zamir & Ritov, *supra* note 200.

264. *See supra* note 74.

265. *See supra* note 201.

includes the formal elements, rather than deciding the elements serially (which generates the paradoxes). Moreover, error allocation under the explanatory conception matches the underlying goals of the rules—the risk is roughly equal in cases under the preponderance standard (the risk applies to the selection of contrasting explanations) and shifts accordingly with higher standards of proof.

One might object to the explanatory conception at the macro-level, however, by challenging the link between explanatory considerations and truth. The explanatory conception depends on the assumption that, other things being equal, a better explanation is more likely to be true than a worse explanation. But, according to this potential counterargument, the better explanations might be false. That is true, but that fact does not undermine the explanatory conception. The counterargument misses the mark because it relies on the general "problem of induction."[266] The "problem" arises because any nondeductive inference (i.e., one in which the premises do not entail the conclusion) may be false. This potential problem, however, affects *any* theoretical account of legal proof, which by its nature involves inductive inferences.[267] Explanatory considerations are not guarantees of truth, but neither are any other criteria that could be used to guide inferences in the context of legal proof (probabilistic or otherwise). Moreover, within the domain of legal proof, there are no other criteria by which to judge whether the better explanation is false in a given case—no way to peer behind the explanations, as it were, and to evaluate truth conduciveness directly.[268] Any considerations that might bear on questions about the truth of explanations arise *within* the explanatory framework. Reasons for thinking an otherwise plausible explanation is *unlikely* are reasons for thinking it is not a good or better explanation (depending on the alternatives).

---

266. DAVID HUME, TREATISE ON HUMAN NATURE 89 (1978) (originally published 1739-40); *see also* John Vickers, *The Problem of Induction*, STANFORD ENCYCLOPEDIA PHIL., http://plato.stanford.edu/entries/induction-problem/ (last visited Oct. 16, 2012) (describing differing accounts of induction).

267. The more likely propositions expressed in probabilistic terms might also turn out to be false. This is a consequence of the fact that both conceptions involve inductive inferences. Inferences may be *deductive* (in which truth is guaranteed, if the premises are true) or *inductive* inferences (in which truth is not guaranteed). Inductive inferences may be *abductive*, *see supra* note 213, or *enumerative* (e.g., after observing ten white swans one may infer the eleventh will be white or that all swans are white). The probabilistic conception relies on enumerative inferences (or a "more of the same" principle), and the explanatory conception relies on abductive inferences (or an "inference to the best explanation" principle). Both are inductive and thus subject to the problem of induction.

268. *Cf.* Wendel, *supra* note 213, at 1074 ("IBE presents itself as a pretty good alternative to doing without theories of explanation.").

Considerations that make an explanation appear unlikely ought to be incorporated into the explanation-evaluation process, rather than used to trump that process. For example, if cognitive psychology can identify distortions that occur as an empirical matter in the evaluation of competing explanations, then these effects should be taken into account when critically examining why one explanation is thought to be better than another.[269] One possibility along these lines concerns what Professor Dan Simon has called "coherence shifts."[270] When fact finders formulate a story of what happened, they may overvalue evidence that supports their story and undervalue or discount evidence that challenges it. If these shifts (or other types of cognitive biases) occur, then this information should inform the explanation-evaluation process, exposing situations where beliefs deviate from what is epistemically warranted.[271]

C. The Integration Constraint

The micro- and macro-levels cohere in a straightforward way, and this consilience provides another theoretical virtue.[272] Explanations as a whole animate the macro-level, and the quality of a macro-level explanation will depend on the relevance and probative value of evidence at the micro-level. For example, an explanation that is consistent with the evidence will be better than a contrasting explanation that is inconsistent with the evidence. Likewise, relevance

269. *See* Saks & Kidd, *supra* note 18 (discussing cognitive biases and legal fact-finding); *see also*, *e.g.*, Jerry Kang et al., *Implicit Bias in the Courtroom*, 59 UCLA L. REV. 1124 (2012) (describing how implicit bias may affect an array of decisionmaking tasks in criminal and civil cases). Because the explanatory conception depends on relationships between explanations and evidence that exist outside the mind of decisionmakers, it provides a more plausible epistemic framework for correcting inferential mistakes due to cognitive biases; subjective probabilistic judgments are more likely to simply replicate these biases. Normative assessments of explanations should, of course, also take into account cognitive biases of judges. *See* Guthrie, Rachlinski & Wistrich, *supra* note 16.

270. *See* Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. CHI. L. REV. 511, 511 (2004) (discussing "coherence shifts" in which subjects discount evidence that does not cohere with their accepted accounts and overvalue evidence that does).

271. These issues should be evaluated from within the explanatory framework at least until there is a better theoretical conception of evidence and proof within which to work. Nothing I have argued for precludes this possibility. Some new theory may turn out to be superior in light of the three theoretical criteria, or some new metatheory may be able to articulate better criteria for evaluating evidence theory. But notice that it will do so by better explaining the proof process or evidence theory than the current alternative explanations—precisely the methodology inherent in the explanatory conception itself.

