Relevance, Standards of Proof, and Evidence Theory

Michael S. Pardo*

The past 50 years have seen an explosion in theoretical work on the evidentiary proof process. Much of this work began with, explored, and currently surrounds the application of statistics or probability theory to various aspects of evidence and the proof process.¹ Robust debates initially sprung up around such matters as (1) the feasibility and desirability of basing civil and criminal judgments on explicitly statistical evidence;² (2) defining fundamental concepts such as relevance and probative value in probabilistic terms;³ and (3) explaining structural features of the proof process such as the various decision standards (for example, “preponderance of the evidence,” clear and convincing evidence,” or “beyond a reasonable doubt”) in probabilistic terms;⁴ (4) whether the introduction of statistical evidence or techniques

---


---

* Henry Upson Sims Professor of Law, University of Alabama School of Law.
will improve or detract from accurate fact-finding; and (5) the proper analysis of the various “proof paradoxes,” in which outcomes implied by a statistical conception of evidence or a probabilistic conception of the proof process appear to clash with our intuitions or considered judgments about what the proper outcomes should be. These debates have in turn informed recent theoretical work exploring epistemological and moral aspects of evidence and proof.

Looking out at the landscape of these debates, it is now possible to see two distinct but related projects at the heart of the discussions about statistical evidence and probability theory. These projects also animate many of the current debates about statistical evidence as well as more recent theoretical accounts of juridical proof. It helps to keep them distinct. The first project is the integration, as a practical matter, of statistical information and statistical techniques

---


6 For an overview of the literature, see Mike Redmayne, Exploring the Proof Paradoxes, 14 Legal Theory 281 (2008).

into various stages of the litigation process and the law more generally. The second is the integration, as a theoretical matter, of statistical (and other) evidence into a conception or theoretical account of evidence and proof that accords with the legal system’s evidentiary proof process and its underlying normative goals and purposes. These projects are, of course, related—the theoretical account provides guidance for practical integration as well as a standard by which to assess and perhaps critique these practices. For example, we have no idea whether evidence (statistical or otherwise) is or is not sufficient to satisfy the “preponderance” or “beyond a reasonable doubt” standards without some prior conception or presupposition regarding what is or ought to be required by these standards. Although related, the projects involve distinct methodological issues. The practical issues of integration involve a host of difficult empirical questions that are distinct from the conceptual task of articulating an underlying theoretical account of the nature and structure of proof in light of its normative goals.

In addition to and alongside these projects, a robust empirical literature has emerged providing a powerful psychological model of the behavior of legal fact-finders in general and varying degrees of information regarding a host of specific evidentiary questions. The well-confirmed model of jury behavior—the Story Model—posits that legal fact-finders assimilate evidence into competing narratives of the underlying events and select among the most plausible or satisfying of the available accounts. This general model is then supplemented with a variety of claims regarding the behavior of juries and judges with regard to specific types of evidence,

---


instructions, scenarios, and influences. The relationships between the practical and theoretical projects outlined above and the issues explored in the empirical literature on fact-finder behavior raise additional theoretical, practical, and empirical questions.

In this essay, I focus on the theoretical project writ large. I do so with both general and specific aims in mind. As a general matter, I attempt to step back from this theorizing to ask and answer a few methodological and meta-theoretical questions: What would a successful theory of evidence look like? By what criteria can or ought we assess such a theory? And what might the purpose of such theorizing be?

Based on this general discussion, I arrive at a few specific conclusions. The familiar, dominant account of evidence and proof in probabilistic terms fails as a general theory of proof and with regard to particular items of evidence. Under this probabilistic conception, contested propositions may be assessed according to their probability, given the evidence (or vice versa), with such assessments conforming to the axioms or rules of probability theory: for example,

--they fall on a scale of 1 (which equals certain truth) to 0 (which equals certain falsity);

--the probability of a proposition and its negation equal 1;

--the probability of two independent propositions conforms to product rule;


By “probabilistic” I am referring to Pascalian (or mathematical) probability as opposed to Baconian (or inductive) probability. For discussions of the latter in the context of legal proof, see L. Jonathan Cohen, The Probable and the Provable (1977); Stein, supra note 7.
the probability of a proposition in light of new evidence conforms to Bayes’ Theorem; and so on. I illustrate the theoretical failures of this conception of juridical proof with several examples, including a number of examples involving statistical evidence. (These examples are particularly significant because they provide the most plausible test cases for probabilistic accounts.) In pointing out several difficulties for probabilistic accounts of proof, I will argue that a probabilistic theory is not the best way to integrate even overtly statistical evidence into the evidentiary proof process. I sketch an alternative “explanatory” conception of evidence and proof that better satisfies the general criteria for a successful theory and a better framework for integrating any kind of statistical evidence. Under the alternative conception, contested propositions are assessed according to how, if true, they would explain the evidence and events when compared with competing, contrastive propositions.

Although my topic is theoretical, the consequences are practical and significant. At stake are not only issues regarding admissibility and sufficiency of evidence at trials, but also important issues regarding who gets to trial in the first place, which verdicts will be allowed to stand, and which convictions ought to be overturned. In our era of so-called “vanishing trials” on the civil side and the (justified) attention and concern with false convictions on the criminal side, these practical doctrinal issues may be as important as any facing the law today. And, as I

---


will demonstrate, they depend, at root, on a satisfactory account of evidence and proof—in other words, on evidence theory.

I. What is Evidence Theory?

By “evidence theory” or a “conception” or “theoretical account” of proof, I mean a theory, conception, or account of the structure and nature of the evidentiary proof process and how it relates to the underlying goals of factual accuracy and allocating the risk of factual errors among the parties. There is, of course, much more theoretically and practically going on in evidence law, including a host of moral, political, and economic issues, but my focus will be on the core epistemological issues that are necessary—although not sufficient—for a satisfactory theoretical account of evidence and proof.

At a minimum, a successful theory of evidence and proof must satisfy three constraints. First, it must provide a plausible account of the relevance and probative value of individual items of evidence. I refer to this as the “micro-level” constraint. Second, it must provide a plausible account of the various decision standards and sufficiency of the evidence. I refer to this as the “macro-level” constraint. Third, it must provide a plausible account of the relationship between the accounts at these two levels. I refer to this as the “integration” constraint. I elaborate on the nature and significance of each constrain below.

