

Spotting a Preponderance of the Evidence in the Wild: Inference to the Best Explanation and Sufficiency of the Evidence

Craig R. Callen *

I.ON QUANTIFICATION.....	1524
II.INFERENCE TO THE BEST EXPLANATION	1532
A. Everyday Life and Sherlock Holmes.....	1532
B. Inference to the Best Explanation in the Pardo/Allen Theory 1536	
1. Professors Pardo and Allen on the Basics of Inference to the Best Explanation.....	1536
2. Inference to the Best Explanation and Sufficiency of Evidence	1537
3. Objections to the Pardo/Allen Theory.....	1539
4. Pardo and Allen on the Implications of the Pardo/Allen Theory.....	1542
IV.RESEARCH ON PROBLEM-SOLVING, THE STORY MODEL AND ALTERNATIVE VIEWS OF ABDUCTION AS SOURCES OF DESCRIPTION.....	1543
A. Stories, Cognition, and Confidence in Judgment	1545
1. The Story Model.....	1546
2. Oblique-Stories.....	1548
B. How Theories of Explanation Can Account for Lack of Confidence.....	1549
C. The Utility of Burdens and Sufficiency Tests	1552
1. As a Guide for Jurors.....	1552
2. Sufficiency Tests and Errors of Law	1555
D. Cognition, The Pardo/Allen Theory, and the Court’s View of Fact Finding	1557
1. The Fact Finder’s Story in <i>St. Mary’s</i>	1557
2. Celotex, Reeves and The Pardo/Allen Theory.....	1559
3. When Evidence or Explanations Are Lacking.....	1560
V.AN ALTERNATIVE	1563
A. Optimization Is Not An Option	1564

* Craig R. Callen (1950-2008), Professor, Michigan State University College of Law.

B. Employing ‘Story’ or ‘Explanation’ in Preponderance Instructions	1566
CONCLUSION.....	1567

When facts are in dispute, the enforcement of legal rights depends on standards of proof. Such standards tell fact finders (who are seldom completely certain of the facts) how to resolve issues of fact even though they are uncertain. The typical standard in civil litigation is that the party to whom the burden of persuasion is allocated must prove its contention by a preponderance of the evidence. Wigmore doubted that verbal explanations of the preponderance of the evidence standard were necessary, or even helpful. He said that the “simple and suggestive phrase [preponderance of the evidence] has not been allowed to suffice; and in many precedents sundry other phrases—‘satisfied,’ ‘convinced,’ and the like—have been put forward as equivalents, and their propriety as a form of words discussed and sanctioned or disapproved, with much waste of judicial effort.”¹ If his confidence in the clarity of the term was ever justified, it is no more. “Preponderance” is at best an uncommon word nowadays. Moreover, issues such as whether mathematical models should guide jury deliberations put considerable strain on that “simple and suggestive phrase.”

Professor Ronald Allen has championed a comparative interpretation of the preponderance test.² In a recent article, *Juridical Proof and the Best Explanation*,³ Professor Michael Pardo and Professor Allen offer a new variation of the comparative model. They contend that, with respect to the preponderance standard, judges should instruct jurors to “select the best explanation of the evidence (or the most plausible version of the litigated events) and that something has been proven by a preponderance of the evidence if it is part of their selected explanation or version of events.”⁴ With respect to summary judgments and judgments as a matter of law, they take the position that “a judge ought to grant such a motion only if a jury would have to find one side’s explanation more plausible than the other side’s

¹ 9 JOHN HENRY WIGMORE, EVIDENCE § 2498, at 325–26 (3d ed. 1940).

² See, e.g., Ronald J. Allen, *A Reconceptualization of Civil Trials*, 66 B.U. L. REV. 401, 425–34 (1986); Ronald J. Allen, *Rationality, Algorithms and Juridical Proof: A Preliminary Inquiry*, 1 INT’L J. EVIDENCE & PROOF 254, 273 (1997) (“[W]hen a plaintiff produces evidence of some state or states of affair that are favourable to its claim, that is sufficient, until the defense produces some to the contrary, and so on. This in turn means that legal proof is comparative: the question is the relative plausibility of the parties’ cases; the question is not the cardinal probability, conceived as a relative frequency or a subjective belief state, of a certain state of affairs.”).

³ Michael S. Pardo & Ronald J. Allen, *Juridical Proof and the Best Explanation*, 27 LAW & PHIL. 223 (2007).

⁴ *Id.* at 266–67.

explanation.”⁵

Even though Pardo and Allen advocate a comparative understanding of the preponderance test, they accept a “typical” pattern jury instruction explanation of the preponderance standard: “When you have considered all the evidence in the case, you must be persuaded that it is more probably true than not true.”⁶ That does not sound particularly like a comparative standard. It seems to fit better with a more conventional explanation of the preponderance of the evidence standard, “Preponderance is not a comparative standard as such: The question is not whether plaintiff’s case is better than defendant’s, but whether the evidence makes the points that the plaintiff must prove more probable than not.”⁷ The prime significance of the difference between the two is that Pardo and Allen’s suggested instruction on the preponderance of the evidence test would require jurors to find in favor of the party with the burden of proof when the best explanation favored that party, even though the jurors considered that explanation (or the evidence that supported it) insufficient to warrant a verdict.⁸

⁵ *Id.* at 264–65.

⁶ *Id.* at 266 (quoting FEDERAL CIVIL JURY INSTRUCTIONS OF THE SEVENTH CIRCUIT 34).

⁷ “[T]his standard means the jury is persuaded (acting as reasonable persons) that the points to be proved are more probably so than not. . . . At least in theory, it is satisfied if the factfinder believes by the thinnest conceivable margin that the points to be proved are so, and anything less would not be a standard of proof at all. . . .” 1 CHRISTOPHER B. MUELLER & LAIRD C. KIRKPATRICK, FEDERAL EVIDENCE § 3:5, at 437–38 (2007) (footnotes omitted). Often, under a preponderance of the evidence standard, it is assumed that the trier of fact piles up the evidence arguably on the plaintiff’s side and the evidence arguably on the defendant’s side and determines which pile is greater. *Cf.* *In re Winship*, 397 U.S. 358, 367–68 (1969) (noting “the accuracy of the observation of commentators that ‘the preponderance test is susceptible to the misinterpretation that it calls on the trier of fact merely to perform an abstract weighing of the evidence in order to determine which side has produced the greater quantum, without regard to its effect in convincing his mind of the truth of the proposition asserted’” (quoting Norman Dorsen & Daniel A. Reznick, *In Re Gault and the Future of Juvenile Law*, 1 FAM. L. Q. no. 4 (1967) at 26)). In fact, a more accurate notion of the preponderance of the evidence standard is “evidence which as a whole shows that the fact sought to be proved is more probable than not.” *United States v. Montague*, 40 F.3d 1251, 1254–55 (D.C. Cir. 1994) (some citations and footnotes omitted). For a slightly different approach, see, for example, 2 MCCORMICK ON EVIDENCE § 339 at 484 (Kenneth S. Broun, et al. 6th Practitioners ed. 2006) (“The most acceptable meaning to be given to the expression, proof by a preponderance seems to be proof which leads the jury to find that the existence of the contested fact is more probable than its nonexistence. Thus the preponderance of evidence becomes the trier’s belief in the preponderance of probability.”) (citing MODEL CODE OF EVIDENCE RULE 1.3; EDMUND MORGAN, SOME PROBLEMS OF PROOF 8485 (1956)).

⁸ Neither of those two approaches fits well with the heuristic approach that many courts and some commentators use, conceiving of a preponderance of evidence as anything greater than 50%. Where “50%” is merely an analogy to a state of cognitive agnosticism (where the trier finds no party’s evidence and argument superior to the others’), it becomes very problematic when one considers questions such as those in the debates over the use of mathematical models at trial and the sufficiency of so-called “naked” statistical evidence to support a verdict.

Moreover, their approach to summary judgments and other tests of the sufficiency of the evidence does not seem to fit with the Supreme Court's opinions in *Celotex Corporation v. Catrett*,⁹ and *Reeves v. Sanderson Plumbing Products, Inc.*¹⁰ Each of those cases permits the court to enter judgment against the nonmoving party for insufficiency of evidence without considering evidence and explanations in favor of the moving party.¹¹ Of course, it is not unknown for conventional approaches to be wrong, or confused, and Pardo and Allen undoubtedly think that non-comparative approaches to the preponderance tests are in error. This paper will argue that the conventional understanding, viewed in light of studies of inference under uncertainty, seems superior to the approach that Professors Pardo and Allen take.

These flaws are unfortunate, because the theory Professors Pardo and Allen advance is very interesting, particularly its reliance on inference to the best explanation, a form of non-deductive reasoning.¹² That is to say that inference to the best explanation does not guarantee the truth of its conclusions,¹³ although conclusions based upon it may well be adequately warranted.¹⁴ (I will use the term "the Pardo/Allen Theory" to distinguish their theory from the actual practice of inference to the best explanation, or from scholarly commentary on that process with which they may not agree.)¹⁵ They view the process of inference to the best explanation in litigation as occurring in two steps. The first is generation of potential explanations of the evidence—typically a process that the parties to a dispute initiate, but one to which the fact finders contribute. Then the triers of fact select "the best explanation from the list of potential ones as an actual explanation or the truth."¹⁶ While they rely on the reader's intuitive sense of what an explanation is, they note that the context in which one makes

⁹ 477 U.S. 317 (1986).

¹⁰ 530 U.S. 133 (2000).

¹¹ Merely asserting that the non-moving party had insufficient evidence could not count as an explanation, or else making an inference to the best explanation would merely mean deciding in favor of the party with the best argument. In that case, inference to the best explanation would not add anything to what we already know.

¹² For illustrations see *infra* Section II, part A.

¹³ Pardo & Allen, *supra* note 3, at 227.

¹⁴ Gilbert H. Harman, *The Inference to the Best Explanation*, 74 PHIL. REV. 88, 90–91 (1965).

¹⁵ On possible differences between the Pardo/Allen theory and the process of inference to the best explanation, see, for example, Larry Laudan, *Strange Bedfellow: Inference to the Best Explanation and the Criminal Standard of Proof*, 11 INT'L J. EVIDENCE. & PROOF 292, 300 (2007) (contending that the reasonable doubt standard is not, contrary to Pardo and Allen's theory, consistent with inference to the best explanation because the reasonable doubt standard prohibits a conviction where there is a reasonable explanation in favor of the defendant, regardless of whether it is the best explanation)

¹⁶ Pardo & Allen, *supra* note 3, at 229.

inferences will tend to suggest the characteristics of acceptable explanations.¹⁷

Juridical Proof and the Best Explanation analyzes, as the title suggests, the entire process of proof. In the process it naturally devotes significant attention to the standards of proof. I will focus on the preponderance standard since (i) it seems to be the standard of proof that would reflect inference to the best explanation, if any does,¹⁸ and (ii) courts often pose it in terms that might sound quantitative, for example, as requiring the party bearing the burden to prove that its contentions are more likely than not the case, or more likely so than not so.¹⁹

Formulations of the preponderance test in terms of likelihood or probability, such as “When you have considered all the evidence in the case, you must be persuaded that it is more probably true than not true,”²⁰ have tempted advocates of the use of mathematical models of evidence over the last few decades to make the argument that statistical norms should govern fact finding or formulation of evidentiary rules. Pardo and Allen’s piece is, in significant part, a criticism of such models. Issues about the preponderance standard arose most starkly in discussions of the sufficiency of so-called “naked” or exclusively statistical evidence,²¹ often presented through the “blue bus” hypothetical. Those discussions raised two specific questions that are particularly salient here: First, what sort of superiority of evidence (or inferences from it) is necessary for a preponderance of evidence? Second, and related, what sort of superiority of evidence or likelihood of inferences is necessary to justify a directed verdict or summary judgment? I will argue that the Pardo/Allen theory’s response to each of them is flawed.

¹⁷ Pardo & Allen, *supra* note 3, at 229 & n.17.

¹⁸ See, e.g., Laudan, *supra* note 15 at 300 (arguing that the reasonable doubt standard does not reflect inference to the best explanation in that the reasonable doubt test requires a verdict for the defense if any reasonable explanation would be in defendant’s favor).

¹⁹ 1 MUELLER & KIRKPATRICK, *supra* note 7, §3:5, at 437 n.4 (“It is irresistible [to] draw an analogy to likelihood or comparative probability: 51-49 (or for that matter 50.001-49.999) is enough.”).

²⁰ Pardo & Allen, *supra* note 3, at 266 (quoting FEDERAL CIVIL JURY INSTRUCTIONS OF THE SEVENTH CIRCUIT 34). Cf. *Bourjailly v. United States*, 483 U.S. 171, 175–76 (1987) (discussing examples of “more likely than not” standard in preliminary factual issues for the court).

²¹ See, e.g., Daniel Shaviro, *Statistical-Probability Evidence and the Appearance of Justice*, 103 HARV. L. REV. 530, 543–44 (1989); David H. Kaye, *Naked Statistical Evidence*, 89 YALE L. J. 601 (1980) (reviewing MICHAEL FINKELSTEIN, *QUANTITATIVE METHODS IN LAW: STUDIES IN THE APPLICATION OF MATHEMATICAL PROBABILITY AND STATISTICS TO LEGAL PROBLEMS* (1978)).

The Pardo/Allen theory has a great deal going for it. Juries can seldom be certain about the truth.²² Their ability to resolve questions of fact under such conditions of uncertainty depends on their selection and application of cognitive strategies they have learned from their experience given limited availability, time and cognitive resources. Pennington and Hastie's work on the story model was designed to delineate one such strategy, central to understanding discourse, i.e., to the jury's comprehension and assessment of evidence,²³ on which the Pardo/Allen theory relies, in part. Moreover, it is undoubtedly true that jurors, who are at least outwardly passive, often rely heavily on the parties' gathering of evidence and formulation of theories, or explanations of the evidence. When jurors have no reason to question the adequacy of the evidence or of the hypotheses that the parties have constructed to explain the evidence, then it makes perfect sense for jurors to compare those hypotheses in the process of reaching their decision. In addition, Professors Pardo and Allen make the telling point that more formal theories of evidence and inference, in particular Bayes' Theorem, depend in unacknowledged ways on narrative and other accounts of causation. Finally, their emphasis on comparison does not require that the fact finder consider all of the evidence that might bear on the ultimate issue. Exhaustive evaluation of evidence is utterly unworkable.²⁴

Even so, Pardo and Allen's analysis of the preponderance test seems to fall short. Research on human cognition suggests that successful decision makers employ strategies that sometimes call for an assessment of whether the information they have on an issue is sufficiently complete to support a decision, and to apply the default rule, or seek more information, when the data at hand seem insufficient.²⁵ Pardo and Allen's position bars

²² I will use "jury," "juror" or "jurors" to refer to fact finders in general, unless the context clearly indicates otherwise.

²³ See, e.g., Nancy Pennington & Reid Hastie, *A Cognitive Theory of Juror Decision Making: The Story Model*, 13 CARDOZO L. REV. 519 (1991).

²⁴ See Craig R. Callen, *Rationality and Relevancy: Conditional Relevancy and Constrained Resources*, 2003 MICH. ST. L. REV. 1260-71.

²⁵ Indeed, the story model itself posits that fact finders will consider the completeness of evidence when asking whether they have enough confidence in the stories they construct to support a verdict. Moreover, at least some scholars take the position that inference to the best explanation requires that an explanation or hypothesis meet a threshold requirement of plausibility (regardless of whether there are plausible competitors) before it can be considered "best." Laudan, *supra* note 15, at 304. Setting aside questions about the adequacy of an information base for drawing conclusions, Josephson argues that inference to the best explanation (or, as it is sometimes called, abduction) suggest that it must take account of two possible explanations even if the parties fail to offer them: (i) that the evidentiary basis is false (or just "noise") and (ii) that some new explanation not yet formulated will be better. John R. Josephson, *On the Proof Dynamics of Inference to the Best Explanation*, in *THE DYNAMICS OF LEGAL PROOF: COMPUTATION, LOGIC AND COMMON SENSE* 287, 293 (Peter Tillers & Marilyn MacCrimmon eds. 2002).