272. *See supra* notes 38, 85–87.

and probative value at the micro-level will depend on whether the evidence forms part of a party's explanation at the macro-level (or a challenge to the other side's explanation), and the significance of the evidence at the micro-level will depend on the significance of the evidence to these macro-level explanations. Evidence that renders one side's explanation plausible or the other side's explanation implausible will be highly probative. For example, evidence that one would expect to see if an explanation were true renders that explanation more plausible, and evidence that requires the truth of several extraordinary assumptions for both the evidence and an explanation to be true renders that explanation less plausible.

Although the accounts at both levels are mutually reinforcing, a potential counterargument might contend that this interaction is problematically circular. The same explanatory considerations provide guidance, constraint, and a basis for evaluation at both levels. Therefore, so the counterargument would go, if the explanations explain the evidence, on one hand, and the evidence gets its value from the role it plays in supporting the explanations, on the other, have we stumbled into a circle? Any apparent circularity here is virtuous, not vicious. Trials are fundamentally about the alleged events and circumstances that gave rise to litigation in the first place. Proffered evidence is admitted to prove, test, or challenge the competing explanations regarding these events and circumstances. What makes evidence relevant, and what justifies a particular explanation, is not the simple fact that the explanation may explain the evidence; it is that the evidence combines with reasons for thinking that a particular explanation is true or is better than the alternatives.[273]

Given this relationship between evidence and competing explanations, it should not be surprising that the micro- and macro-levels reinforce each other. The counterargument misses the mark because judgments at the micro-level should inform and be informed by the same considerations that will ultimately form macro-level judgments. Any test that employed one set of considerations for determining whether some condition has been met as a general matter but employed entirely different considerations when testing for whether specific instances of that condition obtain—without any explanation of how the two sets of considerations related—would be a

273. For example, evidence that the defendant was out of the country when a crime was committed makes "defendant did not commit the crime" a better explanation of what occurred than "defendant is guilty" because the former explanation coheres with the reason "people can't be in two places at once" and the latter explanation does not.

pretty bad test.[274] The consilience of the explanatory conception on this point is thus a theoretical virtue. What should be surprising is the survival of the probabilistic conception despite its failure to cohere and provide this kind of mutually reinforcing support.

CONCLUSION

The theoretical project of articulating an account of the evidentiary proof process and its components has largely proceeded in piecemeal fashion through the application of formal methods to specific evidentiary or legal issues. There is, of course, nothing wrong with this approach. Indeed, it may be a superior methodology for making progress on specific issues—to the extent their domains can be cabined and made tractable for formal applications. This Article, however, stands for the idea that sometimes grand theoretical ambitions have their place as well. Doing so allows for a broader perspective on how the various issues and components hang together, reveals flaws and limits of more local theorizing, and allows one to get a grip on the various metatheoretical issues at stake in the discussions.

This is the broad perspective taken up in this Article. In doing so, we can draw some general conclusions: the *nature* of evidence theory, whatever other features it possesses, necessarily includes considerations of its epistemological core; the *purpose* of evidence theory, whatever other purposes it serves, must be to provide or rely upon a satisfactory account of this epistemological core in light of the considerations of factual accuracy and allocating the risk of error. This Article elaborated on this epistemological core and these metatheoretical commitments by articulating three necessary constraints on evidence theory (micro, macro, and integration), and concluded that the explanatory conception meets them while the probabilistic conception does not.

A final observation about the significance of evidence theory: any application of evidence doctrine presupposes *some* conception of the key evidentiary concepts.[275] The theoretical constraints identified in Part I provide a check on whether any given explicit or implicit

274. For example, suppose we have a test that purports to assess mathematical ability, but none of the specific questions actually measures mathematical ability.

275. The practical issues that presuppose these conceptions include: decisions to admit or to exclude evidence, motions to dismiss, summary judgments, judgments as a matter of law, motions for new trials, and challenges to sufficiency of the evidence in criminal cases before, during, and after trial and on appeal.

conception makes sense in light of the underlying considerations of accuracy and the risk of error. Issues of procedural justice and constitutional and substantive rights all depend on sound applications in the evidentiary proof process in light of these considerations.[276] Evidence theory provides a way to examine the soundness of these applications. A theory that fails these constraints—as Part II concluded with regard to the probabilistic conception—is not a theory worth having. Given the significant issues at stake, a satisfactory conception is needed. The explanatory conception provides the best available conception in light of the theoretical criteria. This conclusion demonstrates more generally that there is value in evidence *theory*, as distinct from, but in tandem with, empirical and doctrinal investigations of the evidentiary proof process.

276. *See supra* notes 22–28, 49–52 and accompanying text. Indeed, the political legitimacy of law itself depends on adequate evidentiary procedures. *See supra* note 51.