---

15 For an overview of recent empirical and theoretical projects related to the law of evidence, see Roger C. Park and Michael J. Saks, Evidence Scholarship Reconsidered: Results of the Interdisciplinary Turn, 47 B.C. L. Rev. 949 (2006).

16 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 73 (1994).
The Micro-Level Constraint

A theoretical account of evidence and proof should be able to explain the relevance (or irrelevance) and probative value of evidence.\(^{17}\) The two 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.\(^{18}\) The presuppositions underlying these tenents are that irrelevant evidence cannot contribute to a rational assessment of disputed factual issues or improve the factual accuracy of decisions, 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 more accurate decisions.\(^{19}\) A satisfactory theory of evidence and proof ought to be able to explain whether evidence is relevant or not, along with an explanation why. In providing plausible explanations on these issues, a satisfactory theory will serve not only an explanatory theoretical function (along with any clarity and illumination that may follow) but may also serve a variety of normative roles. These normative roles include 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.

\(^{17}\) Fed. R. Evid. 401, 402.

\(^{18}\) James Bradley Thayer, A Preliminary Treatise on Evidence at the Common Law 264-66 (1898). The reasons to exclude otherwise relevant evidence include considerations falling one of the following categories: (1) economic, (2) the advancement of non-epistemic policy goals, (3) party control, and (4) jury control.

\(^{19}\) To be clear, “accuracy” refers to whether the outcomes match \textit{what actually happened} not necessarily whether the outcome matches what is implied by the presented evidence and the proof rules. For a discussion of the distinction, see Michael S. Pardo, \textit{Pleadings, Proof, and Judgment: A Unified Theory of Civil Litigation}, 51 B.C. L. Rev. 1451 (2010).
A theory of evidence and proof will be potentially problematic if the answers it provides or implies about the relevance (or irrelevance) of evidence fail to accord with our intuitions or considered judgments about the relevance (or irrelevance) of evidence. Neither the theory nor particular judgments ought to necessarily have priority, however.\textsuperscript{20} If there is a mismatch 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 satisfactory theory. In this context, the underlying judgments ought to give way, for example, if the outcomes provided or implied by a satisfactory theory would better meet the underlying goals of factual accuracy or would better identify the evidence that ought (i.e., is relevant) and ought not (i.e., is irrelevant) factor into a rational assessment of the disputed factual issues. Although it is thus possible for a theory’s implications to override particular judgments—and it must be for the theory to play a normative role—the burden is on the proponent of a theory to explain why our (or a judge’s) considered concrete judgment on an issue ought to give way to the dictates of an abstract theory.

Similar theoretical and methodological considerations apply to probative value. The probative value of evidence refers to the strength of relevant evidence in proving a disputed factual proposition. Judgments on probative value are highly contextual and depend not only on the logical or empirical relationship between evidence and disputed propositions, but also on such issues as the importance of the evidence, the parties need for the evidence, the availability of other similar evidence, and the extent to which the issue is contested and evidence is introduced by the other side. At the micro-level, as with relevance, probative value provides a fundamental consideration for the admissibility of relevant evidence; judges may exclude

\textsuperscript{20} See Nelson Goodman, Fact, Fiction, and Forecast (4\textsuperscript{th} ed. 1983).
virtually any item of evidence if its probative value is substantially outweighed by the potential dangers of unfair prejudice, confusing the issues, misleading the jury, or for efficiency considerations (undue delay, waste of time, needlessly cumulative). Thus, any satisfactory theory of evidence and proof must provide a plausible explanation of the probative value of individual items of evidence, along with why it has that value. Similarly, if there is mismatch between judgments about probative value and the answers provided by a theory of evidence of proof, the burden is on the proponent of the theory to explain why, for example, evidence that appears to have low probative value in fact has high probative value, or vice versa.

The Macro-Level Constraint

A theory of evidence and proof must also be able to explain 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 evidence as a whole is sufficient to satisfy particular decisions standards—in other words, whether or not a fact has indeed been “proven” for purposes of legal action.

At the macro-level in any given civil or criminal case, one party has the burden of proof on each litigated issue that forms an element of a claim, crime, or affirmative defense. The burden of proof is made of two components: a burden of production and a burden of persuasion. The burden of production is a function of the burden of persuasion—in the sense that parties will have satisfied a burden of production when they have presented evidence from which a

---

21 Fed. R. Evid 403; but see Fed. R. Evid. 609(a)(2). Probative value is also important at the macro-level 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.
reasonable fact-finder could conclude that the burden of persuasion has been satisfied.\textsuperscript{22} The burden of persuasion is determined by the various decision standards: for example, “preponderance of the evidence,” “clear and convincing evidence,” or “beyond a reasonable doubt.”\textsuperscript{23}

The decision standards specify and instruct fact-finders on when to conclude that the evidence as a whole proves a particular disputed fact. Two overarching goals underlie and animate the standards: accuracy and allocating the risk of an erroneous decision. Each goal provides an independent criterion by which to assess the acceptability of the standards. For example, the “preponderance of the evidence” standard, other things being equal, functions (or appears to function) to minimize errors and maximize accuracy. This follows from the fact that decisions under this standard are the ones that are more likely given the evidence (than their alternatives).\textsuperscript{24} Assuming the evidence is a good indicator of truth, decisions that appear more likely given the evidence will be more accurate in the long run than decisions that appear less likely given the evidence. Even if the preponderance rule maximizes accuracy, it does so, however, by spreading the risk of error roughly equally\textsuperscript{25} among the parties (either side may have the less likely case and the decision may be an error). Although this risk of error is acceptable in


\textsuperscript{24} For a discussion of these policies underlying the preponderance rule, see Stein, supra note 7, at 143-44.

most civil cases (where equality among the parties is an important procedural consideration), this
is well-accepted to be an unacceptable risk of error for criminal defendants to bear for the
elements of criminal offenses.\textsuperscript{26}

Likewise, a focus solely on the risk-of-error allocation leaves out the key component of
accuracy. For example, an interpretation of the “beyond a reasonable doubt” standard consistent
with the Blackstonian admonition that one false conviction is as problematic as ten false
acquittals implies plainly unacceptable consequences. It implies that in 100 cases, the following
would be an acceptable distribution of outcomes: 90 false acquittals; 9 false convictions; and 1
accurate determination.\textsuperscript{27}

A satisfactory theory of evidence and proof ought to be able to explain how the decision
standards function in light of these goals. This includes, for example, how the preponderance
rule functions to maximize accuracy and allocate the risk of error roughly evenly among the
parties and how the “beyond a reasonable doubt” and “clear and convincing evidence” standards
skew the risk of error to one side and attempt to produce one type of error rather than another
(while also sacrificing some degree of total accuracy).\textsuperscript{28} Moreover, a successful theory must not
only be able to do so “in theory,” as it were. It must also be able to provide some criteria or
guidance for decision-makers applying the rule in light of the rule’s underlying goals regarding

\textsuperscript{26} See In re Winship, 397 U.S 358 (1970).