2018]

SPOTTING A PREPONDERANCE

1523

consideration of the adequacy of evidence to support an explanation, so long as one explanation seems to be better than any competitor. It is also inconsistent with any requirement that fact finders consider whether some unarticulated explanation might be better. Accordingly, as a descriptive theory, it cannot fit with the extant summary judgment and judgment as a matter of law standard, nor, in a real sense, does it seem consistent with the requirement of proof by a preponderance.

The main point of this paper is to show that inference to the best explanation provides an inadequate model of the preponderance standard if it omits consideration of whether the plaintiff²⁶ has offered sufficient evidence and an adequate explanation to justify a finding in its favor, regardless of the theories and evidence in defendant's favor. Moreover, the fact that the explanations in favor of plaintiff are better is necessary, but not always sufficient to warrant a verdict in its favor. In contrast, research on human problem solving (in conjunction with inference to the best explanation) strongly supports judicial examination of sufficiency of the evidence (primarily through motions for summary judgment and judgment as a matter of law) and jurors' consideration of the adequacy of evidence and inferences to support a verdict.

In order to lay the groundwork for that argument, I have to discuss some other research on evidence to which it responds, or on which I plan to build it. First is a brief analysis of the utility of statistical evidence and models of the process of proof, with particular attention to some descriptive and prescriptive weaknesses. Mathematical models are, at best, normative at very high levels of generality. As a guide to practical decisions, they fail to recognize several realities of practical decision making, creating an irrational model of rationality. The paper then summarizes the Pardo/Allen theory, which is in fair part an attempt to formulate a theory of the proof process that does not have the defects of mathematical models. Finally, the paper briefly sketches a theory of the preponderance of the evidence that is more consistent with research on cognition and with the legal significance of the burden of persuasion than either mathematical models or Pardo and Allen's theory could be.

²⁶ Again, I use the word "plaintiff" as short hand for the party with the burden of persuasion, unless the context clearly indicates otherwise.

I. ON QUANTIFICATION

The status of mathematical models of inference from evidence has been a persistent issue in evidence scholarship over the last forty years or so. Advocates of the use of mathematical models ranged from those who argued that jurors should be taught to use statistics in the process of deliberation about evidence that witnesses did not quantify,²⁷ to what is probably the predominant current view: that the models can be helpful for understanding the inferential process, even though judges or jurors do not, and probably cannot, use the models to evaluate specific pieces of evidence and their effect on the verdict in the time available to them at trial.²⁸ I will summarize the reasons for skepticism about mathematical models here, only insofar as Pardo and Allen's work is a response to them—I have addressed probabilistic modeling in its own right elsewhere.²⁹

A surprising amount of the debate crystallized around a set of hypotheticals. The best known is the Blue Bus case, one version of which Professor Shaviro offered in the Harvard Law Review:³⁰ plaintiff is hit by a bus that she knows is blue but cannot otherwise describe. She establishes that the defendant bus company operates eighty percent of all the blue buses in the area. In the absence of other relevant evidence, this evidence establishes an eighty percent chance that the bus company is liable. No one thinks that a similar situation is likely to occur in real life. For example, it seems unlikely that a plaintiff would ever testify that a bus hit her without providing details such as the speed of the bus, her own conduct, and so forth. The stakes for the academics involved in the debates over the blue bus hypothetical were not limited to resolving such cases. The debates over the blue bus example, and related exchanges over models based on Bayes' Theorem, were really about the nature of inference under conditions of uncertainty—the context in which almost all trials take place. Essentially, the Bayesian idea was that our inferences from evidence can, and should, conform to the principles of statistics, that we can use Bayes' Theorem to

²⁷ See Stephen E. Fienberg, & Mark J. Schervish, *The Relevance of Bayesian Inference for the Presentation of Statistical Evidence and for Legal Decisionmaking*, 66 B.U. L. REV. 771, 792 (1986) (advocates of use of Bayesian model to train jurors).

²⁸ E.g., Richard D. Friedman, *Answering the Bayesioskeptical Challenge*, 1 INT'L J. EVIDENCE & PROOF 276, 289–91 (1997).

²⁹ E.g., Craig R. Callen, *Notes on a Grand Illusion: Some Limits on the Use of Bayesian Theory in Evidence Law*, 57 IND. L. J. 1 (1982); Craig R. Callen, *Adjudication and the Appearance of Statistical Evidence*, 65 TUL. L. REV. 457 (1991); Craig R. Callen, *Adams and the Person in the Locked Room*, 1 INTERNATIONAL COMMENTARY ON EVIDENCE (1998).

³⁰ Daniel Shaviro, *Statistical-Probability Evidence and the Appearance of Justice*, 102 HARV. L. REV. 530, 530–31 (1989). He drew his hypothetical from Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1340–41 (1971). Tribe adapted it from *Smith v. Rapid Transit, Inc.*, 317 Mass. 469, 58 N.E.2d 754 (1945).

test to guide, assess or represent inferential processes.³¹ In other words, their argument was that the use of models demonstrated rational reasoning under conditions of uncertainty (“rational” being a tricky word),³² or, at least, a rational method for dealing with uncertainty.³³

Now it seems only human to have a desire to try to calculate the odds of guilt after receiving impressive statistical evidence (such as a positive DNA test with a very low random match frequency).³⁴ Even so, the premise that we can, or should, conform to mathematical probability theory is subject to some question. In that sense, the Blue Bus debates (or, more prosaically, questions about the status of mathematical models) raise questions about how we should resolve disputes of fact. And, given that the most common expression of the preponderance of evidence standard is “more likely than not,” or “more likely so than not so,” the preponderance standard is among the likelier targets of mathematical modelers.

To illustrate the application of Bayes’ Theorem without a lot of mathematics, assume the question is whether the defendant was the source of blood that bears on the identity of the person who committed a crime. Assume that a DNA test indicates that the sample and the defendant’s blood yield the same results when subjected to certain probes. Only 1 in 100 randomly selected persons’ blood would have those same traits. The frequency of lab error is 5%. Now, if the juror expressed her pre-DNA assessment of the likelihood that defendant committed the crime in numerical form, the following chart shows the results of applying Bayes’ Theorem.

³¹ One expression of Bayes’ Theorem is

$$P(B|A) = \frac{P(B) * P(A|B)}{[P(B) * P(A|B)] + [P(not-B) * P(A|not-B)]}$$

where the probability that A is true is P(A), that A is false is P(not-A), and the probability that A is true given that B is true is P(A/B). 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, 12–14 n. 46 (1982). Further, (i) P(B) and P(not-B) must equal 1, (ii) P(A) must equal P(A&B) + P(A & not-B), and (iii) P(A&B) must equal P(B) x P(A/B). Craig R. Callen, *Statistical Methods and a Paradox of Forensic Expertise*, in FORENSIC EXPERTISE AND THE LAW OF EVIDENCE 129, 130 n. 2 (1993).

³² See Richard Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof*, 66 B.U. L. REV. 439, 452–53 (1986) (“[T]he law may be Bayesian rational without mandating a fact-finding process that combines evidence with conformity with Bayes’s Theorem or its underlying axioms.”).

³³ See Richard D. Friedman, *Answering the Bayesioskeptical Challenge*, 1 INT’L J. EVIDENCE & PROOF 276, 278 (1997) (“The conventional [Bayesian] theory is not contrived; it reflects rational treatment of uncertainty based on intuitively appealing and experientially successful premises.”).

³⁴ E.g., 1 MUELLER & KIRKPATRICK, *supra* note 7, §3:5, at 437 n.4 (“It is irresistible [to] draw an analogy to likelihood or comparative probability: 51–49 (or for that matter 50.001–49.999) is enough.”).

Probability of Guilt before Learning of DNA Test Result	Probability of Guilt after Learning of Test (Rounded)
0%	64.1%
25%	84.3 %
50%	94,2%
75%	98%

In order to arrive at the probability of guilt without knowing about the DNA evidence, one could ask what the likelihood of receiving the other evidence would be if the defendant were guilty as compared to the likelihood of innocence and set the initial probability of guilt accordingly.³⁵ That is not to say that there would be hard data on the probability of guilt based on non-quantified evidence—it is subjective. “Application of the conventional probability theory to subjective probability assignments is the essence of Bayesianism.”³⁶ Professor Shaviro’s version of the hypothetical highlights several of the problems with Bayesian models because it is such a stark mathematical argument. First, Professor Shaviro’s view exemplified a problem that Professor Tribe earlier called the “dwarfing of soft variables”³⁷ when he argued that, absent other evidence, the court should direct a verdict in favor of the plaintiff.³⁸ Even assuming that the Blue Bus Co. operated 80% of blue buses in the area, there would be reasons to suppose that the likelihood of *liability* would be less than market share—absent further information. Thorough application of mathematical models would require discounting for several reasons. For example, one must discount (i) for possible errors in eyewitness identification,³⁹ and (ii) to reflect that the hypothetical presents few, if any, facts to show that the bus company breached a duty to the plaintiff, or that, if it did so, its breach was the proximate cause of the injury—a worrisome failure.⁴⁰ In short, as Tribe suggested might happen, the seeming ease of inference from the market share

³⁵ See Pardo & Allen, *supra* note 3, at 247 n. 70; Callen, *Notes on a Grand Illusion*, *supra* note 29, at 4–5. If the DNA evidence were the first evidence the jury received, one might set the prior probability at .5, on the theory that the jurors would not know who had the better case. See RICHARD A. POSNER, *FRONTIERS OF LEGAL THEORY* 372 (2001). There is a great deal more discussion about prior probabilities in the literature, but it is largely beyond the scope of this piece.

³⁶ Richard D. Friedman, *Answering the Bayesioskeptical Challenge*, 1 INT’L J. EVIDENCE & PROOF 276, 277 (1997).

³⁷ Tribe, *supra* note 30 at 1361–62.

³⁸ Shaviro, *supra* note 30, at 533.

³⁹ See, e.g., Craig R. Callen, *Cognitive Science and the Sufficiency of “Sufficiency of the Evidence” Tests*, 65 TUL. L. REV. 1113, 1127 (1991).

⁴⁰ POSNER, *supra* note 35, at 371 n.66; Callen, *Adjudication and the Appearance of Statistical Evidence*, *supra* note 29, at 473.

led Shaviro to ignore other portions of the fact finding task that would not be so easy to quantify.

Second, even if Professor Shaviro had developed an equation that incorporated the soft variables, we simply do not (and probably cannot) think about questions of fact as zealous Bayesians might wish.⁴¹ For example, in order for human inferences to correspond to subjective probabilities, all of our beliefs would have to be consistent, i.e., no contradictions or conflicts would be permissible. All signs are that we do not have the cognitive capacity to check all of our beliefs for latent conflicts. That is not to say that we consciously entertain conflicts, such as believing at the same time that Bush is president and that Gore is. But it is to say that, given the vast range of experiences we have, the constraints on our decision-making resources,⁴² and the unlikelihood that avoiding latent inconsistencies would have a significant payoff, there is no reason to suppose that we do, can, or should keep our beliefs consistent.⁴³ So there is substantial room for question about the usefulness of subjective probability statements, particularly for decision making in realistic conditions, i.e., real time.

Third, mathematical logic of the sort Bayesian models employ is reversible. Suppose one calculates the probability of a conclusion as .67. There is nothing in mathematical logic to prevent the calculator from concluding that he or she simply applied the mathematical formula incorrectly, developing a new equation (or reassigning probabilities) and recalculating, until satisfied that the result seems reasonable, and that the application of the theorem incorporates the relevant criteria.⁴⁴

That is related to the fourth point. Advocates of the use of subjective or personalist probability in conjunction with Bayes' Theorem might criticize the reversibility argument on the ground that it smacks too much of manipulating the formula to procure a particular result. Probabilistic models can only guide decisions well to the extent they reflect useful information and the appropriate inferential relationships among items of data and the

⁴¹ E.g., Richard O. Lempert, *Of Flutes, Oboes and the As If World of Evidence Law*, 1 INT'L J. EVIDENCE & PROOF 316, 317 (1997).

⁴² Including time and cognitive capacity.

⁴³ Professor Redmayne suggests that, given that we only have a limited number of beliefs in play at any time, we may be able to limit checks for consistency to those beliefs. See Mike Redmayne, *Rationality, Naturalism and Evidence Law*, 2003 MICH. ST. L. REV. 849, 864–65 (2003). Unless our beliefs are inherently probabilistic, however, in order for assignments of probability to be reliable, they must be consistent with some well-founded beliefs about odds formation, which we have tested in a number of other contexts. Since our everyday cognition does not seem to be Bayesian in any comprehensive sense, Redmayne's suggestion would still be intractable.

⁴⁴ Ronald J. Allen, *On the Significance of Batting Averages and Strikeout Totals: A Clarification of the "Naked Statistical Evidence" Debate, the Meaning of "Evidence," and the Requirement of Proof beyond a Reasonable Doubt*, 65 TUL L. REV. 1093, 1110 (1991).

conclusions to which they may point.

Nothing in an equation picks out one piece of data rather than another as important, or even worth including. Instead, probabilistic formulae rest on our experience, and we may have to adjust the fit of the formulae and the experience so that they can work appropriately. Nor does anything in an equation measure whether (i) the variables in the equation exhaust the types of information that bear (or reasonably could bear) on the ultimate questions or (ii) whether the evidence supporting the result of a probabilistic model is as complete as one would expect. Research suggests that jurors do consider completeness of the evidence in resolving questions of fact.⁴⁵

Jurors would, indeed, be somewhat irrational if they did not ask themselves not only whether the evidence was consistent with the parties' contentions, but also whether the parties' offers of evidence and hypotheses to explain it were consistent with the social import of the issue confronting them. To illustrate, consider a standard explanation for the intuition that the facts in the blue bus hypothetical are insufficient to support a verdict: that the absence of other evidence suggests that plaintiff did not bother making such a search for other evidence. In such a situation, the statistical information should be insufficient.⁴⁶ If a Bayesian model yielded what seemed to be an anomalous result and the decision maker noticed the absence of such a consideration from the model, the decision maker could rationally choose to reverse the process and reformulate the model. The key consideration would be whether the model and the result seemed to fit the decision-making process, not whether the decision maker adhered to the model.

In some ways, debates over the utility of probabilistic models for fact finding came to a critical point in an English case, *Denis Adams (No.2)*.⁴⁷ There, the Court of Appeal relied on its intuitive view of human decision making (which empirical studies largely corroborate) to reject the use of comprehensive Bayesian models in jury trials. An eminent statistician testified in a rape trial and, in the process, provided the jury with a questionnaire based on one Bayesian model⁴⁸ of the factual issues posed by

⁴⁵ See, e.g., Pennington & Hastie, *supra* note 23.

⁴⁶ See, e.g., David H. Kaye, *The Paradox of the Gatecrasher and Other Stories*, 1979 ARIZ. ST. L. J. 101, 107–09; Tribe, *supra* note 30, at 1349–50. Cf. POSNER, *supra* note 35, at 371–72 (making argument at lower ratios).

⁴⁷ R. v Adams (Denis John) (No. 2) [1998] 1 Crim. App. 377.

⁴⁸ The model that Professor Donnelly employed was only one of any number of possible models of the fact-finding task. To illustrate: the transcript of the judgment reproduces a portion of the trial judge's summing up. According to the judge, one of the prosecution experts, identified only as Mr. Lambert, raised doubts about the questionnaire, on the grounds that it did not cover "all the relevant evidence or all that [the jury] might think was relevant." *Id.* at 381. The trial judge himself pointed out that the questionnaire did not "include a box" in which the jury might take account of "Adams' own evidence and how he gave it." *Id.* at

both quantified and unquantified evidence in Adams' case,⁴⁹ and explained to them how they might use it to calculate the likelihood of Adams' guilt.

Adams was convicted of rape. The trial judge, the Court of Appeal said, had "made it abundantly plain to the jury that if they found the Bayesian approach helpful then they were at complete liberty, having had it explained to them, to use it in their deliberations."⁵⁰ The judgment went on, however, to hold that the trial judge should not have admitted the evidence at all. The *Adams* judgment is emphatic that a trial court should exclude such evidence in a case "lacking special features" absent there (or here).⁵¹ "[E]xpert evidence should not be admitted to induce juries to attach mathematical values to probabilities arising from nonscientific evidence adduced at trial."⁵²

As the Court of Appeal read it, the statistician's questionnaire omitted some evidence and arguments that the jury might have wished to credit, and might have included other lines of thought that they would have disregarded.⁵³ The victim in *Adams* had failed to identify the defendant as her attacker in an identity parade. The trial court mentioned that failure in summing up the evidence.⁵⁴ The Court of Appeals noted that jurors might well ask themselves about the identification and about the significance of Adams' own evidence,⁵⁵ another apparent omission from the questionnaire.⁵⁶

Fifth, Bayesian models of fact finding are difficult to square the notion of proof of elements of a claim by a preponderance of evidence with mathematical models. Suppose a claim has four elements, and the truth of each is independent of the others. Further, suppose that, according to an interpretation that is standard for Bayesians⁵⁷ and many others, a

381–82.

⁴⁹ *Id.* at 380.

⁵⁰ *Id.*

⁵¹ *Id.* at 385.

⁵² *Adams* (Denis John) (No. 2) [1998] 1 Crim. App. at 385.

⁵³ It might seem that jurors could ignore the boxes for arguments that they did not consider worth their attention. That argument would conflict with a number of assumptions that seem to underlie the defendant's use of Professor Donnelly. Arguing that jurors might disregard arguments to which Professor Donnelly's questionnaire referred, and attempt to use the questionnaire without relying on them, assumes that the jury has considerable sophistication in statistics. Yet, if the jury has such knowledge, it is not clear why Professor Donnelly's testimony should be admissible. In addition, jurors who were inclined to defer to the expert witness' knowledge of logic might be very reluctant to set aside lines of thought that he suggested. In any event, jurors who lacked statistical sophistication would have difficulty deciding whether they could ignore an item in the questionnaire without conflict with the underlying mathematical model.

⁵⁴ *Adams* (Denis John) (No. 2) [1998] 1 Crim. App. at 382.

⁵⁵ *Id.* at 384–85.

⁵⁶ *Id.* at 382.

⁵⁷ Friedman, *supra* note 28, at 278 ("This, of course, is the familiar 'more likely than

preponderance is any probability greater than .5. So, according to mathematical theory, plaintiff could show four elements by just more than .5, and the probability of their conjunction would be roughly one sixth, or .514.

There are two ways to reconcile that requirement with the preponderance standard, each of which is problematic. One can read the preponderance standard to apply only to the conjunction, which means that the probability of each element of a two element claim should be more than .7 (on average) and of a three element claim, more than .8. It is difficult to see why the number of elements should affect the burden of persuasion for each or why the burden should be as high as .7 given the typical instruction. Alternatively, one might say that the jury need only believe that a set of evidence that shows all of the elements of the claim is more likely than .5 to be accurate. That argument saves, in a sense, mathematical models, but only by stipulating that they cannot be used for anything more detailed than a global assessment of the evidence.⁵⁸

One final argument is important for my discussion of the Pardo/Allen theory. What is called the reference class problem in probabilistic models reinforces the importance of experience as a foundation for the application of logic. In the blue bus case above, the class of buses in town may not be the most appropriate one for assessing the likelihood of liability. Instead, it may be preferable to ask what percentage of the vehicles that the plaintiff would identify as blue buses are owned and operated by the defendant, what percentage of vehicles that hit people are owned by the defendant, what percentage of the buses on the relevant street are blue, or what percentage of vehicles that the plaintiff would identify as blue buses that drive on the relevant street are defendant's.⁵⁹ Probabilistic logic does not tell us which reference class is optimal, or how picky we should be about reference classes. Certainly, as Professors Pardo and Allen have pointed out, if we built a probabilistic model that took account of all possibly relevant factors, gathered all of the information relating to those factors and sought to calculate the probability of liability, and applied that model correctly, the result would be a probability of liability of either 1.0 (certainly liable) or 0.0

not,' or 'balance of probabilities,' standard.”).

⁵⁸ One Bayesian has argued that jury instructions tend to be ambiguous on the question of whether the plaintiff must prove each element of the claim by a preponderance, or just the conjunction of them. Dale A. Nance, *A Comment on the Supposed Paradoxes of a Mathematical Interpretation of the Logic of Trials*, 66 B.U. L. REV. 947, 949 (1986). That seems overstated. He cites as ambiguous an instruction that says, in pertinent part, “If the proof should fail to establish any essential element of plaintiff’s claim by a preponderance of the evidence in the case, the jury should find for the defendant.” *Id.* (quoting E. DEVITT & C. BLACKMAR, FEDERAL JURY PRACTICE AND INSTRUCTIONS § 71.14 (3d ed. 1977)).

⁵⁹ Ronald J. Allen & Michael S. Pardo, *The Problematic Value of Mathematical Models of Evidence*, 36 J. LEG. STUD. 107, 109 (2007).

(certainly not liable).⁶⁰ Absent that information, Bayesians must confront the question of the appropriateness of the reference classes to which their models refer. There is no mathematical way to answer that question.

That concern is related to a phrase that keeps popping up in cases and commentary. Debates over blue bus hypotheticals often touched on the status of “naked”⁶¹ or “solely”⁶² statistical evidence, and whether it is sufficient to support a verdict. Discussions of the question are confusing, because the distinction between naked statistics and other sorts of evidence is ill-founded. A “naked” or “sole” statistic would simply be a number, with no information about what it quantified, or the context in which one might make inferences based on it. There would be no reason to regard it as probative of anything.⁶³ Reluctance to premise a verdict on a statistic such

⁶⁰ Pardo & Allen, *supra* note 3, at 260. [

⁶¹ See Kaye, *supra* note 21.

⁶² See Shaviro, *supra* note 30, at 545.

⁶³ See, e.g., Callen, *Adjudication*, *supra* note 29, at 466–69. To adapt a famous thought experiment by Professor John Searle,

Suppose we were to conduct an experiment in which I sat in a locked room with a personal computer and an instruction manual. At some point, I received an e-mail message listing “G” as a conclusion, and designating four or five other variables as evidence. I then consulted the manual and sent a message with a series of questions derived from the instructions therein, such as “What is P(A|G)?”. On receiving a response stipulating mathematical values in response to each question, I inserted them into a spreadsheet according to the directions in the manual, and depressed the enter key. On reading the result, 0.563, I then transmitted it as an e-mail message to my unknown correspondent. Would my role (as compared to that of, e.g. a person sending me e-mail messages) be fairly described as fact-finding or problem-solving?

Craig R. Callen, *Adams and the Person in the Locked Room*, 1 INT’L COMMENTARY ON EVIDENCE: iss. 1, art. 3., at 3–4 (1998). There seems to be no reasonable argument that my response to the formal symbols would constitute fact-finding. Hence, there is a good deal more to fact-finding than formal logic. Searle originally sought to show that a formal logical system cannot accurately duplicate human thought. Searle took the position that digital computers cannot understand what they are doing solely on the basis of computer programs, because complying with formal procedures does not constitute thought. Instead it is merely a matter of manipulating meaningless symbols without semantic content. JOHN SEARLE, MINDS, BRAINS AND SCIENCE 36 (1984). He later refined that point, arguing that a computer running a program would not thereby be intelligent because the computer would be incapable of interpreting its own operations to refer to the world. JOHN R. SEARLE, THE MYSTERY OF CONSCIOUSNESS 14–15 (1997). For a similar argument, see TERRENCE W. DEACON, THE SYMBOLIC SPECIES: THE CO-EVOLUTION OF LANGUAGE AND THE BRAIN 444–45 (1997). If the person sending me email messages were instead to refer, for example, to the likelihood that the defendant’s bus negligently struck the plaintiff, instead of P(N), then I would have some degree of awareness of the possible application of my work. Even then I could hardly be said to have done the fact-finding, since I could not seek or exclude evidence to which the

as the percentage of blue buses in town has less to do with whether the evidence is naked than it has to do with questions such as (i) whether the statistic refers to the correct reference class, since blue busses might not be uniformly distributed through the area,⁶⁴ (ii) whether reliance on statistical evidence would promote or hinder the purposes of the substantive law,⁶⁵ and (iii) whether such reliance would give other plaintiffs incentives to refrain from further investigation or submission of evidence.⁶⁶ Some courts have thought that they confronted the issue of whether they should permit a verdict based on “naked” statistical evidence, but they had a great deal more information than a single, solitary number. Generally the question in such cases was whether, when the evidence would not be sufficient to support a verdict without the statistical probability evidence, the evidence would be sufficient if the statistical information were added.⁶⁷ One of the issues likely to arise in any such situation is whether the proponent of the evidence established an adequate factual foundation for it, which is closely related to, if not the same as, the reference class issue.

II. INFERENCE TO THE BEST EXPLANATION

A. *Everyday Life and Sherlock Holmes*

In response to the flaws in mathematical models, particularly as they relate to jury decision making in practice, Professors Pardo and Allen sought to offer an alternative grounded in a better description of fact finders’ decision making processes at trial. The Pardo/Allen theory focuses on inference to the best explanation. Inference to the best explanation is not an uncommon process, although few of us who use it think of it in those terms.⁶⁸ So I will digress slightly to offer a few illustrations of it.

At my gym one night, I was walking past a man who was talking loudly into a cell phone and said “Four foot long turkey and three foot long ham. I’m on my way.” The first idea (technically a hypothesis or explanation) that popped into my head was that the man was discussing a topic that related to some very large farm products. I quickly realized that there was another

instructions referred. In other words, at best, mathematical models of reasoning under uncertainty rely heavily on the decision maker’s use of her experience to identify, organize, and assess relevant evidence. That ability is not itself mathematical, and reliance on it may, in some senses, render the statistical results at best, a summary of underlying inferences rather than a guide for inference.

⁶⁴ Pardo & Allen, *supra* note 3, at 260 & n. 121.

⁶⁵ Pardo & Allen, *supra* note 3, at 260 & n. 121.

⁶⁶ *See, e.g.*, Hart v. Sec’y of Health and Human Services, 60 Fed. Cl. 598, 607 (2004). *See* Callen, *Adjudication*, *supra* note 33, at 497–98.

⁶⁷ *E.g.*, Baker v. Bridgestone/Firestone Co., 966 F. Supp. 874, 876 (W.D. Mo. 1996).

⁶⁸ *See generally* Harman, *supra* note 14.

possible explanation: that he was on his way to pick up four foot long turkey, and three foot-long ham, sandwiches. My recall of an additional piece of information, that Jimmy John's Sandwiches was around the corner, seemed to strengthen that hypothesis. Moreover, had he been on his way to pick up a four foot long turkey, he would probably have said "a four foot long turkey," or in the case of the ham "a three foot long ham." Taking a more whimsical view, given that everyone was getting pretty tired of the Super Tuesday stories in the media, a four foot long turkey or a three foot long ham would probably have been the lead story on the local television news, and I had seen no such story.⁶⁹

In effect, I used inference to the best explanation, in that I started with some pieces of information, generated some explanations from it, assessed the quality of those explanations, chose the best one, and accepted it for my purposes. In the process, I unconsciously searched my memory for other information relating to the hypothesis, in this case, information about the location of a sandwich shop, grammar, and the typical size of farm animals.⁷⁰

While Pardo and Allen do not mention it (and may not be too fond of Holmes),⁷¹ inference to the best explanation has a relatively unique distinction among theories of inference: a relatively standard fictional source of illustrations, Arthur Conan Doyle's Sherlock Holmes. Citations to Holmes seem to be fairly common in academic scholarship on the relation of inference to the best explanation to fact finding and investigation.⁷² Criminologists use the fictional Holmes' methods, and may acknowledge him as a source of ideas and useful techniques.⁷³ Charles Peirce originated the concept of abduction, which is closely related, if not identical, to inference to the best explanation.⁷⁴ Professor David Schum noted Peirce was

⁶⁹ There were other alternatives, including the possibility that he was seeking to obtain the services of bad child actors, selected by size.

⁷⁰ See Harman, *The Inference to the Best Explanation*, *supra* note 14, at 89.

⁷¹ For fairly strong criticism of the Holmes stories' depiction of circumstances in which he "probably should be wrong and he isn't," see, for example, D. Michael Risinger, *Boxes in Boxes: Julian Barnes, Conan Doyle, Sherlock Holmes and the Edalji Case*, 4 INT'L COMMENTARY ON EVIDENCE iss. 2, art. 3, at 13 (2006), <http://www.bepress.com/ice/vol4/iss2/art3/>; Richard Posner, *CSI: Baker Street*, THE NEW REPUBLIC (Oct. 11, 2004), <http://www.tnr.com/doc.mhtml?i=20041011&s=posner101104>.

⁷² See, e.g., David A. Schum, *Species of Abductive Reasoning in Fact Investigation in Law*, in THE DYNAMICS OF JUDICIAL PROOF: COMPUTATION, LOGIC AND COMMON SENSE 307, 314 (Marilyn MacCrimmon & Peter Tillers, eds., 2003); Thomas A. Sebeok & Jean Umiker-Sebeok, "You Know My Method": A Juxtaposition of Charles S. Peirce and Sherlock Holmes, in THE SIGN OF THREE: DUPIN, HOLMES, PEIRCE 11 (Umberto Eco & Thomas A. Sebeok, eds., 1983).

⁷³ Marcello Truzzi, *Sherlock Holmes: Applied Social Psychologist*, in THE SIGN OF THREE, *supra* note 72, at 55, 57–58.

⁷⁴ Peirce is often credited as the originator of abduction. "Abductive inference is widely taken to be the same as inference to the best explanation." DOUGLAS WALTON, ABDUCTIVE

writing at the same time that Conan Doyle wrote the Holmes stories. Schum observed that Holmes' "investigative strategies seem almost identical to the concept of abductive reasoning described by Peirce." Nevertheless, "there is no evidence that Peirce and Conan Doyle ever shared ideas on the subject."⁷⁵ Illustrations from Conan Doyle may be more dramatic than typical abduction is, and, indeed, Holmes was not too critical of his own hypotheses. Even so, as illustrations of the importance of explanations, they work pretty well.

One of Conan Doyle's shorter stories illustrates a typical characteristic of hypotheses that compete to be the best explanation—they are usually causal. Conan Doyle⁷⁶ wrote a story entitled *How Watson Learned the Trick*, which was prepared as a miniature book for the Queen's doll house's library.⁷⁷ In that story, Sherlock Holmes' companion, Dr. Watson, sought to demonstrate his mastery of Holmes' "trick," the ability to draw correct conclusions from seemingly inconsequential matters. He announced his conclusion that Holmes was preoccupied upon arising that morning, based on Watson's observation that Holmes was unshaven. Further, Watson inferred that Holmes had been unsuccessful in an investigation for a client named Barlow, based on the fact that Watson saw Holmes open an envelope with the name Barlow on it and groan. Watson also believed that Holmes had taken up financial speculation because he turned to the financial section and uttered an exclamation of interest. Finally, Watson concluded that Holmes expected an important visitor momentarily because Holmes was wearing his coat rather than a dressing gown.