\textsuperscript{28} Total accuracy may be sacrificed because cases proven beyond a preponderance of the evidence (but not
to the higher standard) will go against the party with the proof burden, even though these decisions are less likely to
be accurate.
accuracy and the risk of error. 29 The decision standards are instructions to fact-finders on how to arrive at judgments consistent with the goals of the rules. A theory of evidence and proof that can explain formally how the standards fit with their goals, but not how fact-finders could implement the standards consistent with the goals, fails as a theory. 30 Concrete applications of the standards in actual cases are what need explaining and are the object of a successful theory.

What is more, a successful theory at the macro-level must provide some guidance or criteria for separating “reasonable” from “unreasonable” applications of the various standards. This important doctrinal issue has, unfortunately, not been the subject of sustained dialog in theoretical discussions of evidence and proof. 31 But it is critical not only for juridical proof at trial—it ramifies through 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 non-moving party. 32 And what is reasonable or not is a function of the evidence, the burden of proof, and the decision standard. For example, at issue at summary judgment may be whether a reasonable fact-finder could find for the plaintiff on a particular issue (e.g., causation) by a preponderance of the evidence. This obviously implies some understanding or presupposition of what is required by

29 See Allen & Leiter, supra note 7 (discussing the “ought implies can” constraint in epistemology).

30 It is not a legitimate criticism, however, if the standards by themselves do not produce optimal results or a desired distribution of errors. How the standards ultimately operate in practice will depend on their interaction with a host of other issues, including, among others, 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.

31 I discuss these issues in more detail in Pardo, supra note ___ and Michael S. Pardo, Second-Order Proof Rules, 61 Fla. L. Rev. 1083 (2009).

the decision standard (i.e., some applications are reasonable and some are unreasonable). The same standard and distinction applies for judgment-as-a-matter-of-law determinations. Therefore, whether parties get to trial in the first place and whether particular verdicts will be allowed to stand depend at root on the decision standards.

Likewise, in criminal cases whether the evidence is sufficient to support a conviction depends on whether, given the evidence, a reasonable jury could find the issues proven beyond a reasonable doubt. These determinations depend on some conception of what is required by the “beyond a reasonable doubt” standard. Whether cases will be brought to trial, dismissed at trial, or reversed on appeal will depend on a distinction between reasonable and unreasonable applications of the decision standard. A successful theory of evidence and proof must explain this critical distinction on the civil and criminal sides, and provide a normative standard or criteria by which to guide, evaluate, justify, and critique particular judgments.

The Integration Constraint

Finally, a successful theory of evidence and proof must provide a plausible explanation of how its accounts at the micro- and macro-levels fit together. A theory will fail if the theory’s accounts do not do so. A theory that provides plausible accounts at both levels may nevertheless fail if these accounts are inconsistent with one another. For example, a theory that, at the micro-level, accounts for probative value solely in terms of the subjective beliefs of each fact-finder

---

(i.e., the value just is what any given jury thinks is it) cannot then, at the macro-level, explain sufficiency of the evidence in terms that imply the subjective beliefs of individual jurors may be mistaken about the value of evidence, or vice versa. Either there is more to the value of evidence than subjective 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 desirable.  

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 cardinal probabilities and decision standards in terms of ordinal probabilities (or vice versa), or a theory that explains one level in terms of subjective probability assessments (i.e., credal mental states) and one level in terms objective relative frequencies, must explain how the micro-level determinations translate at the macro-level, and vice versa. In sum, a successful theory will bridge the gap between the two levels.

II. Distractions, Confusions, Mistakes: Clarifying Evidence Theory

Given the explosion of theoretical work in evidence law, contemporary discussions of evidence theory such as this one travel on some well-worn and hotly contested ground (or at least what passes for “hotly contested” within the evidence world).  

Given the complexity of these debates, it is not surprising that sometimes issues are confusingly run together, mistaken assumptions or presuppositions are relied upon, and false implications or dichotomies are asserted. Therefore, it will help to clarify what is not at stake or contested—or at least what I

---

36 Such an explanation may be constructed, for example, around the idea that there should be more judicial control at one level rather than the other.

37 If the evidence professor listserv is any indication, there does not appear to be much of a consensus about these fundamental conceptual issues. See supra note 1.
think is not at stake or contested—within evidence theory generally (or within the statistical-evidence debates specifically) in order to avoid further muddying the waters with potential confusions.

1. Statistical evidence is distinct from an account (probabilistic or otherwise) of aspects of the proof process. In other words, whether statistical evidence is relevant, how probative it is, and whether it is sufficient to satisfy a decision rule are distinct issues from whether relevance, probative value, or decision standards are probabilistic or ought to be conceived in probabilistic terms. The relevance, probative value, and sufficiency of evidence, for example, that the Blue Bus Co. owns 60 percent of the buses in town or that the “random match probability” of a DNA sample is 1 in a billion depend not only on what inferences may be drawn from these statistics, but also on a prior conception of the evidentiary concepts (which may, but need not, be conceptualized in probabilistic terms). One cannot draw evidentiary conclusions about statistical evidence for purposes of legal proof without reliance upon some conception of what is required by the evidentiary concepts. For example, one cannot determine whether statistical evidence (of any kind) is sufficient to satisfy the “preponderance of the evidence” standard without some prior conception of what is required by the preponderance rule.

2. Similarly, the practical issues surrounding whether statistical evidence or statistical techniques ought to be introduced or not at trial (or other stages) in particular cases or types of cases are distinct issues from the theoretical issues regarding conceptions of evidence and proof.