Although Holmes applauded Watson's initiative, he told Watson that he did not yet have the trick. Holmes was wearing his coat in anticipation of his forthcoming visit to his dentist, Barlow, whose letter had just confirmed the time for Holmes' appointment, resulting in the groan. The financial page was next to the cricket scores, and Holmes was interested in Surrey's outcome against Kent. Finally, Holmes was unshaven because he sent his razor out to be sharpened.⁷⁸ Each part of each explanation rested on a causal link, in that it suggested a cause for the evidence.

Watson, of course, was the character Conan Doyle used as Holmes' foil, and it would not have done for Watson to master Holmes' method, even

REASONING 34 (2004).

⁷⁵ Schum, *supra* note 72, at 314.

⁷⁶ Umberto Eco has done a good deal of work on abduction (which many theorists regard as another name for inference to the best explanation, and vice versa). He cites Holmes as an exemplar of abduction. Umberto Eco, *Horns, Hooves, Insteps: Some Hypotheses on Three Types of Abduction*, in *THE SIGN OF THREE*, *supra* note 72, at 198, 215–20.

⁷⁷ ARTHUR CONAN DOYLE, *How Watson Learned the Trick*, in *THE RETURN OF SHERLOCK HOLMES* 324 (Richard L. Green ed., 1993).

⁷⁸ *Id.* at 325–26.

though Doyle referred to that method as nothing but “systematized common sense.”⁷⁹ The story does, however, depict the use of explanations to assess evidence and draw conclusions from it, along with an implicit suggestion that the method is not infallible.

Another Holmes story suggests that an explanation may, in fact, show that some characteristics of a situation may be probative when we might, if we failed to formulate the explanation, consider those aspects of no inferential value. In Doyle’s story, *The Silver Blaze*, Holmes solves a mystery, in large part, because a possible explanation suggested that evidence might be relevant. In that story, Holmes was trying to determine how a racehorse had disappeared. He surmised that the disappearance might have been part of a plot to prevent the horse from winning a race, by making a slight cut in the horse’s tendon. Assuming that someone as careful as the perpetrator would not undertake the operation without practice, he asked whether some nearby sheep were in good condition, found out that they had become lame without any particular reason, and regarded that as confirmation of his hypothesis.⁸⁰

On its face, the condition of the sheep would have little bearing on the horse snatching. If, however, Holmes had in mind a possible theory about the perpetrator’s motive, then the sheep’s condition would become relevant. It is important to note here, by the way, that the point of these illustrations is not to argue or even suggest that jurors always, or often, engage in processes as dramatic or as risky as Holmes’. Nor, when jurors are considering explanations as surprising as Holmes’ often are, should they probe them as little as Holmes seems to do. As Umberto Eco pointed out, detectives can afford to take more chances in abduction than scientists can,⁸¹ and given their relative roles, jurors as the final authority must exercise more caution than detectives. “The greatest weakness in Holmes’ applications of inference—at least as Watson related them to us—was Holmes’ failure to test the hypotheses which he obtained through abduction.”⁸² On the other hand, however, jurors do routinely evaluate the quality of the explanations or stories⁸³ that the parties advance for the evidence, and it is their job (as it was not Watson’s) to analyze the stories critically.

⁷⁹ ARTHUR CONAN DOYLE, *The Blanched Soldier*, in *THE CASE-BOOK OF SHERLOCK HOLMES* 151, 169 (W.W. Robson, ed. 1993).

⁸⁰ ARTHUR CONAN DOYLE, *Silver Blaze*, in *THE MEMOIRS OF SHERLOCK HOLMES*, 3, 23, 28 (Christopher Roden ed. 1993).

⁸¹ Umberto Eco, *Horns, Hooves, Insteps: Some Hypotheses on Three Types of Abduction*, in *THE SIGN OF THREE: DUPIN, HOLMES, PEIRCE* 198, 220 (Umberto Eco & Thomas A. Sebeok eds., 1983).

⁸² Marcello Truzzi, *Sherlock Holmes: Applied Social Psychologist* in *THE SIGN OF THREE*, *supra* note 72, at 55, 70.

⁸³ Or theories, or hypotheses.

B. *Inference to the Best Explanation in the Pardo/Allen Theory*

1. Professors Pardo and Allen on the Basics of Inference to the Best Explanation

Pardo and Allen contended that attention to what they refer to as “explanatory considerations” would “provide a better account of micro-level proof issues regarding the relevance and probative value of evidence . . . which justifies the open ended nature” of rules such as Federal Rules 401 and 403.⁸⁴ The Pardo/Allen theory, they say, offers an alternative protocol for inferences from the evidence that may, in turn, produce conclusions that are better justified than those that mathematical models may yield.⁸⁵ Further, they argued that explanatory considerations afford a better view of standards of proof such as the preponderance test than conventional theory does, and concomitantly, a better justification for various sorts of motions that allow the court to rule against the party with the burden of proof when evidence is insufficient or fails to support a verdict adequately.⁸⁶

Explanations, they say, are prior to inference from evidence, and also help to direct it, in the sense that explanations help to assess the likelihood of particular hypotheses or conclusions. Pardo and Allen depict the process of inference to the best explanation as involving two steps: first, generation of hypotheses or explanations and, second, selection of the best explanation from among the potential hypotheses. When it comes to formulation the “domain of the inferential task,” which probably means our knowledge and experience with the subject matter of the questions at hand, will help us formulate explanations of data and, at the same time, constrain us from relying too much on sheer imagination.⁸⁷ In legal disputes, the substantive law requires that the parties address certain facts, and successful explanations must pertain to one or more of those facts. As Pardo and Allen see it, the only limits on the generation of hypotheses that address those critical facts are the limits of the creativity of the parties, the court, and the jurors.⁸⁸

The second phase of inference to the best explanation, in the authors’ eyes, involves a comparison of explanations and choice of the best one. They cite criteria for choice such as the extent to which a hypothesis (i) is simple;

⁸⁴ Pardo & Allen, *supra* note 3, at 225–26.

⁸⁵ Pardo & Allen, *supra* note 3, at 225.

⁸⁶ Pardo & Allen, *supra* note 3, at 226.

⁸⁷ Pardo & Allen, *supra* note 3, at 229. As they point out, explanations that serve for a chemist will not serve for a musical composer, so that acceptability of hypotheses may be a function of the decision maker’s cognitive interest. Pardo & Allen, *supra* note 3, at 229–30. Non-law disciplines are not of much interest here, although the jurors’ cognitive interest does have some importance.

⁸⁸ Pardo & Allen, *supra* note 3, at 230.

(ii) accounts for numerous and diverse pieces of evidence; (iii) fits with our general knowledge, or “background beliefs;” and (iv) is systematic rather than makeshift.⁸⁹ That measure of the quality of an explanation is contested. For example, one philosopher sets a higher bar: as a threshold for acceptance as the best account of the facts, a hypothesis must (i) “explain . . . different kinds of facts; (ii) possess “simplicity” (a quality inversely related to “the number and nature” of the assumptions on which the hypothesis’s explanatory power depends); (iii) exhibit analogies with other explanations known to be successful and (iv) “strongly cohere” with more basic or “background” beliefs.⁹⁰

Pardo and Allen argue that decision-makers who are choosing among explanations will choose the hypothesis that best explains the aspects of evidence that are most of interest to them. So, for example, if we are considering the explanation “the butler did it,” that is helpful if we wonder who committed a murder, but much less useful as an explanation if the question is whether the dining room silver was laid out in a precise pattern.⁹¹ Finally, in explaining the force of explanations, they note that there is something of a circularity. As they conceive of explanations, an explanation can help us to understand facts, and yet the facts also tend to make the explanation seem more accurate. They argue that the circularity “is not vicious or problematic” but merely an accurate description of human inferential practice.⁹²

2. Inference to the Best Explanation and Sufficiency of Evidence

Turning to trials, they argue that the parties are primarily responsible for generating theories or hypotheses to explain the evidence. Those with the burden of persuasion must address all of the elements of the claim or defense; their opponents may address fewer since a verdict in their favor on one will defeat their opponent. Parties may offer alternative explanations, and triers of fact may develop explanations of their own.⁹³ Once the hypotheses are formulated, Pardo and Allen believe that the triers of fact

⁸⁹ Pardo & Allen, *supra* note 3, at 230.

⁹⁰ Laudan, *supra* note 15, at 294 (2007) (discussing the views of Paul Thagard). In his seminal article on inference to the best explanation, after mentioning most of the considerations that the authors list, Professor Harman wrote, “I do not wish to deny that there is a problem about explaining the exact nature of these considerations; I will not, however, say anything more about this problem.”

⁹¹ Pardo & Allen, *supra* note 3, at 232–33.

⁹² Pardo & Allen, *supra* note 3, at 233. It is not clear how this circularity is consistent with the ability of the fact finder to evolve theories on his or her own.

⁹³ Pardo & Allen, *supra* note 3, at 234. In addition, see the discussion of *St. Mary’s*, *infra* note 192 and accompanying text.

choose the most plausible,⁹⁴ accepting it as accurate and finding in favor of the party that the explanation supports. They rely on the work of Pennington and Hastie, which argues that jurors reach their conclusions through the construction and assessment of stories in terms of criteria such as internal consistency, completeness, and uniqueness.⁹⁵ Once again, in determining the extent of detail necessary for hypotheses and the facts that they should explain, the substantive law and the contentions of the parties will determine the contours that a hypothesis must map in order to be helpful. Where stories differ in terms of details that do not matter or at least are not critical, Pardo and Allen argue that triers of fact may aggregate the stories in order to assess the likelihood of critical facts.⁹⁶

Sometimes, however, development of a good theory may be more difficult. Noting that cases may occur in which “neither party offers a particularly plausible explanation of the evidence, either because neither side can explain key pieces of evidence or because there is such a paucity of evidence that it can be explained in multifarious ways,” none of which is better, or more likely, than its rivals.⁹⁷ According to the authors, the comparative aspect of their theory requires a verdict in the first case (where no explanation offered or constructed is particularly plausible) unless no explanation is of better quality than the others.⁹⁸ If the explanations are all of the equivalent quality, then the party with the burden of persuasion would fail to meet it. If the evidence was so thin that it did not permit the trier to distinguish among explanations, the party with the burden of persuasion would, once again, have failed.⁹⁹ As the authors note, summary judgments and judgments as a matter of law are among the means of enforcing the

⁹⁴ The authors probably mean to use “most plausible” as the equivalent of “best,” although they do not specify the significance of plausibility. Some writers on abduction consider plausibility a tricky concept. See *Appendix B: Plausibility*, in *ABDUCTIVE INFERENCE: COMPUTATION, PHILOSOPHY, TECHNOLOGY* 266–67 (John R. Josephson & Susan G. Josephson eds., 1996). I do not understand Pardo and Allen to be staking out a position in that debate.

⁹⁵ Pardo & Allen, *supra* note 3, at 234 & n. 35. In discussing uniqueness, Pennington and Hastie say that the availability of multiple explanations for the evidence will weaken belief in one over the others. On the other hand, if there is only one story that satisfies the criteria of consistency, completeness, and uniqueness, jurors will tend to use it in reaching their decision. See Nancy Pennington & Reid Hastie, *A Cognitive Model of Juror Decision Making: The Story Model*, 13 *CARDOZO L. REV.* 519, 528–29 (1991).

⁹⁶ Pardo & Allen, *supra* note 3, at 235–37. The Pardo/Allen theory would also accommodate aggregation of alternative defenses, such as “I didn’t do it. But if I did, it was in self-defense. And if it wasn’t in self-defense, I was coerced to do it.” Pardo & Allen, *supra* note 3, at 256. Authors devote little attention to that possibility since they doubt that defendants are likely to use such an approach.

⁹⁷ Pardo & Allen, *supra* note 3, at 237–38.

⁹⁸ That assumes, of course, that at least one of the explanations of equivalent quality favors each party.

⁹⁹ Pardo & Allen, *supra* note 3, at 238.

burdens of proof—in particular the burden of production.¹⁰⁰

Having looked at the process of proof in civil cases (and in other situations not relevant here) Pardo and Allen turn briefly to the relationship of the Pardo/Allen theory to rules of evidence, revisiting their somewhat circular theory of the relationship between evidence and explanation (that explanations explain evidence, which, in turn, makes the explanations seem more accurate). Relevant evidence, they say, is evidence that the offering party's explanation explains (assuming that the explanation bears on a fact that the substantive law considers important). Probative value, they argue, is the extent to which an explanation explains the available evidence.¹⁰¹

3. Objections to the Pardo/Allen Theory

After they set out the degree to which they believe their conception of inference to the best explanation is reflected in fact-finding, Pardo and Allen turn to possible objections to it. One such objection¹⁰² is that the choice of an explanation as “best” is too subjective. They argue that criteria such as the degree to which the hypothesis explains the evidence are objective and that the prevalence of inference to the best explanation in problem-solving suggests that such practices have been successful over time and in a number of contexts.¹⁰³

Pardo and Allen freely concede that there is no guarantee that the explanation that the Pardo/Allen theory requires the decision maker to select as the best is, in fact, true. If an argument in deductive logic is valid, and the premises of that argument are true, the conclusion must be.¹⁰⁴ But fact-finding at trial seldom turns on deductive arguments. Instead, reasoning at trial almost always involves premises that, even if true, do not guarantee the truth of the conclusion to which they lead. So, for example, a juror might believe that it is difficult to see a traffic light suspended over an intersection if the sun is setting behind it, and a witness might testify that she saw a red light even though she was facing into the setting sun, but that would not

¹⁰⁰ Pardo & Allen, *supra* note 3, at 238 n.42.

¹⁰¹ Pardo & Allen, *supra* note 3, at 241–42.

¹⁰² I will not recapitulate objections from the philosophy of science, which do not seem to have much bearing on the preponderance standard.

¹⁰³ Pardo & Allen, *supra* note 3, at 23.

¹⁰⁴ Of course, if the premises of a deductive argument are wrong, then the conclusion may be as well. For example, the following deductive argument is perfectly valid, yet its conclusion is inaccurate:

The sky is green.

All things that are green are on Mars. Therefore, the sky is on Mars.

necessarily require the conclusion that she did not see the light. On the other hand, if the juror concluded that the witness was accurate, and if the applicable law held that the motorist was negligent for driving ahead despite the red light, the deduction would require the conclusion that the driver was negligent.¹⁰⁵ Their point is that, if we cannot rely exclusively on deductive logic to guide fact-finding, we must accept the possibility that our conclusions will be in error. The Pardo/Allen theory, they contend, is no greater potential source of error than any other logic that is inductive, that is, that seeks to guide inferences from the evidence that is inconclusive.¹⁰⁶

Readers familiar with the Supreme Court's jurisprudence on scientific evidence, among others, will recall the stress that the philosopher Karl Popper, on whom the Court relied in *Daubert v. Merrell Dow Pharmaceuticals*,¹⁰⁷ put on testing as a justification for inference. Pardo and Allen take the position that the parties test explanations of evidence through cross-examination and submission of rebuttal evidence.¹⁰⁸ Further, if the defendant¹⁰⁹ fails to adequately oppose a hypothesis in favor of the plaintiff, the authors point out that the jury may still find for defendant if the evidence is such that there is no reason to choose plaintiff's hypothesis over theories favoring the defendant. They view that result as implicit in the plaintiff's burden of proof.¹¹⁰

Given the controversy over the use of mathematical models in evidence, Pardo and Allen address some likely objections from probability theory. Fact finders rarely, if ever, apply comprehensive Bayesian models, so controversy between advocates of quantitative models and advocates of alternative approaches tend to focus on whether the quantitative models can demonstrate the behavior of ideal fact finders, or whether they have normative implications for fact-finding or the law of evidence.¹¹¹ Ideal fact finders, or whether they have normative implications for fact finding or the

¹⁰⁵ See Pardo & Allen, *supra* note 3, at 227, 246.

¹⁰⁶ See GILBERT HARMAN, *THOUGHT* 164–68 (1973).

¹⁰⁷ 509 U.S. 579, 593 (1993).