---

38 See Roth, supra note 35.

39 The “greater than 0.5” standard is one such conception, but it is not the only one. What is required by the preponderance standard is a conceptually distinct issue from the sufficiency of particular statistical evidence in a given case.
The practical issues surrounding the introduction of statistical evidence or techniques into the litigation process depend on whether it will improve or detract from accurate fact-finding. These issues depend on answers to a number of 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.40

3. From an epistemological standpoint, statistical evidence is neither necessarily inferior nor superior as compared with non-statistical evidence. As with non-statistical evidence, it will sometimes be good (relevant and highly probative) and sometime not, depending on the context. There is no abstract qualitative distinction, a priori or otherwise, between evidence expressed in statistical terms and evidence expressed in non-statistical terms.41

4. Non-statistical evidence is also based on generalizations, so if there is a distinction between “specific” and “general” evidence it is at best one of degree not of kind. For example, the relevance, probative value, and sufficiency of a witness’s testimony depend on a number of generalizations about this witness’s propensity for accurate perception, memory, narration, and for truthfulness, as well as a number of generalization about the behavior of people in general in the witness’s situation. These generalizations link the testimony to the facts of consequence in a particular case in ways that are substantially similar to the generalizations that make up statistical evidence.

40 Moreover, even if jurors are not particularly adept at assessing statistical evidence, see Koehler, supra note __, the introduction of this evidence may still do some 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 (2010).

41 For an example of this confusion, see Amit Pundik, *What is Wrong with Statistical Evidence? The Attempts to Establish an Epistemic Deficiency*, 27 Civil Justice Q. 461 (2008) This is not to deny that there may be moral or political reasons to distinguish between statistical and non-statistical evidence.
5. Within evidence theory, particular conceptions may be useful as teaching tools or for capturing and drawing attention to certain aspects of the theoretical issues, even if they fail as a general theory of the evidentiary or proof issues. 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.

6. With regard to these theoretical issues, there is sometimes a false dichotomy presented between probabilistic conceptions and the Story Model. 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. As noted in the introduction, the Story Model is a descriptive psychological account of jury behavior, and a particularly good one at that. And any theoretical account would be wise to take account of the best empirical account of fact-finder behavior. But . . . whether the account is true, and whether lawyers present cases in this form of overarching narratives or theories, are distinct issues from theoretical accounts of the nature and structure of proof. In short, it is possible to have non-probabilistic descriptive and explanatory accounts of legal proof that also provide normative guidance and constraint.

III. The Micro-Level of Evidence Theory: Relevance and Probative Value

The micro-level constraint provides that a theory of evidence ought to account for the relevance (and irrelevance) of evidence and probative value. It must be able to explain these

42 For comparisons of these two models, see Kenworthy Bilz, We Don’t Want to Hear it: Psychology, Literature and The Narrative Model of Judging, 2010 U. Ill. L. Rev. 429 (2010); Legal Evidence and Proof: Statistics, Stories, and Logic (Kaptein et al eds. 2009).

43 Pennington & Hastie, supra note __.
concepts; whether particular evidence is relevant or not, and its probative value; and why it is relevant or not and has the value it does. It must also guide and constrain judgments about particular items of evidence. Evidence is “relevant” if it “has any tendency to make the existence of any fact of consequence to the determination of the action more probable and less probable that it would be without the evidence.” Evidence that is not relevant is not admissible. “Probable value” refers to the strength of the evidence—or how strong a “tendency” is has—in proving or disproving a fact of consequence, talking into other contextual factors such as, for example, a party’s need for evidence, the importance of the issue, and the extent to which it is contested by the other side. Evidence may be excluded if its probative value is substantially outweighed by one or more “dangers.”

These doctrinal terms naturally invite probabilistic interpretations. And, indeed, the dominant theoretical conceptions of relevance and probative value are in probabilistic terms. Under these conceptions, relevance is expressed either in terms of a “likelihood ratio” or simply in terms of an increase (or decrease) in the probability of a proposition from its prior probability without the evidence. These conceptions are offered as a description or explanation

---

44 Fed. R. Evid. 401
45 Fed. R. Evid. 402
46 Fed. R. Evid. 403
47 Id.
48 See Lempert, supra note 3.
of the meaning of Rule 401’s definition of “relevance” and rule 403’s concept of “probative value” and as a normative standard that provides a criterion by which to assess possible relevance and probative value and by which to evaluate judgment about relevance and probative value.

According to the “likelihood ratio” approach, we examine the probability of receiving evidence if the proposition is true, divided by the probability of receiving the evidence if the proposition is false. Consistent with Bayes’ Theorem, we can multiply the likelihood ratio by the prior probability to arrive at an updated, posterior probability. Consider the following example,

Example #1: 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.

See Id. at 1025 (describing likelihood ratio as “mathematical equivalent” of “relevant evidence” in Fed. R. Evid 401).

See Kaye, supra note __; Jonathan J. Koehler, When Do Courts Think Base Rate Statistics are Relevant? 42 Jurimetrics J. 375 (2002) (“The ratio of these likelihoods (known as the likelihood ratio) captures the probative value of the evidence.”)

Ids; Lempert, supra note 3, at 1023 (“this is a normative model”).

Id.

Odds Guilt Given the Evidence = (Probability of the Evidence Given Guilt / Probability of the Evidence Give not-Guilt) x Prior Odds of Guilt

Lempert at 1024
Professor Lempert, assuming that 50 percent of the population has the same blood type, analyzes the evidence as follows: the probability of the evidence, given that he is the perpetrator is 1, and the probability of this evidence if he is not the perpetrator 0.5. Dividing the first number by the second yields an increase in the odds of guilt (from even odds) to 2:1 (or 0.67).\textsuperscript{56}

As a general matter, any likelihood ratio other than 1 yields a conclusion that the evidence is relevant. If the odds are greater than 1:1, then the evidence is relevant for proving the proposition at issue. If the probability of the evidence, assuming the proposition is true, is lower than the probability evidence, assuming the proposition is false, then the evidence is relevant for disproving the proposition at issue. Probative value is expressed under this conception based on how greater (or lower) than 1 the likelihood ratio is.\textsuperscript{57} A likelihood ratio of 1 means the evidence is irrelevant under this conception. For an example of irrelevant evidence under this conception, consider the following example:

Example # 2: In a criminal trial, there is evidence that the defendant supports liberal political candidates.

As Lempert explains, “[a]bsent 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

\textsuperscript{56} Id.

\textsuperscript{57} See Sources cited in supra note 51.
The odds of guilt remain what they were before considering the evidence and thus, under this conception, the evidence is irrelevant.