¹⁰⁸ It seems very likely that Pardo and Allen meant this term to include opposing evidence that is offered in a case-in-chief rather than confining it to rebuttal examination.

¹⁰⁹ Shorthand here for the party that bears neither burden of persuasion nor production.

¹¹⁰ Pardo & Allen, *supra* note 3, at 246–47. For a reason that is not clear, Pardo and Allen say that the jury should find for defendant when the plaintiff's hypothesis seems better than the one defendant offered, but no more persuasive than others that favor the defendant, *because the plaintiff has the obligation to produce "sufficient evidence from which a reasonable factfinder could differentiate among the potential contrasting explanations."* Pardo & Allen, *supra* note 3, at 246–47 (emphasis added). That is almost certainly a misstatement. Setting aside occasional baroque attempts to instruct on civil presumptions, jurors do not enforce the burden of production as such, and a failure to offer sufficient evidence would normally be addressed by a judgment as a matter of law or cognate motion.

¹¹¹ Pardo & Allen, *supra* note 3, at 248.

law of evidence.¹¹²

One set of objections concerns the question of aggregation since it is in theory possible to formulate an infinite set of stories favoring one party or the other. If one must select the best story, some Bayesians argue, any given story may be little if any more likely than its closely related colleagues, while a mathematical model can arithmetically combine the probabilities of all stories that favor a particular viewpoint, and then conduct a comparison.¹¹³ Pardo and Allen point out, however, that substantive law and the actual conflicts between the parties at trial help fact finders to identify differences among stories that they can ignore for purposes of aggregation. If, for example, one witness testifies that the plaintiff was wearing black on a dark night, and another midnight blue, the difference would seldom matter if the defendant could only be negligent if he could have seen the plaintiff. Accordingly, each of those accounts would weigh in favor of the defendant, since the black/midnight blue conflict would be both factually and legally unimportant.¹¹⁴ Moreover, they point out that the Bayesian approach is dependent on “explanatory considerations” because such considerations guide the identification of circumstances that may bear on fact-finding, and thus the choice of data that are to be included in a mathematical model.¹¹⁵

Professor Allen is one of the main champions of the conjunction paradox argument¹¹⁶ against the Bayesian theory, so it follows that they believe that the Pardo/Allen theory avoids the paradox. Under the Pardo/Allen theory, once the fact finders have selected the best explanation of the evidence, that explanation either includes and confirms the elements of a claim or defense, or it does not. If it does confirm (or “instantiate”) the elements, then the party bearing the burden of persuasion wins, otherwise, it loses.¹¹⁷

¹¹² Pardo & Allen, *supra* note 3, at 248.

¹¹³ Pardo & Allen, *supra* note 3, at 254–61 (discussing Richard D. Friedman, “*E* is for Eclectic:” *Multiple Perspectives on Evidence*, 87 VA. L. REV. 2029 (2001)); Richard D. Friedman, *Infinite Strands, Infinitesimally Thin: Storytelling, Bayesianism, Hearsay and Other Evidence*, 14 CARDOZO L. REV. 79 (1992); Dale A. Nance, *Naturalized Epistemology and the Critique of Evidence Theory*, 87 VA. L. REV. 1551 (2001).

¹¹⁴ Pardo & Allen, *supra* note 3, at 229–30.

¹¹⁵ Pardo & Allen, *supra* note 3, at 253.

¹¹⁶ *E.g.*, Ronald J. Allen, *A Reconceptualization of Civil Trials*, 66 B.U. L. REV. 401, 407 (1986).

¹¹⁷ Pardo & Allen, *supra* note 3, at 231–34.

Bayesian theory, they argue, requires that the court be aware of all possible explanations for the evidence so that the probability of the explanations will add up to 1.0.¹¹⁸ Pardo and Allen point out that parties at trial behave otherwise, often seizing on the best explanation in their favor, possibly due to the fear that proffering a large number of explanations would suggest to the fact finder that none of them had support. They argue that “the standard problem” for trials is not the difficulty of accumulating and aggregating all of the stories favoring the parties. Instead, they think that resolving cases in which one would quantify the probability of the plaintiff’s¹¹⁹ story as less than .5, but substantially greater than the probability of the defendant’s. To use their example, suppose that the probability of the plaintiff’s case is .4 and the defendant has two defenses, which each have a probability of .1. A Bayesian approach, they argue, would require a defense verdict, because the probability of plaintiff’s case would not exceed .5. As they conceive of it, the Pardo/Allen theory requires that the plaintiff win, because plaintiff’s explanation would be the best.¹²⁰

4. Pardo and Allen on the Implications of the Pardo/Allen Theory

Pardo and Allen contend that clarifies standards of persuasion because (i) it is a better description of trial processes than probabilistic models are, (ii) is consistent with the best current research on jury processes, (iii) does not produce the conjunction paradox “and [iv] allows the standards to fulfill their function of distributing errors among the parties.”¹²¹

In light of jurors’ use of inference to the best explanation, they believe that procedures for testing the sufficiency of evidence such as summary judgment or directed verdict are vital. Assuming that the moving party does not have the burden of persuasion, such motions pose the question of whether there is sufficient evidence to allow the jury reasonably to infer that the

¹¹⁸ I am not sure what Professors Pardo and Allen mean in this regard. In their favor, Bayesian models typically assume that $P(A)$ and $P(\text{not-}A)$ sum to 1.0. On the other hand, given constraints of time and cognitive resources, it seems doubtful that any advocate of Bayesian models is likely to argue that the probability of all of the explanations that jurors must consider must sum to 1.0.

¹¹⁹ Assuming that the burden of persuasion is on the plaintiff.

¹²⁰ Pardo & Allen, *supra* note 3, at 256–57. After that point, Pardo and Allen argue that the Pardo/Allen theory gives a better account of probative value than does a Bayesian model, relying on their analysis of an illustration from Kahneman and Tversky. Pardo & Allen, *supra* note 3, at 257–61. It is an interesting point but does not seem to have any implications for the burden of persuasion that the remainder of their argument does not suggest. On Kahneman and Tversky’s research program in general, see, for example, Peter M. Todd, *Fast and Frugal Heuristics for Environmentally Bounded Minds*, in *BOUNDED RATIONALITY: THE ADAPTIVE TOOLBOX* 51, 52–55 (2001).

¹²¹ Pardo & Allen, *supra* note 3, at 261–62.

aggregate explanation in favor of the non-moving party could be the better explanation.¹²² To put the summary judgment standard in Pardo and Allen's terms "a judge ought to grant such a motion only if a jury would have to find one side's explanation more plausible than the other side's explanation."¹²³

With respect to Pardo and Allen's interpretation of the preponderance test, the authors accept that "[w]hen you have considered all the evidence in the case, you must be persuaded that it is more probably true than not true," as correct in theory.¹²⁴ They note, however, that jurors seldom receive further guidance about the preponderance test. Jurors, they believe, compare explanations in fact-finding. In that light, Professors Pardo and Allen hold that the courts should clarify civil jurors' task by telling them "to select the best explanation of the evidence (or the most plausible version of the litigated events) and that something has been proven by a preponderance of the evidence if it is part of their selected explanation or version of events."¹²⁵

IV. RESEARCH ON PROBLEM-SOLVING, THE STORY MODEL AND ALTERNATIVE VIEWS OF ABDUCTION AS SOURCES OF DESCRIPTION

Professors Pardo and Allen argued that their theory reflects jurors' actual decision-making much more than probability-based accounts do.¹²⁶ Not surprisingly, they believe that it has significant explanatory and normative implications as well.¹²⁷ There is significant agreement (although hardly unanimity) that jurors do not behave as decision-makers following a comprehensive Bayesian model would. The Pardo/Allen theory seems to be a better description of jury behavior, if only because it takes as an implicit premise that jurors can make good decisions without hearing and assessing all possibly relevant evidence or all conceivable explanations.

¹²² Or, the only plausible explanation if the non-moving party is the prosecution in a criminal case. Pardo & Allen, *supra* note 3, at 262–63. Pardo and Allen's text refers to "devices like summary judgment" as testing the sufficiency of "the explanation that the moving party has provided." Pardo & Allen, *supra* note 3, at 262–63. That must be a mistake because summary judgment and analogous motions test the adequacy of the nonmoving party's evidence, and jurors can consider explanations that they formulate, in the aggregate with whatever the non-moving party may offer. Moreover, if summary judgment focused on the moving party's explanation and the non-moving party did not bear the burden of persuasion, all that party would have to show to avoid summary judgment is that a reasonable jury might conclude that the moving party's aggregate story was no better than the non-moving party's.

¹²³ Pardo & Allen, *supra* note 3, at 265–66.

¹²⁴ Pardo & Allen, *supra* note 3, at 266, (quoting FEDERAL CIVIL JURY INSTRUCTIONS OF THE SEVENTH CIRCUIT 34).

¹²⁵ Pardo & Allen, *supra* note 3, at 267.

¹²⁶ Pardo & Allen, *supra* note 3, at 225–26. *See generally*, Ronald J. Allen, *The Nature of Juridical Proof*, 13 CARDOZO L. REV. 373, 413–20 (1991).

¹²⁷ Pardo & Allen, *supra* note 3, at 239–46.

The Pardo/Allen theory seeks to describe jurors' behavior primarily in terms of a philosophical theory of reasoning under uncertainty—inference to the best explanation. That is not, however, to argue for a rigid distinction between philosophy and empirical research. Indeed, when analyzing reasoning under conditions of uncertainty, it is impossible to draw clean distinctions among the roles of philosophy, cognitive science, psychology and other social sciences.¹²⁸ My suggestion, indeed, is that analysis of the preponderance standard needs a heavier measure of the teachings of empirical research on problem-solving than the Pardo/Allen theory includes. Reconsidering the Pardo/Allen theory in light of that research leads to a somewhat different view of the preponderance standard, and of the justification for enforcement of sufficiency of the evidence standards through summary judgments and judgments as a matter of law.

Professor Harman, who first advanced the concept of inference to the best explanation, later likened it to the practical reasoning process of “satisficing,” rather than maximizing or optimizing.¹²⁹ “Satisficing” is the term for choosing “an alternative that meets or exceeds specified criteria” when “it is impossible to optimize, or where the computational cost of doing so seems burdensome.”¹³⁰ Constraints on jurors' time, knowledge, access to evidence and cognitive capacities mean that jurors (or the legal system itself) can only satisfice in the resolution of issues of fact—they cannot hold out for perfection. That is not to say that if they can achieve improvement, or even seeming perfection, without wasting cognitive resources, they should not do so. Whenever the benefit of expending resources outweighs their costs—including opportunity costs—the fact finder should expend them. That may include formulating stories of their own. At the same time, however, when there is nothing they can do to increase the comprehensiveness of the evidence or explanations they can consider, jurors' experience as decision makers in everyday life may suggest to them that they lack information that they should have before taking a particular action, i.e., entering a particular verdict.

The discussion in this section will begin with a discussion of constraints on decision-making resources, and the cognitive strategies and default rules that we use to make good decisions despite those constraints. Professors Pennington and Hastie developed the story model some years ago. Their research concluded that jurors use stories or narratives as strategies for

¹²⁸ See, e.g., JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES (Daniel Kahneman, Paul Slovic & Amos Tversky, eds. 1982); NICHOLAS RESCHER, COGNITIVE ECONOMY: THE ECONOMIC DIMENSION OF THE THEORY OF KNOWLEDGE (1989); PAUL THAGARD, COMPUTATIONAL PHILOSOPHY OF SCIENCE (1988).

¹²⁹ GILBERT HARMAN, CHANGE IN VIEW: PRINCIPLES OF REASONING 68 (1986).

¹³⁰ See HERBERT A. SIMON, MODELS OF BOUNDED RATIONALITY: EMPIRICALLY GROUNDED ECONOMIC REASON 295 (1997).

organizing, understanding, and assessment of evidence, despite cognitive constraints. Pennington and Hastie's research on that use of stories, which Pardo and Allen suggest serve the function of explanations in the Pardo/Allen theory,¹³¹ suggests that the Pardo/Allen theory's view of the preponderance test may place excessive constraints on both judges and jurors. Moreover, some research on inference to the best explanation suggests that the Pardo/Allen theory focuses too narrowly on specific alternative explanations at the expense of the possibility that fact finders may regard the explanations or evidence on offer as insufficient to support a judgment as serious as the one they are asked to make.

A. *Stories, Cognition, and Confidence in Judgment*

We use what one might call cognitive strategies to organize and evaluate information that bears our decisions, in order to make good judgments in light of the constraints on human memory and the limitations on the time and material resources that we can devote to various issues.¹³² Research on cognition developed a theoretical construct called a schema. Schemata depict the organization of knowledge in memory.¹³³ When we receive new information, schemata help us understand it by reflecting our

¹³¹ Pardo & Allen, *supra* note 3, at 225 n.3, 235 n.35.

¹³² For example, chess players' advantage over relative novices lies in their ability to isolate critical aspects of a position, rather than in some superhuman extension of mental capacity. *E.g.*, Herbert A. Simon & William G. Chase, *Perception in Chess*, in 1 MODELS OF THOUGHT, 386, 386–87, 402 (Herbert A. Simon ed. 1979). Chess masters even seem to organize their memory for the position of the pieces in ways that reflect actual gameplay, resulting in improved memory of actual positions. Herbert A. Simon & Michael Barenfeld, *Information Processing Analysis of Perceptual Processes in Problem Solving*, in 1 MODELS OF THOUGHT at 371. Even Deep Blue, the extremely powerful chess computer that defeated Kasparov could not have been victorious if programmers had not written strategies in its program to limit the number of possible outcomes, *i.e.*, the amount of evidence, it would evaluate before making a move. DANIEL HILLIS, *THE PATTERN ON THE STONE: THE SIMPLE IDEAS THAT MAKE COMPUTERS WORK* 83–87 (1998). Conversely, human players may rely on computers to their detriment. In particular, a player who has memorized a set of computer-generated moves can find his position suddenly falling apart because he does not understand the principles that motivated the moves.

Where problems are more complex, strategies are even more important, and brute processing power may be of relatively little help. For example, in the game of Go, "humans still reign," because there are so many more possible positions in a game of Go. HILLIS at 87. The number of possible positions in Go after *two* move-and-response cycles is more than 16 times the number of possible positions in chess after *three* such cycles. *See* Callen, *supra* note 24, at 1265 n. 108. Human Go masters' cognitive strategies allow them to find good moves in limited time despite constraints on their cognitive resources, without assessing all of the possible moves. Supercomputers programmed with simple search strategies such as the ones that Deep Blue employed have not been victorious at Go, despite their vast computational capacity. HILLIS, *supra* note 132, at 87.

¹³³ Micheline T.H. Chi & Robert Glaser, *Problem-Solving Ability*, in HUMAN ABILITIES: AN INFORMATION-PROCESSING APPROACH 227, 241 (Robert J. Sternberg ed. 1986).

memories of normal sequences of events, objects, the roles of persons in similar contexts and frequent situations, which we can use to comprehend new situations. When information satisfies preconditions for the operation of a schema, sometimes called “slots,” the schema is activated, and it suggests a decision, or an action such as searching for a piece of evidence that may be critical, or even formulation of a new schema.¹³⁴ It may be more difficult for decision-makers to reach good decisions when a schema is only a very general guide. In some cases, when the schema is too general, or the actor chooses an inappropriate one, it may be impossible to reach a conclusion.¹³⁵ The stories that Pennington and Hastie discuss are essentially the products of schemata, that is, of cognitive strategies.

1. The Story Model

With respect to the preponderance standard, the most salient among cognitive strategies for allocating limited resources effectively¹³⁶ is the story model. (Particularly relevant to the Pardo/Allen theory, Professors Hastie and Pennington have even referred to the story model as an example of “Explanation-based Decision Making.”)¹³⁷ Jurors seem to use strategies that psychologists call the “story model” to organize the evidence they receive and to evaluate the parties’ contentions.¹³⁸ Pennington and Hastie note that cognitive psychologists believe that comprehension of any sort of communication “is inherently a constructive process, even for the simplest discourse.”¹³⁹ In that light, some process of formulation and utilization of stories may be inevitable for any trier of fact.¹⁴⁰ In his study of the role of narrative and intelligence, Professor Schank found that:

In the end all we have, machine or human, are stories and methods of finding and using those stories. Knowledge, then, is experiences and stories, and intelligence is the apt use of

¹³⁴ Chi & Glaser, *supra* note 136, at 241–42. One way “of describing the weight of evidence is to say that it increases with the quality and importance of the slots which evidence fills in the schema.” Craig R. Callen, *Second-Order Considerations, Weight, Sufficiency and Scheman Theory: A Comment on Professor Brilmayer’s Theory*, 66 B.U. L. REV. 715, 723 (1986).

¹³⁵ Chi & Glaser, *supra* note 132, at 242.

¹³⁶ See CHRISTOPHER CHERNIAK, *MINIMAL RATIONALITY* (1986); ROGER C. SCHANK, *THE CONNOISSEUR’S GUIDE TO THE MIND* (1991) and Paul Thagard, *Explanatory Coherence*, 12 BEHAVIORAL & BRAIN SCI. 435 (1989).

¹³⁷ Reid Hastie & Nancy Pennington, *The O.J. Simpson Stories: Behavioral Scientists’ Reflections on The People of the State of California v. Orenthal James Simpson*, 67 U. COLO. L. REV. 957, 957 (1996).

¹³⁸ See, e.g., Pennington & Hastie, *supra* note 23.

¹³⁹ Pennington & Hastie, *supra* note 23, at 523, n.11.

¹⁴⁰ See, e.g., JOHN JACKSON & SEAN DORAN, *JUDGE WITHOUT JURY: DIPLOCK TRIALS IN THE ADVERSARY SYSTEM* 217–21 (1995) (use of a model to explain judicial fact-finding).

2018]

SPOTTING A PREPONDERANCE

1547

experience and the creation and telling of stories. Memory is memory for stories, and the major processes of memory are creation, storage, and retrieval of stories.¹⁴¹

Information about the events in suit is a crucial ingredient for story formation, but not the only one. Jurors also utilize knowledge about similar events and “generic expectations about what makes a complete story (*e.g.*, the knowledge that human actions are usually motivated by goals).”¹⁴² In addition to the evidence offered at trial, they will make inferences to “fill out the episode structure” of the story—whether suggested by parties or on their own initiative.¹⁴³ Individual jurors use pre-existing knowledge, new information and inferences to construct one or more narratives, which they use in reaching their verdict.¹⁴⁴ If jurors could not rely on the organizational structure and links to prior experience that stories provide, the sheer quantity of information that jurors receive and the “disconnected” manner in which they receive it would make the evidence “unwieldy.”¹⁴⁵

The story model takes as a premise that jurors’ use of stories in everyday life gives them the knowledge they can employ to assess the “completeness of the evidence, or the extent to which a story has all its parts.”¹⁴⁶ “Missing information, or lack of plausible inferences about one or more major components of the story structure will decrease confidence in the explanation.”¹⁴⁷ The research concluded that each juror constructs a story to account for the evidence (or “best” story if she constructs alternatives) is the basis of her verdict.¹⁴⁸ One should not read that to say that jurors will necessarily skate on thin narrative ice.

According to the story model, jurors who believe that a particular aspect of a story necessary to establish a claim or defense lacks adequate support will rely on the default rule and find against the party with the burden of proof.¹⁴⁹ If, on the other hand, the evidence makes jurors confident about

¹⁴¹ ROGER C. SCHANK, TELL ME A STORY: NARRATIVE AND INTELLIGENCE 16 (1995).

¹⁴² Pennington & Hastie, *supra* note 23 at 522.

¹⁴³ Pennington & Hastie, *supra* note 23, at 527.

¹⁴⁴ Pennington & Hastie, *supra* note 23, at 522–23.

¹⁴⁵ Pennington & Hastie, *supra* note 23, at 523.

¹⁴⁶ Pennington & Hastie, *supra* note 23, at 527.

¹⁴⁷ Pennington & Hastie, *supra* note 23, at 528.

¹⁴⁸ Hastie & Pennington, *supra* note 137, at 959–60.

¹⁴⁹ Pennington & Hastie, *supra* note 23, at 530–31. Jurors will similarly decide in favor of the default if the story does not coincide with one or more elements of the claim or defense. (“We allow for the possibility that the best story is not good enough or does not have a good enough fit to any verdict option and, therefore, a default verdict would have to be available.”) *Id.* at 531 n. 24; Nancy Pennington & Reid Hastie, *The Story Model for Juror Decision Making*, in INSIDE THE JUROR: THE PSYCHOLOGY OF JUROR DECISION MAKING 192, 201 (1993).

their judgment, and fills in necessary elements of the story, they may reach a verdict by comparing the likelihood of stories.¹⁵⁰

While the story model holds that development of a story or explanation precedes the jury's decision, it is possible for the jurors to go through a number of cycles of story formation before reaching a final verdict. In such a situation, the juror's tentative decision may influence the final story formation.¹⁵¹ Regardless, story formation continues through the trial, and during jury deliberation.¹⁵² In contrast to the Pardo/Allen theory, the story model theory does not assume that the jurors will construct at least one explanation in favor of each party, although it does recognize that jurors may construct more than one story and pick the best.¹⁵³

2. Oblique-Stories

As Professors Pennington and Hastie conceive of them, stories may explain the origin or value of the evidence without directly showing how a crime, or central events in controversy, occurred. That is, they may be designed to discredit the opposing evidence. For example, in discussing the O.J. Simpson case, Professors Hastie and Pennington discuss two stories that the defense advanced to account for some or all of the evidence without giving an explanation of events leading to the death of Nicole Simpson or Ronald Goldman. The first explanation was that the investigators prematurely concluded that Simpson was the murderer and were motivated to frame Simpson. The second was that LAPD forensic investigators simply bungled evidence collection.¹⁵⁴ One might call those stories "oblique-stories," that is, explanations of the evidence focusing on how the opposing party developed items of evidence in a manner that is not, strictly speaking, logically inconsistent with any ultimate fact that the opposing party alleged. One test of the quality of a descriptive theory is whether it "saves the data,"

¹⁵⁰ Nancy Pennington & Reid Hastie, *Evidence Evaluation in Complex Decision-Making*, 51 J. PERSONALITY & SOC. PSYCHOLOGY 242, 245, 254 (1986).

¹⁵¹ Pennington & Hastie, *INSIDE*, *supra* note 149, at 201–02.

¹⁵² Pennington & Hastie, *INSIDE*, *supra* note 149, at 201–02. Pardo and Allen agree that story formation may continue through the trial. Pardo & Allen, *supra* note 3, at 234. It follows that their references to the "self-evidencing" nature of explanations are somewhat misleading. (For examples see Pardo & Allen, *supra* note 3, at 233, 235.) Similarly, their claim that evidence is "relevant if it is explained by the particular party offering the evidence" is flawed. If explanations are under construction as evidence is offered, evidence must be relevant if it confirms or weakens an existing explanation or could lead jurors to trigger the formation of a new story.

¹⁵³ Pennington & Hastie, *supra* note 23, at 531 n.24; Pennington & Hastie, *INSIDE*, *supra* note 149, at 201. Professors Vidmar and Hans quote transcripts from the Arizona jury study that tend to suggest that jurors constructed only one story in their deliberations. *See, e.g.*, NEIL VIDMAR & VALERIE P. HANS, *AMERICAN JURIES: THE VERDICT* 135–40, 184–87, 292–95 (2007).

¹⁵⁴ Hastie & Pennington, *supra* note 137, at 966–67.

that is, whether it accounts for the empirical phenomena.¹⁵⁵ Regarding oblique-stories as stories that the trier should compare to narratives that include the ultimate facts could require jurors complying with the Pardo/Allen theory to act in a way that is inconsistent with the theory.

The use of oblique-stories in the story model suggests a flaw in the Pardo/Allen theory as a descriptive theory. The Pardo/Allen theory holds that the stories jurors may consider include stories they formulate (in part from scraps of the parties' stories, but clearly with some material or imagination of their own). Hastie and Pennington say that accounts of the evidence that depict it as fabricated or flawed count as stories. Those two points seem to add up, at a minimum, to the conclusion that, in deciding whether any one story is better than the others, jurors should consider doubts they have about the evidence or the explanation that the parties did not raise. If such doubts form part of a story that the jurors might consider the best one, then the Pardo/Allen theory's requirements for a preponderance mean little more than limiting the competing accounts that jurors may consider on behalf of the non-burden bearing party to theories that jurors can articulate to themselves. That is, where a preponderance is more likely than not, the Pardo/Allen theory might limit the stories that jurors can consider on behalf of not to articulated explanations in favor of that party. At most, it would suggest that the burden of persuasion requires that evidence and explanations in favor of the burden-bearing party be (i) more likely in the aggregate than those explanations articulated in favor of the non-bearing party, and (ii) sufficiently comprehensive to overcome reluctance to abandon the default rule stemming from absence of evidence or doubts about the quality of the explanations in favor of the burden-bearing party.

B. *How Theories of Explanation Can Account for Lack of Confidence*

An inference to the best explanation is often called abduction, a term that Professors Pardo and Allen use as well.¹⁵⁶ Umberto Eco, whose scholarship includes work on abduction, suggests that when we use abduction to create an explanation of facts, as is usually the case in litigation, we must also engage in meta-abduction, to ask whether the explanation we imagine matches up with the real world as we know it.¹⁵⁷ Regardless of our

¹⁵⁵ See Pardo & Allen, *supra* note 3, at 243 (discussing work of Bas van Fraassen).

¹⁵⁶ WALTON, *supra* note 74, at 34; Pardo & Allen, *supra* note 3, at 228. That is not to say that equating them is universal, nor that everyone accepts the distinction between, e.g., enumerative induction and inference to the best explanation. Pardo and Allen equate abduction and inference to the best explanation and distinguish between enumerative induction and abduction, and possible differences among the theories do not seem to be important here.

¹⁵⁷ Umberto Eco, *Horns, Hooves, Insteps: Some Hypotheses on Three Types of Induction*,

cognitive interests, we must ask whether we consider the explanation and its supporting evidence adequate for action.

Detectives are rewarded by society for their impudence in betting by meta-abduction, whereas scientists are socially rewarded for their patience in testing their abductions. Naturally, in order to have the intellectual and moral force to test, and to ask for new tests, and to entertain stubbornly an abduction before it has been definitely tested, scientists also need meta-abduction. Their difference from detectives stands in their refusal to impose their beliefs as a dogma, in their firmness not to repudiate their motivated conjectures.¹⁵⁸

John Josephson argues that, in order for inference to the best explanation to work well, it is important to consider the exhaustiveness of the set of explanations to which we compared the eventual winner.¹⁵⁹ His position is that before we decide that the superiority of explanation E_i , as compared to explanations E_{ii} - E_n , warrants taking E_i to be true or acting on it, we should consider whether the group E_{ii} - E_n exhausts the alternatives. To take an example from Conan Doyle, through Sherlock Holmes, “It is an old maxim of mine that when you have excluded the impossible, whatever remains, however improbable, must be the truth.”¹⁶⁰ That assumes that the search for other evidence and explanations has been sufficiently thorough to make us confident that “whatever remains” is true, or at least an acceptable basis for action.

In particular, Josephson considers it important that the decision maker comparing explanations include two special sorts of stories. The first is what he calls the “noise” hypothesis, that all of the evidence is incorrect, or just noise.¹⁶¹ “[I]t may be that what seems to require explanation is merely coincidence misperception, miscategorization, fraud, perjury . . . or some similar phenomenon. Sometimes the data should be ‘explained away’ rather

in THE SIGN OF THREE: DUPIN, HOLMES, PEIRCE 198, 207, 215–20 (Umberto Eco & Thomas Sebeok eds. 2002).

¹⁵⁸ *Id.* at 220.

¹⁵⁹ John R. Josephson, *On the Proof Dynamics of Inference to the Best Explanation*, in THE DYNAMICS OF JUDICIAL PROOF: COMPUTATION, LOGIC AND COMMON SENSE 287, 293 (2002).

¹⁶⁰ Arthur Conan Doyle, *The Adventure of the Beryl Coronet*, in THE ADVENTURES OF SHERLOCK HOLMES 476 (The Floating Press 2008). The idea might have come down from Poe. See explanatory notes in THE ADVENTURES OF SHERLOCK HOLMES, *supra*, at 390. Conan Doyle repeats the point a number of times. *E.g.*, ARTHUR CONAN DOYLE, THE SIGN OF FOUR 29 (n.d.) (ebook) (“How often have I said to you that when you have eliminated the impossible, whatever remains, however improbable, must be the truth?”).

¹⁶¹ John R. Josephson, *On the Proof Dynamics of Inference to the Best Explanation*, in THE DYNAMICS OF JUDICIAL PROOF: COMPUTATION, LOGIC AND COMMON SENSE 286–93 (Marilyn MacCrimmon & Peter Tillers, eds., 2002).

than explained.”¹⁶² What I referred to earlier as oblique-stories in the Simpson case would fall into this classification.

The second of Josephson’s “special” explanations is really more of a meta-explanation. That sort of hypothesis would posit that the set of explanations that the decision maker is considering rely on past experience, and as a result may be inadequate to explain the current situation, which is a new possibility. A trier with little or no experience may fail to make sufficient allowance for that lack of experience and jump to conclusions. A case-hardened fact finder, on the other hand, may be overconfident and prematurely limit the alternatives he or she is considering. In assessing the degree to which an unspecified new explanation may be better than the extant ones, it may be possible to estimate the strength of the potential new explanation by analogy to other situations in which the trier was able to form a new explanation, and by assuming that the new explanation would account for current anomalies in the evidence.¹⁶³

Professor Thagard similarly said that his theory of explanatory reasoning needed to be “enriched by taking into account judgments about the evidence and availability of hypotheses” and that in dealing with novel problems a prudent decision maker might suspend acceptance of what might be the only explanatory hypothesis. He suggested that the decision maker might, in fact, entertain a sort of “place-holder hypothesis” as competition for the only fully developed candidate hypothesis.¹⁶⁴

In a response to a related article by Professor Pardo, Professor Laudan argued that to qualify as an acceptable “best” explanation, an explanation must not only be better than any competitor, but must “exhibit a high degree of coherence with background beliefs, . . . must be internally consistent, . . . should explain different kinds of facts, . . . must be simple, and so on.” In the absence of those traits, he argues, inference to the best explanation does not justify any inference except possibly agnosticism on the factual issues in question.¹⁶⁵

¹⁶² *Id.* at 293.

¹⁶³ *Id.* at 293–94. Judge Posner notes that juries help to mitigate judges’ tendency to use short cuts based on their experience to reach decisions. POSNER, *supra* note 35, at 350–51.

¹⁶⁴ Paul Thagard, *Defending Explanatory Coherence*, 14 BEHAV. & BRAIN SCI. 745, 745–46 (1991).

¹⁶⁵ Laudan, *supra* note 15, at 304–05. Laudan does, however, seem to agree with Professor Allen’s earlier relative plausibility theory with respect to the preponderance standard. “Under current rules, if the plaintiff has a better story than the defendant, he must win even when his theory of the case fails to satisfy the strictures required to qualify his theory as the best explanation.” Laudan, *supra* note 15, at 304–05.