This notion of the likelihood ratio, to be sure, captures something intuitive and important about relevance. If the ratio is greater than 1, then when we multiply it by the prior probability, it will increase the number, and vice versa if the ratio less than 1. The likelihood ratio also serves as a useful tool for drawing the attention of students and courts to the issue of how likely evidence would be even if the proposition for which it is offered were false—an important, and sometimes overlooked, aspect of relevance and probative value.

But as a theory of the meaning of these concepts or as a normative theory for their application, the likelihood-ratio theory does not work. This is so because a ratio other than 1, or an increase in posterior probability more generally, is neither necessary nor sufficient for evidence to be relevant for proving the truth of a contested proposition at trial. As a methodological matter, if there are cases of middle-of-the-road logical (or “rational”) relevance and probative value that do not fit the dictates of the theory, then the theory does capture the “meaning” of these concepts (in philosophical jargon, it is not an analytically true claim about them), nor is it a normative standard by which we can measure their applications. Instead, the theory will sometimes coincide with relevance and probative value and sometimes not.

---

58 Lempert, supra note 3 at 1025

59 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 in the case and has no independent or direct relevance; general background details; res gestae). Moreover, if evidence truly is not relevant it should not be admitted over a relevance objection. See Fed. R. Evid. 402. It is sometimes claimed 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
In any given trial, there may be relevant evidence that each side claims supports their theory of what happened and that, from the perspective of a reasonable or rational jury, does not distinguish between the cases. Here is one example from a criminal trial transcript. The defendant, an inmate at a maximum security prison, is charged with two counts of battery on guards at the prison. The charges arise from an altercation between the defendant and several guards after the defendant refused to return a food tray in his cell. The prosecution’s theory was that the defendant charged and battered the officers when they opened the cell door to retrieve the tray. The defendant testified that one of the guards rushed in and began hitting him first, and his attorney argued in addition that even if the defendant made contact first with the officer he was acting in self defense.

In sorting out these competing explanations of what happened, one piece of information a reasonable jury might like to know is what precipitated the altercation. Why did the defendant withhold the tray in the first place? And both sides provide a number of details regarding this seemingly relevant evidence. Both attorneys mention during opening and closing arguments that the defendant had not received a package sent to him, and that after several attempts to speak with a sergeant and retrieve this package, the defendant refused to return his food tray.

---

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.

60 The case is referred to as “People v. Johnson” and is reprinted (with the names changed) in Allen, Kuhns, Swift & Schwartz, Evidence: Text, Problems & Cases 2-75 (4th edition 2006).

61 The Prosecution’s Opening: “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.

The Defense Opening: “the 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.
Multiple witnesses for each side testify to details about the package and the procedures at the prison for receiving them. Each side uses 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, or (2) the defendant was frustrated about not receiving the package, withheld his tray, and in response the guard charged at him (to retaliate or punish him for this behavior). Now, all this seemingly relevant evidence does not appear to distinguish between the two theories; as with the “supporting liberal candidates” example above, there is no reasonable basis for concluding that this evidence supports one theory over the other. In other words, the likelihood ratio is 1. Under the likelihood-ratio ratio conception, this implies the evidence is irrelevant (and a fortiori has no probative value) and thus should have been excluded. This would further imply that the jury should have been forced to choose between who started the altercation without giving them any information about the events that precipitated it. The relevance of this evidence—evidence that overlaps with both theories but does not distinguish them—provides a reductio ad absurdum for the likelihood-ratio approach.

Prosecution Closing: “for some reason inmate Johnson chose to draw the line that day over some package that he claimed he wasn’t getting.” Id. at 66.

Defense Closing: “Johnson chose to draw the line. I agree. He drew the line. “I want something done; I want to see a seargent . . . And its obvious form the testimony that if he just kept on drawing the line at some point he would have seen a seargant. That doesn’t mean Johnson made a choice to go out and start battering the officer or officers.” Id at 71.

62 [Quote Testimony]

63 See supra note 61.

64 Id.

65 See Lempert, supra note 3, at 1025, and notes 57 and 58 supra.

66 Fed. R. Evid. 402
The prevalence of such evidence may be a matter of dispute, but no doubt at least some examples exist. And the failure of the likelihood-ratio approach to explain such evidence is problematic for the theory. This problem can be brought into focus with a couple of more stylized examples.

Example # 3: In a murder case, a witness will testify that he saw the defendant in the area the night of the crime. The defendant’s theory was that he was in the area to visit his mother. If the judge concludes that a reasonable jury would have to conclude this evidence is just as likely if he is the murderer as if he was simply visiting his mother, then it has a ratio of 1 and should be excluded as irrelevant.67

Example # 4: 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. Assuming the judge concludes there is no reasonable basis to distinguish the defendant from his twin (i.e., it has a likelihood ratio of 1), then witness’s testimony should be excluded as irrelevant.

Pace the likelihood-ratio theory, the evidence is logically relevant in both of cases, despite a likelihood of 1. In both cases, the evidence provides a reasonable fact-finder with important information in drawing inferences about what happened. The fact that both sides can offer equally plausible explanations of how that evidence is consistent with their theory, does not make the evidence irrelevant. Thus, the likelihood-ratio theory is false as an analytic claim about the meaning relevance and as a normative theory of relevance. The upshot is that evidence can

67 Example of Ronald J. Allen in, supra note 1.
be logically relevant even if it does not distinguish between the parties’ cases or theories of what happened.\textsuperscript{68}

More generally, an increase in the posterior probability of a proposition is neither necessary nor sufficient for evidence to be logically relevance for proving that proposition to be true. To illustrate it is not \textit{necessary}, consider the following example:

Example # 5: Suppose the issue is whether Victim died of poisoning. It is known that at 12:00 pm on the day he collapsed and died, he ingested a poison concealed in his food that is fatal for 90 percent of the people who ingest it. Suppose there is also evidence that: at 12:30 pm he ingested a second poison that completely counteracts the first poison; however, it is fatal for 80 percent of the people who ingest it. Is evidence of the second poison relevant for proving that Victim died of poisoning?\textsuperscript{69}

Evidence regarding the second poisoning reduces probability that Victim died of poisoning from the prior probability without the evidence, but the evidence is relevant for proving Victim died of poisoning.