Eco's, Thagard's, and Josephson's arguments suggest that inference to the best explanation requires decision-makers to consider not only whether the specific stories in favor of one party are better than those in favor of the other, but the extent to which the consideration of explanations is sufficiently exhaustive. In Pennington and Hastie's terms, one might think of that judgment of exhaustiveness as a factor in one's confidence in the specific stories. To put the point another way, Eco, Josephson, Thagard and the story model all suggest that the jury should decide in favor of the party that does not have the burden of persuasion when the burden-bearing party's story is simply not good enough, even if it is better than the story of the burden-bearing party—who might have offered little or nothing. Studies of cognition and the Court's jurisprudence on the sufficiency of the evidence in civil cases seem to agree. Moreover, burdens of proof and rulings on the sufficiency of the evidence suggest that standards more stringent than the Pardo/Allen theory's preponderance test have social benefits that judicial implementation of the Pardo/Allen theory would not.

C. *The Utility of Burdens and Sufficiency Tests*

1. As a Guide for Jurors

In addition to the cognitive strategies that yield stories, our decision-making strategies include defaults or *ceteris paribus* principles.¹⁶⁶ Such rules help us allocate those resources efficiently in at least three ways.¹⁶⁷ First, they simplify inferences about the empirical world which we must draw from information at hand by focusing our attention on critical points—the conditions under which the default rule no longer applies. Second, given a particular state of knowledge, they help us decide whether additional information is important, and to limit our search for that information in our memory or in the world. Third, given our inferences and information about the empirical world, and about the possible costs and benefits of various actions, they help us decide which actions to take.

¹⁶⁶ I meld together here, with default rules, a number of constructs including “fast and frugal heuristics,” (Gerd Gigerenzer & Peter M. Todd, *Fast and Frugal Heuristics: The Adaptive Toolbox*, in SIMPLE HEURISTICS THAT MAKE US SMART 12–18 (Gerd Gigerenzer & Peter M. Todd, eds. 1999)); “rule-based realizations” of “q-morphisms” (J. HOLLAND, K. HOLYOAK, R. NISBETT & P. THAGARD, INDUCTION: PROCESSES OF INFERENCE, LEARNING AND DISCOVERY 46–47 (1986)), and frames with defaults (M. MINSKY, THE SOCIETY OF MIND 243–52 (1986)). See generally, Callen, *Adjudication*, *supra* note 29; Callen, *Cognitive Science*, *supra* note 39.

¹⁶⁷ See Callen, *Adjudication*, *supra* note 29, at 480–81; Callen, *Cognitive Science*, *supra* note 39, at 1118, 1120–21.

The burden of proof standards¹⁶⁸ that lawmakers establish help to define fact finders' task by serving as default decision-making standards for the fact-finding function. They require fact-finders to make one particular decision unless the burden-bearing party provides the requisite quantum of evidence for a fact-in-issue, which may trigger another. By isolating a few critical variables, those facts-in-issue or elements on which the applicability of rules turn, they simplify the fact-finder's task. At the same time, focusing the fact-finder's task more narrowly helps lawmakers, courts or legislators to promote predictable results.¹⁶⁹ Isolating the critical factual inquiries also helps lawmakers to condition rules on the facts they consider most critical, and to establish a default in favor of what they consider the generally desirable outcome.¹⁷⁰ Not to put too fine a point on it, courts and legislatures in formulating legal criteria make the judgment that the factual preconditions of that rule describe a point at which the enforcement of the rule has social utility that other preconditions would not. In that light, society has an interest in accurate and predictable determination of whether those preconditions exist.¹⁷¹ Finally, burdens of proof help organize disputes where multiple rules are involved. For example, rule *R* may apply unless party *P* offers quantum of evidence *Q* of facts *F*₁-*F*₄. If *P* does, in fact, offer the evidence, then *R* applies unless opponent *O* offers quantum of evidence *Q* of facts *F*₅-*F*₈, and so on.¹⁷²

The argument for the comparative standard implicitly ignores the extent to which human decision-makers who think *P*'s story is more plausible than *D*'s, might still hesitate to rely solely¹⁷³ on the difference in the stories. While relying on the better story might lead to the correct decision in many cases, where evidence for any story is thin, in particular when it is thin given the importance of the matter at hand,¹⁷⁴ we might well hesitate to rely on the relative superiority of *P*'s story. Applying that idea in the context of

¹⁶⁸ Including both burdens of persuasion and production, as well as summary judgment standards, which are functions of the burden of persuasion. See *Anderson v. Liberty Lobby, Inc.*, 477 U.S. 242, 253–55 (1986).

¹⁶⁹ See, e.g., Callen, *Adjudication*, *supra* note 29, at 483–85.

¹⁷⁰ See, e.g., Callen, *Adjudication*, *supra* note 29, at 483–85.

¹⁷¹ E.g., POSNER, *supra* note 35, at 341.

¹⁷² See, e.g., Callen, *Cognitive Science*, *supra* note 39, at 1120–21. See the discussion of the *St. Mary's* case, *infra* notes 192 and accompanying text for an example of such an allocation.

¹⁷³ This reservation is important because one might also choose to rely on a default rule that would ordinarily operate in *A*'s favor.

¹⁷⁴ See Ronald J. Allen, Mark F. Grady, Daniel D. Polsby & Michael S. Yashko, *A Positive Theory of the Attorney-Client Privilege and the Work Product Doctrine*, 19 J. LEG. STUD. 359, 388 (1990) (suggesting that the social interest in the accurate application of rules warrants use of resources in litigation, i.e., makes the accumulation of evidence socially useful. Litigation is not merely a stakes-dividing device.).

litigation, jurors might hesitate to abandon a default standard until they had enough information indicating that the balance of utilities tipped in favor of a decision for *P*.¹⁷⁵

Professors Pardo and Allen might argue that courts and fact-finders have no valid interest in examining the completeness of evidence, or the exhaustiveness of the explanations available, so there may be no reason for a fact-finder to apply any standard other than pair-wise comparison.¹⁷⁶ No one doubts that jurors will often conclude that the parties' presentations are comprehensive, or at least adequately representative. The Pardo/Allen theory holds that we should *instruct* jurors to rely on their pair-wise comparisons, not that we should allow them to do so. That argues that jurors should set aside some of the lessons of their cognitive experience, even though the likelihood that they will rely on those lessons relates to an important justification of the jury—their ability to bring everyday experience to bear on questions of fact.¹⁷⁷

Social utility may require more a standard more exacting than a comparison of stories, theories, or explanations, particularly given the rather vague nature of The Pardo/Allen theory's concept of instantiating an element of a claim.¹⁷⁸ The incentives that parties in a given action have to gather evidence, and to formulate stories or arguments, may not correspond to the societal benefit of increasing the amount of information, or of refraining from applying a particular rule in the absence of further information. The parties may have only a small amount at stake or may hesitate to present evidence with uncertain implications for their interests.¹⁷⁹ As a general rule, the parties have no interest in gathering evidence to vindicate broad social norms or to see to it that outcomes at trial are consistent. Usually, their interest is limited to their own dispute.¹⁸⁰ Thus, the courts, in common law

¹⁷⁵ One may argue that terms such as “enough” or “sufficient” do not describe a quantum of evidence with adequate precision. Allen, *supra* note 44, at 1106. Given, though, that (i) computational intractability inhibits the use of mathematical or logical description to model human thought processes and (ii) questions of the adequacy of evidence can arise in differing contexts, it is difficult to sharpen those terms. See Allen, *The State of Mind Necessary for a Judicial Verdict*, 13 CARDOZO L. REV. 485 (1991) (“No logical algorithms can capture, and thus guide very precisely, juror decision-making, . . . with the exception of formal systems such as Euclidean geometry.”).

¹⁷⁶ Cf. Allen, *supra* note 2, at 428.

¹⁷⁷ See, e.g., *Sioux City & Pac. Ry. Co. v. Stout*, 84 U.S. 657, 664 (1873); Pattern Civ. Jury Instr. 5th Cir. 2.18 (2014). Cf. Pardo & Allen, *supra* note 3, at 40 (“[J]urors are necessary to provide the background knowledge to make contextual judgments about the strength of competing explanations and to suggest new ones.”).

¹⁷⁸ See Pardo & Allen, *supra* note 3, at 31–34.

¹⁷⁹ Callen, *Adjudication*, *supra* note 29, at 487. See POSNER, *supra* note 35, at 349 (“[T]he private benefits of searching for evidence may exceed or fall short of the social benefits.”).

¹⁸⁰ Callen, *Adjudication*, *supra* note 29, at 487. See POSNER, *supra* note 35, at 349 (“[T]he private benefits of searching for evidence may exceed or fall short of the social benefits.”).

and interpretive functions, have good reason to use sufficiency tests more exacting than those the comparative standard entails. Jurors, concerned about their own interests as members of a broader society, and relying on their own cognitive standards, have similar reasons to be concerned about the adequacy of evidence before reaching a verdict.

Legal rules themselves reflect judgments about the need to consider adequacy of evidence, even in everyday decision-making. The reasonable person standard in, for instance, negligence law, reflects the need to apportion cognitive resources in light of the task at hand—to exercise reasonable care. Professor Cherniak points out that the reasonable person standard requires persons to commit cognitive resources, to endeavor to obtain information, to the extent warranted by the potential risk and its likelihood.¹⁸¹ It follows that one can make, and that ordinary people often do make, judgments about the adequacy of evidence, rather than relying exclusively on the better of two or more stories as the Pardo/Allen theory would have it. An account of reasoning which includes judgments about the adequacy of evidence is indispensable to any understanding of rationality which accommodates an understanding of human thought processes and yet assumes that interpersonal standards are meaningful.

2. Sufficiency Tests and Errors of Law

We have been considering the effect of limited cognitive resources on the jury's role in fact finding. The judicial system likewise has limited resources: time, personnel, money, and the human capacities of judges and jurors. Moreover, to the extent the court considers the effect of its decisions on the other human and economic resources of the society, the allocation of those resources¹⁸² is constrained not only by their absolute limits, but also by the conflicting demands on the resources. Those resources can be devoted to the production of goods and services, or to other activities which may lead to greater societal benefits than would the resolution of conflicts in a judicial forum. Directed verdicts and summary judgments limit the expenditure of resources in on-going litigation,¹⁸³ as do other devices for early disposition of litigation.

Judicial procedure in civil disputes, including evidentiary rules, is the means by which courts organize data about disputes for resolution of those disputes in accord with the applicable legal rules, setting general criteria for state action through adjudication,¹⁸⁴ not merely a method for conflict

¹⁸¹ CHERNIAK, *supra* note 136, at 102–03.

¹⁸² Professor Lempert points out that unnecessary litigation costs are deadweight losses. Lempert, *supra* note 32, at 470.

¹⁸³ Lempert, *supra* note 32, at 470..

¹⁸⁴ See John H. Leubsdorf, *Constitutional Civil Procedure*, 63 TEX. L. REV. 579, 594–96,

resolution as the Pardo/Allen theory suggests. Unless one looks at the court's role in fact-finding as part of its role in applying and interpreting rules of law,¹⁸⁵ it is difficult to make sense of common law adjudication, or for that matter, interpretation of rules and statutes. Each requires the court to decide whether the facts in a case fit a particular rule, or whether the facts are sufficiently analogous to precedent or the pre-conditions of the rule or statute to require or allow the court to apply the pre-existing principle.

Directed verdicts and summary judgments are two of the procedural means by which courts decide that a party's evidence is simply not good enough to trigger, or prevent, the application of a rule. Directed verdicts and summary judgments set a threshold, in terms of the non-moving party's ability to satisfy the burden of persuasion, for measurement of the accuracy of the non-moving party's case against the information in society at large. In other words, they require that the courts see that facts have a certain closeness of fit with the premises of rules and statutes before applying them.

Courts use burdens of persuasion and production to refine and interpret rules of law. Whenever the courts accept a rule, statute, or interpretation thereof as a valid prescription for their action, the prescription must necessarily be based on assumptions about the way events normally occur in the world.¹⁸⁶ The court must make such assumptions in order to focus its effort, and the litigation, on determining whether the critical triggering conditions for sanctions exist. The allocation of burdens of proof enables the court to implement its interpretation of the rule to fit the world—by making default assumptions about the nature of events or evidence thereof. It is those assumptions, among others,¹⁸⁷ on which the court relies in assigning burdens of proof. The allocation of burdens of proof, particularly burdens of production, allows the court to conserve decision-making resources for the abnormal or problematic case. When the court seeks to create or reinforce incentives for particular behavior, it has a further interest

603 (1984). Cf. Owen Fiss, *The Social and Political Foundations of Adjudication*, in ROBERT M. COVER, OWEN M. FISS & JUDITH RESNICK, *PROCEDURE* 219, 222 (1988) (Structural litigation which enforces public values tends to abandon dispute resolution model of procedure).

¹⁸⁵ Of course, a court can formulate a new rule. Although not completely distinct from interpretation or application of rules, that possibility does not seem to affect the argument, so I have set it aside.

¹⁸⁶ Consequently, they have valid reasons, in some situations, to refuse to allow verdicts to be based on some proffers of statistical evidence.

¹⁸⁷ On the state's interest in accurate enforcement, *see, e.g.*, *Addington v. Texas*, 441 U.S. 418, 426 (1979). Predictability is desirable for another reason—to establish an expectation that the legal system will predictably apply the incentive in the future, which in turn makes incentives more effective. Cf. ROBERT AXELROD, *THE EVOLUTION OF COOPERATION* 112–17 (1984) (In iterated prisoner's dilemma simulation, most effective strategy is not the first to deviate from mutually beneficial course—unwillingness to be first avoids unnecessary conflict).

in promoting accuracy and predictability for substantive reasons, to see that the incentives are properly allocated.¹⁸⁸

The courts, in that light, have an interest in seeing that the party who contends that the courts should apply rules in that party's favor, action which affects the allocation of extra-judicial resources, has sufficiently proven the elements of those rules. Accordingly, as lawmakers and agents of the state, the courts have an interest in promoting accuracy in enforcement.¹⁸⁹ They also have an interest in promoting consistency of enforcement, to clarify incentives for actors in society at large to avoid the acts which give rise to evidence which, in turn, can result in conviction or liability.

D. *Cognition, The Pardo/Allen Theory, and the Court's View of Fact Finding*

1. The Fact Finder's Story in *St. Mary's*

I do not doubt that jurors in many civil cases use one or more of the parties' stories to a great extent in understanding and assessing evidence. Moreover, it seems likely that some jurors assume that each party has put forward the best explanation of the evidence that is in its own favor, and refrained from constructing or adapting a story on their own initiative.¹⁹⁰

Fact finders do, however, formulate their own stories to explain the evidence. For example, suppose that plaintiff, an African-American, brought an action alleging that his former employer discriminated against him on the basis of race in discharging him from employment. Defendant offered a neutral explanation for the discharge under the *McDonnell-Douglas* line of cases, that plaintiff failed to discipline his subordinates adequately. In

¹⁸⁸ See, e.g., Crispin Wright, *Wittgenstein's Rule-Following Considerations and the Central Project of Theoretical Linguistics*, in REFLECTIONS ON CHOMSKY 255 (A. George ed. 1989). For example, William H. Page, *The Chicago School and the Evolution of Antitrust: Characterization, Antitrust Injury and Evidentiary Sufficiency*, 75 VA. L. REV. 1221, 1257–74 (1989) points out the effect of the Chicago School model in antitrust law on questions of characterization of behavior, antitrust injury and evidentiary sufficiency. Communication in general requires that speakers or authors on the one hand, and audiences on the other, share a large number of assumptions about the empirical world. Otherwise, the audience would not be able to understand the content of the communication. See DAN SPERBER & DEIRDRE WILSON, RELEVANCE: COMMUNICATION AND COGNITION 38–51 (1986).

Edmund M. Morgan, *Some Observations Concerning Presumptions*, 44 HARV. L. REV. 906, 910–11 (1931), points out that other, sometimes conflicting, rules of thumb can affect the fixing of the burden of persuasion. “These respectively make it fall upon (1) the party having the affirmative of the issue, (2) the party to whose case the fact in question is essential, (3) the party having peculiar means of knowing the fact, and (4) the party who has the burden of pleading it.” The point in the text is not that the court always relies on a particular principle in allocating the burdens of persuasion or production. It is, rather, that burdens of production and persuasion are important for accuracy in fact finding and application of legal criteria.