To show an increase is not \textit{sufficient}, consider the following example:

Example # 6: Suppose 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. Suppose we know that Defendant purchased 1 of the 1000 tickets and Rival purchased 99

\textsuperscript{68} For an additional example of such evidence, see example by Peter Tillers in supra note 1.

tickets. Suppose there is also evidence that: the other 900 tickets were never sold and have been accounted for. Is evidence about the other 900 tickets relevant to proving Defendant is guilt? Or is it evidence that further supports Defendant’s theory? Or neither?\textsuperscript{70}

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 900 of tickets. Despite this increase in posterior probability, the evidence is not necessarily relevant for proving Defendant’s guilt. Indeed, this is evidence that Defendant will seek to introduce to better support Defendant’s theory that Rival is the true culprit.\textsuperscript{71}

Finally, one additional problem with quantifying \textit{probative value} at the micro-level is the problem of reference classes.\textsuperscript{72} To use a likelihood ratio to quantify probative value typically involves placing the evidence is 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”:

Example #7: If research established that the behavior is equally common among abused and non-abused children, then its likelihood ratio would be one, and evidence of that

\begin{flushleft}
\textsuperscript{70} Id.
\end{flushleft}

\begin{flushleft}
\textsuperscript{71} Perhaps one might also argue the evidence is irrelevant because it keeps the same relative ratio of tickets between Defendant and Rival. But either side would likely try take advantage of evidence of the unknown 900 tickets and thus this may be another case of overlapping evidence.
\end{flushleft}

\begin{flushleft}
\end{flushleft}
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.\textsuperscript{73}

These conclusions depend on a number of further auxiliary assumptions (which may or may not be true). The evidence at issue is certain behavior. Even if this behavior is equally common about both groups of children, it might nevertheless be highly probative in a given case if, for example, abused children exhibiting this behavior possess and non-abused children lack and the particular child at issue possesses (or lacks) this characteristic. Similarly, even if the behavior is 1000 times more likely in abused children, the probative value may be nevertheless be minimal if the child possesses (or lacks) an additional characteristic that places the child in the group of non-abused children who exhibit the behavior. In these examples, the evidence and the likelihood ratios remain constant, but the probative value may vary dramatically. The probative value is thus not the likelihood ratio.

To illustrate the reference-class problem more starkly, consider this famous example from Tversky and Kahneman:

Example # 8: A cab was involved in a hit and run accident and two cab companies, Green and Blue, operate in the city. The evidence: 85% of the cabs in the city are Green and 15% are Blue. What is the probability the cab involved in the accident was Blue?\textsuperscript{74}

The initial response may be 0.15. But why? The distribution of cabs \textit{in the city} is just one class to which this event is a member. The reference-class problem arises because the data depend on

\textsuperscript{73} McCormick, Evidence 277 (Strong ed., 5\textsuperscript{th} ed. 1999).

\textsuperscript{74} Amos Tversky & Daniel Kahneman, \textit{Evidential Impact of Base Rates}, in Judgement Under Uncertainty: Heuristics and Biases (Kahneman, Slovic & Tversky eds. 1982)
placing an event or particular piece of evidence in one class, but the event or evidence is a
member of a virtually infinite number of sets, each with differing rates. For example, suppose
that 85% of the cabs on the street are Blue, and 85% of the cabs running at that time of day are
Green, 85% of the cabs in the service area (which includes several towns) are Blue, and so on.
What is the probability now that the one in the accident was Blue? Inferences from evidence in a
particular reference class involve a host of auxiliary assumptions about, for example, the
appropriateness and homogeneity of that particular class, the likely distribution on other classes,
and several other possible issues. The upshot, in short, is that we cannot take data from one
reference class as a normative standard by which to quantify and judge probative value or as a
theoretical explanation of the content of the concept.

Cataloging these limitations to probabilistic conceptions of the micro-level of proof
reveals criteria that a successful theory ought to meet. It must be able to explain whether and
why evidence is relevance and probative (or not), including evidence that “overlaps” with—but
does not distinguish between—each side’s theory of the case. It must also explain when and
why the relevance and probative value of evidence deviates from the answers implied by cardinal
probabilistic assessments (increases or decreases of posterior probabilities, and likelihood ratios
depending on particular reference classes). Finally, it must provide criteria for guiding,
constraining, and evaluating judgments regarding particular items of evidence.

III. The Macro-Level of Evidence Theory: Decision Standards and Evidential Sufficiency

The macro-level constraint on evidence theory posits that a theory of evidence ought to
explain when evidence is sufficient to satisfy a particular decision standard. This account ought
to explain these outcomes in light on the goals underlying the standards—which include
accuracy and allocating the risk of factual errors—and provide criteria for guiding and constraining applications and for evaluating which applications are reasonable and unreasonable. The integration constraint further provides that macro-level theoretical accounts must cohere with a theoretical conception of evidence and proof at the micro-level.

Probabilistic conceptions at the macro-level are commonplace. It is typically posited or presupposed that “preponderance of the evidence” equals proof of 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. As evidence is sufficient to prove a proposition when the probability of the proposition, given the evidence, exceeds the decision threshold. This is a fairly standard interpretation, about which much has been written. As at the micro-level, this conception capture some important aspects of the issues but, again, it is also deeply problematic as a theoretical account or normative standard for meeting the underlying goals of the standards (i.e., accuracy and allocating risk of error).

First, a caveat. The literature on probabilistic conceptions of decision standards is large, and I do not canvass all the issues, problems, and counter-responses that have been proposed. Instead, I’ll focus on three aspects of probabilistic conceptions of decision standards that make them particularly problematic in light of the theoretical and methodological criteria outlined in Part I. These aspects have not been the focused of sustained discussion in the literature.

See, e.g., Bell, supra note __; see also note __.

See, e.g., the sources cited in note __.

See Pardo, supra note __.
The first problem with probabilistic conceptions concerns the lack of guidance in applying the standards and the lack of criteria by which to determine which applications are reasonable and unreasonable. To apply a probabilistic decision standard, one must also quantify the evidence.\(^78\) This could be done based on “objective” data\(^79\) or “subjective” beliefs. The objective data do not exist for most items of evidence in any given trial.\(^80\) If they did and such standards were feasible, then the only reasonable decisions would be the objective answer, which would render fact-finders largely superfluous. So we must largely rely on subjective probability assessments. Unfortunately, probabilistic decision standards provide little guidance on how to proceed. Suppose you are on a jury in a civil case applying the preponderance standard, and you 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 jury says it is 0.2 likely? Are they both reasonable? Is .99? Is .01? How can we tell?\(^81\)

This limitation affects more than 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 depend on whether a “reasonable” jury could find for the non-moving party.\(^82\) What is

\(^78\) This could be done precisely or with vague judgments about whether the evidence surpasses a threshold.

\(^79\) But note that objective data would still be subject to the reference-class problem. See supra at __.

\(^80\) Even if they did exist, the computational complexity may be too great for most trials. See Ronald J. Allen, Constitutional Adjudication, the Demands of Knowledge, and Epistemological Modesty, 88 Nw. U. L. Rev. 436, 444 (1993).

\(^81\) Moreover, there is no reason to think subjective beliefs would be truth conducive. See Clermont, supra note __, at 481-82.

\(^82\) See supra at p. __.
“reasonable” or not depends on whether an outcome is reasonable or not given the burden of proof and underlying decision standard. Thus, decisions to grant or deny either motion presuppose some conception of the decision standards and what they require. Whether parties will get to trial or not in the first place, whether the case will go to the jury at trial, and whether a jury verdict will be overturned all depend on principled applications of the decision rules. But standards based on subjective probabilistic assessments provide no criteria or guidance how to proceed and no criteria or guidance for separating reasonable and unreasonable applications.

Suppose now that 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 assessments.” First, there is no way to know what any actual jury would do without letting them decide. More importantly, because the criteria are subjective, there is no independent basis for concluding which applications are reasonable or 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 could find guilt “beyond a reasonable doubt” for the elements of charged crimes.\(^{83}\) Thus, whether criminal defendants will go to trial in the first place, whether the cases will go to a jury, and whether convictions will be overturned should as a doctrinal matter depend on principled applications of what is required by the “beyond a reasonable doubt” standard. But subjective probabilistic assessments again provide no guidance or criteria for principled applications.

\(^{83}\) See supra at p. ___.
The second limitation on subjective probabilistic conceptions concerns the integration constraint. There is an inconsistency between macro-level subjective probabilistic decision standards and conventional micro-level probabilistic conceptions of relevance and probative value. At the micro-level, relevance and probative value depend on more than the subjective assessments of individual fact-finders. This is consistent 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. Moreover, an application of any decision standard is essentially an assessment of the probative value of the evidence as a whole, which as we just seen presupposes the possibility of a reasonable or unreasonable application. If there is more to the probative value of evidence as a whole, then subjective probability conceptions at the macro-level fail to explain the macro-level proof process and fail to provide the requisite normative guidance. (And the lack of objective data for most items of evidence means that “objective” probabilistic conceptions of the standards cannot make up this theoretical gap.) Something more—and more plausible—is needed. If, by contrast, a probabilistic theorist argues that there is in fact nothing more to probative value at the macro-level beyond subjective assessments, then at the micro-level the theorist must concede that relevance and probative value also depend solely on subjective assessments. This concession, for the reasons discussed in Part III, would fail to explain relevance and probative value, and it would fail to provide any criteria for admissibility decisions. Either way, the probabilistic theorist is in a bind.

---

84 If it did not, judges would have no basis to exclude evidence on grounds of relevance or Fed. R. Evid. 403.

85 Or this apparent inconsistency must be explained away.
The third problem arises from the non-comparative nature of probabilistic conceptions that focus on whether the evidence of the party with the burden of proof surpasses a fixed threshold. If the standards are understood in this way, it would frustrate the underlying goals of the rules. The probabilistic rules compare the probability of a proposition with its negation, but they should be comparing the competing propositions put forward by each side. This is cryptic.

Let me explain. 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.\(^{86}\) Now, suppose a plaintiff offers a theory of what happened that a reasonable jury concludes is 0.4 likely and the defendant offers a theory of the case they concluded was 0.2 likely.\(^ {87}\) 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 defendant’s alternative account. This is does not make sense if what the legal system attempting with the preponderance standard to minimize errors (in which case it should go with more likely) or equalize the risk of errors (plaintiff should not bear the risk of error for all the unknown probability space). 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.\(^ {88}\)

As at the micro-level, cataloging these limitations provides some insight into what a successful theoretical account of evidence and proof would look like at the macro-level. Such a

\(^{86}\) See supra at p. ___

\(^{87}\) For purposes of this illustration, I am putting aside the other problems with such conceptions.

\(^{88}\) This is so without even considering the various “conjunction” problems, which may further shift the risk of error in undesirable ways. The conjunction problems arise because the decision standards apply to individual elements rather to cases as a whole. See Ronald J. Allen & Sarah A. Jehl, Burdens of Proof in Civil Cases: Algorithms vs. Explanations, 2003 Mich. St. L. Rev. 893 (2003). I put aside these additional complexities for purposes of this discussion.
theory should provide possible and feasible criteria and guidance for applying the standards, and for evaluating which applications are reasonable; it should cohere with an account at the micro-level; and it should explain how the applications it implies are consistent with the underlying goals of the rules.

IV. An Explanatory Conception of Evidence and Proof

Any theoretical conception of some social practice as rich and complex as juridical proof is likely to have some difficulties. So, without a robust alternative, perhaps the probabilistic conceptions are the best we can do?89 I think we can do better. I conclude this essay by sketching the alternative conception and how it better meets each of the three theoretical constraints articulated in Part I.90

The alternative conception examines explanatory relationships between propositions and evidence. Rather than focus on probabilistic relationships—whether the probability of propositions (or their negations) given evidence, or the probability of evidence given propositions (or their negations)—the explanatory conception examines the competing explanations of the evidence and events put forward by the parties and developed by fact-finders. More specifically, the critical issue is whether particular explanations put forward by the parties or developed by fact-finders, if true, would be better or worse at explaining the evidence and the

89 And, as I have conceded, the probabilistic conceptions do capture and illuminate some important aspects of the structure and nature of juridical proof.

90 The explanatory conception has been developed by myself and Ron Allen, both jointly and in our individual work. See Michael S. Pardo & Ronald J. Allen, Juridical Proof and the Best Explanation, 27 Law & Phil. 223 (2008); Allen & Pardo, supra note ___; Pardo, supra note ___; Pardo, supra note ___. The focus on explanatory criteria (or inference to the best explanation) 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 (1994).
underlying events (than the alternatives). Under this conception, juridical proof is an abductive process of inference to the best explanation.\textsuperscript{91} This inferential process occurs in two stages: generating potential explanations and then the selection of one as best or acceptable, depending on the decision standard.\textsuperscript{92} For example, the quality of an explanation needed to satisfy the “beyond a reasonable doubt” or “clear and convincing evidence” standards will be greater than under the preponderance standard.

\textit{The Micro-Level Constraint}

Under this conception, relevance and probative value depend on whether evidence is part of a party’s explanation or a challenge to the other side’s explanation. Unlike the likelihood-ratio theory specifically or probabilistic conceptions generally, the explanatory conception thus accounts for evidence that is relevant but does not distinguish between cases. Evidence is relevant if it is part of one side’s explanation (or put forward by the other side as a challenge to that explanation), regardless of whether it is also part of the other side’s explanation. And its probative value depends on the role it plays in each side’s explanations or in distinguishing between them.

This conception thus explains the examples presented in Part III of “overlapping” evidence that supports both sides but does not distinguish the cases. The evidence about the package in the transcript; the evidence about the defendant’s whereabouts in Example # 3; and

\textsuperscript{91} For a general overview, see Peter Lipton, Inference to the Best Explanation (2004).

\textsuperscript{92} See Pardo & Allen, supra note __. Explanations involve both “acts of explaining” and the explanations themselves, which concern epistemic relationships between propositions. Explanations typically provide answers or understanding about another proposition: e.g., why something happened. On the nature of explanations, see Lipton, supra note 91, at 21-29; Peter Achinstein, The Nature of Explanation 74-102 (1983).
the eyewitness testimony in Example # 4 are part of each side’s explanation of what occurred and thus relevant.\textsuperscript{93} This conception also explains why the statistical evidence in Examples # 5 and # 6 are relevant, despite their deviations from what is implied by a cardinal probabilistic conception.\textsuperscript{94} The second poisoning in Example # 5 is part of the prosecution’s explanation of what occurred, and the evidence about the 900 tickets in Example # 6 supports the defendant’s explanation that Rival was the culprit. Finally, this conception avoids the limitations of the reference-class problem in ascribing probative value to evidence based on its membership in a particular class. In any case, the critical question of what explains the evidence isolates what is important about the evidence without the need to quantify based on reference classes. For example, in Example # 7 the issue is what best or better explains the child’s behavior (and the statistical evidence will be relevant to answering this question). Example # 8, by contrast, provides statistical information with less of an explanatory connection. Although relevant, this lack of an explanatory connection implies that the evidence is nowhere near as probative as has been suggested.\textsuperscript{95}

\textit{The Macro-Level Constraint}

Under this conception, decision rules can be articulated based on explanatory criteria. For example, under the preponderance rule, fact-finders should infer the best explanation and the party with the burden of proof ought to win if that explanation includes the formal elements, and the party with the burden ought to lose if the best explanation does not include one or more of

\begin{itemize}
\item \textsuperscript{93} See supra at p. ___.
\item \textsuperscript{94} See supra at p. ___.
\item \textsuperscript{95} See Pardo & Allen, supra note 7 (critiquing the conclusion by Tversky and Kahneman that subjects err when they discount this evidence)
\end{itemize}
the elements. This may better instantiate the goals of minimizing errors and equalizing the risk of error among the parties—better explanations more likely to be true and each bears an equal risk of error. Moreover, because the standard is comparative, neither side benefits from the unknown probability space.

The explanatory preponderance standard is largely consistent with the empirical Story Model in that many of the criteria that make a “story” persuasive also make an explanation better, and vice versa (consistency, coverage, completeness, simplicity, absence of gaps, fit with background knowledge, and so on). However, there are important differences: explanations may be general and disjunctive, depending on how the parties choose to contrast what happened, rather than the singular, specific, and integrated stories posited by the Story Model. Moreover, under the higher decision standards, the quality of an explanation needed to satisfy the standards rises to beyond merely being better than the alternatives. Under the “clear and convincing” standard, the explanation must be substantially better than the alternatives, and the standard for “beyond a reasonable doubt” 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). Raising the quality of an explanation needed to satisfy these standards will

---

96 See Pardo, supra note __. This approach also resolves any potential “conjunction” issues. See supra note 88.

97 For example, in a res ipsa loquitur case a general explanation may be sufficient.

98 For example, an explanation that “X, Y, or Z happened” may be sufficient if this disjunctive explanation is itself better than the alternative explanation (which may or may not itself be disjunctive).

99 When the two diverge, the explanatory criteria provide normative guidance.

100 See Pardo, supra note __.
correspondingly shift the risk of error in ways consistent with the goals of the rules and ought to implement these goals better than subjective probability assessments.\textsuperscript{101}

The explanatory criteria underlying this conception also provide some guidance and criteria beyond both the empirical Story Model and the theoretical probabilistic conceptions for fact-finders and reviewing courts. There are objective ways to examine and discuss explanatory relationships between evidence and competing propositions, even if there will no doubt be disagreement in some cases about which explanations are better or worse or what makes an explanation better or worse.\textsuperscript{102} Other things being equal, a consistent explanation is better than a contradictory one; an explanation that accounts for more evidence and most important items of evidence will be better than one that cannot account or explain them; an explanation that adheres with background knowledge and does not require extraordinary assumptions will better than one that is inconsistent with background knowledge and requires unrealistic assumptions. And so on. These explanatory criteria connect with deeper doctrinal issues on civil and criminal sides: they provide criteria for reasonable and unreasonable applications and may also provide a burden of explanation and vocabulary for reviewing courts.\textsuperscript{103}

\textsuperscript{101} Although the standards are not as precise as cardinal probability thresholds, they may be better understood by fact-finders and better fit with the reasoning processes of jurors. These features may make “explanatory” based jury instructions more effective than current instructions. For more on this issue, see Pardo, supra note __.

\textsuperscript{102} See Clermont, supra note __, at 482 (“the 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.” Moreover, some areas of law—e.g., employment discrimination—already have frameworks in place for evaluating competing explanations. See Reeves v. Sanderson Plumbing Prods., Inc., 530 U.S. 133, 149-50 (2000).

\textsuperscript{103} 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.
The Integration Constraint

The micro- and macro-levels of the explanatory theory cohere in a straightforward way, and this provides another virtue for this theoretical conception. Explanations as a whole animate the macro-level, and the quality of these explanations depends on the relevance and probative value of evidence at the micro-level. Likewise, relevance and probative value at the micro-level depend on whether 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 evidence at the micro-level depends on the significance of the evidence to these macro-level explanations. The same explanatory criteria provide guidance, constraint, and a basis for evaluation at both levels.