¹⁸⁹ See POSNER, *supra* note 35, at 340–41; Leubsdorf, *supra* note 184, at 596–97.

¹⁹⁰ See Pardo & Allen, *supra* note 3, at 234 n.30.

addition, it showed that two African-Americans sat on the employee board that approved disciplinary action against plaintiff, and that the number of its African-American employees remained constant. Nevertheless, the plaintiff offered evidence that led the fact finder to conclude that the neutral explanation (firing for misconduct) was false, given that other employees who failed to discipline subordinates received less severe penalties. Under those circumstances, could the fact finder find that the discharge was non-discriminatory?

Readers familiar with Title VII jurisprudence will note the hypothetical's resemblance to *St. Mary's Honor Center v. Hicks*.¹⁹¹ There, the District Court held that plaintiff failed to prove by a preponderance that racial discrimination was the motivation for plaintiff's discharge¹⁹² rather than personal hostility to him (which the defendant did not offer as an explanation for the discharge). While the court recognized that disbelief of the defendant's explanation would tend to show intentional discrimination,¹⁹³ it reasoned that, even if the trier thought the employer was lying about the neutral explanation, such a conclusion would not automatically preclude the trier from finding that there was another nondiscriminatory reason for the discharge.¹⁹⁴

We do not know all the details of the trial, but *St. Mary's* seems to be a good example of a case in which the fact finder worked out a story on his own, at least to a significant extent. Although they have not said so, Pardo and Allen would be very likely to accept *St. Mary's* as a situation in which the trier developed an explanation of his own for the evidence, and thought it was at least as good as the one that the plaintiff offered. Nevertheless, it does demonstrate a situation in which the judge as trier of fact found the explanations offered wanting, and so formulated one of his own.

Comparing *St. Mary's* to the Pardo/Allen theory, one might pose the following hypothetical. Suppose that jurors applying the preponderance standard regarded explanations in favor of each party in a civil dispute as possible, and the plaintiff's as best, but ultimately thinner than they would rely on in their own lives. Should the preponderance standard require them to find in favor of the plaintiff? The Pardo/Allen theory seems to say yes.¹⁹⁵ As a descriptive matter, this does not seem to save the data. In particular, it is very difficult to square with current practice in regard to motions for

¹⁹¹ 509 U. S. 502, 508 & n.9 (1993).

¹⁹² *Hicks v. St. Mary's Honor Ctr.*, 756 F. Supp. 1244, 1252 (E.D. Mo. 1991).

¹⁹³ *St. Mary's*, 509 U.S. at 511; *see also* *Reeves v. Sanderson Plumbing Prods. Inc.*, 530 U.S. 133, 148–49 (2000) (evidence sufficient to find employer's explanation to be a pretext, together with evidence showing a *prima facie* case of intentional discrimination, can be sufficient to support a finding of age discrimination).

¹⁹⁴ *St. Mary's*, 509 U.S. at 520–21.

¹⁹⁵ *E.g.*, Pardo & Allen, *supra* note 3 at 246–47, 267–68.

2018]

SPOTTING A PREPONDERANCE

1559

summary judgment and judgment as a matter of law.

2. *Celotex, Reeves and The Pardo/Allen Theory*

In *Celotex Corporation v. Catrett*,¹⁹⁶ the Court held that a movant for summary judgment who did not bear the burden of persuasion need not produce any affidavits in support of her contentions of fact, but, instead, may obtain summary judgment

after adequate time for discovery and upon motion, against a party who fails to make a showing sufficient to establish the existence of an element essential to that party's case, and on which the party will bear the burden of proof at trial. . . . "[T]h[e] standard [for granting summary judgment] mirrors the standard for a directed verdict under Federal Rule of Civil Procedure 50(a)."¹⁹⁷

In *Reeves v. Sanderson Plumbing Prods, Inc.*,¹⁹⁸ the Court considered whether the district court correctly refused to enter a judgment notwithstanding the verdict under Federal Rule 50(a) after a plaintiff's verdict in an age discrimination action. In upholding the district court's judgment and reversing the court of appeals, the Court said that, when a trial judge assesses whether the nonmoving party's evidence was sufficient to support a finding in its favor on

a motion for judgment as a matter of law, the court should review all of the evidence in the record. In doing so . . . the court must draw all reasonable inferences in favor of the nonmoving party, and it may not make credibility determinations or weigh the evidence. . . . Thus, although the court should review the record as a whole, it must disregard all evidence favorable to the moving party that the jury is not required to believe.¹⁹⁹

Neither *Celotex* nor *Reeves* seems to reflect a comparative approach to the standard of proof. If the standard of proof asked whether the plaintiff's story was better than the defendant's, then *Celotex* would require the defendant to offer some affirmative evidence in support of its own motion. Otherwise, with the exception of complete absence of evidence for the non-moving party or certainty that the defendant's story was right,²⁰⁰ a motion

¹⁹⁶ 477 U.S. 317 (1986).

¹⁹⁷ *Id.* at 232–33 (quoting *Anderson v. Liberty Lobby, Inc.*, 477 U.S. 242, 250 (1986)) (brackets in original).

¹⁹⁸ 530 U.S. 133 (2000).

¹⁹⁹ *Id.* at 150–51 (citations omitted).

²⁰⁰ Neither was the case in *Celotex*. *Celotex*, 477 U.S. at 320. In *Catrett v. Johns-Manville*

for summary judgment without a showing of facts in favor of the moving party would be completely pointless. On the other hand, suppose the Court in *Reeves* had conceived of the burden of proof standard as comparative, that is, as a test of whether a reasonable juror could consider the nonmoving party's explanation as better to any alternatives. It would have made no sense to restrict the evidence that the district court could consider in favor of the movant to evidence that the jury was required to believe. Evidence that the jury was not required to believe might nevertheless be a potent ingredient in a powerful story for the movant.

Of course, defendants or other non-burden-bearing parties who make motions for summary judgment under *Celotex* may argue that the plaintiff's evidence is consistent with any number of stories. That still would not make the defendant's story any better—unless we think of defendant's story as “the plaintiff doesn't have enough evidence to support a verdict.” There is no reason to believe that the Pardo/Allen theory encompasses such a story. So that possibility does not seem to give the Pardo/Allen theory any sustenance.

3. When Evidence or Explanations Are Lacking

The Pardo/Allen theory contends that, in cases in which “neither party offers a particularly plausible explanation of the evidence . . . because neither side can explain key pieces of evidence,”²⁰¹ the jury should resolve the situation by asking what is the best explanation and ruling accordingly. If there is no best explanation, the trier should find against the party with the burden of persuasion.²⁰² The authors contrast situations in which “there is such a paucity of evidence that it can be explained in multifarious ways.”²⁰³ The Pardo/Allen theory requires that where, as in the first case, all explanations are equally bad, the trier should find against the party with the burden of persuasion.²⁰⁴

Further, analyzing what they seem to regard as a typical problem in civil litigation, they say that, in cases in which the trier believes the probability of the plaintiff's²⁰⁵ story is less than .5, but still substantially

Sales Corp, 826 F.2d 33, 37–41 (D.C. Cir. 1987), the court of appeals held on remand that the plaintiff submitted evidence sufficient to allow a reasonable jury to conclude that she satisfied the burden of persuasion.

²⁰¹ Pardo & Allen, *supra* note 3, at 237–38.

²⁰² That assumes, of course, that at least one of the explanations of equivalent quality favors each party.

²⁰³ Pardo & Allen, *supra* note 3, at 237–38.

²⁰⁴ Pardo & Allen, *supra* note 3, at 237–38.

²⁰⁵ Assuming that the burden of persuasion is on the plaintiff. Pardo & Allen, *supra* note 3, at 246–47, once again, argue that judgment should go for the defendant if the evidence does not enable the jury to differentiate among explanations.

greater than the probability of the defendant's. Hypothesizing a situation in which the trier believes the probability of the plaintiff's case is .4 and the defendant has two defenses, which each have a probability of .1, they argue that the trier should find for the plaintiff. The Pardo/Allen theory requires that the plaintiff's more probable explanation prevail.²⁰⁶

Those arguments expose at least two different flaws in the Pardo/Allen theory's position of the burden of persuasion and sufficiency of the evidence. The first flaw is the conflict between Pardo and Allen's quantification of the value of evidence and their criticisms of Bayesian models that yield quantitative probability assignments for inferences at trial.²⁰⁷ In their work on the Pardo/Allen theory and elsewhere they have argued that that fact finders' ultimate decisions cannot be a Bayesian personalist probability.²⁰⁸ For instance, an interpretation of Bayesian personalist theory to fit the burden of persuasion test²⁰⁹ is not convincing because subjective probability assumes that we can make calculations in our minds which are computationally intractable.²¹⁰ Accordingly, it is extremely unlikely that there are mental states that correspond to their .4 or .1 probability assignments.²¹¹

²⁰⁶ Pardo & Allen, *supra* note 3, at 256–57. After that point, Pardo and Allen argue that inference to the best explanation gives a better account of probative value than does a Bayesian model, relying on their analysis of an illustration from Kahnemann and Tversky. *Id.* at 257–61. It is an interesting point, but does not seem to have any implications for the burden of persuasion that the remainder of their argument does not offer. On Kahnemann and Tversky's research program in general, *see, e.g.*, Todd, *supra* note 120 at, at 52–55.

²⁰⁷ Pardo & Allen, *supra* note 3, at 240 (“Even if the strength of a party's total evidence could be quantified”) & n.49 (citing Ronald J. Allen & Michael Pardo, *The Problematic Value of Mathematical Models of Evidence*, 31 J. LEG. STUD. 36 (2007) for reasons why it cannot), 247–48 & n.70; Allen, *Nature*, *supra* note 126, at 376.

²⁰⁸ Allen, *supra* note 126, at 379–82.

²⁰⁹ At least insofar as we are concerned with fact-finder's actions in real time, that is in actual decision-making, rather than abstract hindsight. Real time decision-making seems to be the only issue at hand here.

²¹⁰ Allen, *supra* note 126, at 380 n.25 notes Professor Cherniak's dramatic example. Testing 138 logically independent propositions by the truth table method would take the fastest serial computer imaginable longer than the “big bang” theory believes to be the life of the universe to the present. *See* CHERNIAK, *supra* note 136, at 93–94 & n.13.

²¹¹ Professor Nance questions the strength of the computational complexity argument against Bayesianism, that is, the argument that comprehensive Bayesian models would require processing capacity that humans do not have. He notes that physicists use complex models of physical systems, systems that do not have the intelligence to do the calculations incorporated in the model. Dale A. Nance, *Naturalized Epistemology and the Critique of Evidence Theory*, 87 VA. L. REV. 1551, 1600–02 (2001).

But Bayesian models of fact finding apply to mental states. Given the computational demands of simply assuring that beliefs are logically consistent, there is no reason to believe those states are analogous to numerical probabilities, at least in any way that would be useful for a systematic prescription of deliberations in a practical amount of time. Claims of more abstract utility should be viewed with caution. Professor Nance's attempt to analogize human thought to physical systems seems to break down.

Similarly, all the conjunction paradox²¹² shows is that Bayesian models do not fit very well with human application of multi-element rules. That simply does not require formulation of a new comparative model of inference. In other words, while inference to the best explanation is very useful for understanding fact finding, there is nothing in the conjunction paradox that requires us to stop appraising the adequacy of evidence before we act, even if the theory on which we might act is the best we can come up with at the moment. The conjunction paradox suggests rather strongly that comprehensive Bayesian models of fact finding do not work well, but that is all.

Suppose we set aside the numbers in Pardo and Allen's thought experiment in which the probability of plaintiff's case is .4.²¹³ That would convert the experiment into one in which the fact finder would regard the plaintiff's explanation as more likely (or a great deal more likely) than the theory of the party's opponent, but also think of the fact finder's uncertainty as equally significant—plaintiff's case and explanations of an unknown nature each having a probability of .4 in the mathematical model.²¹⁴ The Pardo/Allen theory seems to implicitly assume that the party with the better story is more likely to be right than its opponent, and so should win the verdict, despite the uncertainty.

The Pardo/Allen theory, then, is clearly in conflict with research on human cognition and abduction that suggest that decision makers may validly adhere to default rules (such as finding in favor of the party that does not bear the burden of persuasion) when evidence is too thin to justify a different decision. Professors Pardo and Allen have referred to “distributing errors evenly among the parties”²¹⁵ as a norm in their favor. I understand²¹⁶ that a premise of that argument is, essentially, the assumption that the parties' arguments are representative, not only of explanations they have conceived, but of those that they might conceive given enough time, wisdom, and evidence. Then, the argument contends, the party whose story is most likely is the one who deserves to win.²¹⁷ On that basis, the Pardo/Allen theory requires fact finders to decide in accord with the best explanation, ignoring gaps in the information, even those that make jurors doubt the wisdom of

²¹² See Pardo & Allen, *supra* note 3, at 253–56.

²¹³ See Pardo & Allen, *supra* note 3, at 256–57.

²¹⁴ The remaining .4 in Pardo and Allen's example if the parties' explanations have a likelihood of .4 for the plaintiff, and .1 for each of defendant's two explanations. Under the Pardo/Allen theory, the likelihood of the two defenses should be aggregated, given that they are alternatives. Pardo & Allen, *supra* note 3, at 256.

²¹⁵ Pardo & Allen, *supra* note 3, at 254 & n.100, 261–62; Ronald J. Allen, *Factual Ambiguity and a Theory of Evidence*, 88 N.W. U. L. REV. 604, 614 (1994).

²¹⁶ With help from a conversation with Professor Pardo.

²¹⁷ Pardo & Allen, *supra* note 3, at 254 n.100.

doing so. That will often be the case. That is, parties' incentives will often be sufficient to cause them to gather and offer enough evidence to produce a result that is accurate—at least insofar as observers of litigation could be expected to agree that a particular result was accurate.

Even so, there is no reason to believe that the parties' presentations and the jurors' imagination *must necessarily* yield sufficient evidence or explanations to be representative of the information they would have if the substantial uncertainty were resolved. Parties have incentives to gather information and formulate new hypotheses to the extent that their benefits for the parties (primarily improvement of the likelihood of an outcome in their favor) warrant the costs in resources. After they have done so, a large gap in information may still remain. Jurors may sometimes be able to close that gap with inferences, as in the *St. Mary's* case, or in situations in which they formulate an explanation that fills the gap. Their resources are constrained, and the evidence is a further constraint, but the jury may sometimes be able to evolve a new theory subject to those constraints, without wasting resources. Where the gap in information is a gap that reasonable inferences cannot close, the court should rule against the burden bearing party on a motion challenging the sufficiency of the non-moving party's evidence. When it is a gap over which the fact finders are unwilling to make a leap of faith, they should not come to a verdict in favor of the party with the burden.

V. AN ALTERNATIVE

The primary goal of this paper has been to suggest that any theory of the preponderance standard based on inference to the best explanation should take account of empirical research on problem-solving and fact finding, and, in that light, that the preponderance standard cannot be wholly comparative. Even so, the discussions of empirical research and inference to the best explanation in this article provide some underpinnings for an alternative to the Pardo/Allen theory's conception of the preponderance standard.²¹⁸ It seems appropriate then, in closing, to sketch that alternative, at least in a preliminary fashion.

Research on human problem-solving—including the story model—suggests that the preponderance standard for proof of element *E* requires that the fact finder conclude that (i) *E* is more likely than not-*E*, given the evidence in the record and the stories or explanations the fact finder is considering and (ii) the possible disutility of finding *E* in the absence of

²¹⁸ For related analyses, see Craig R. Callen, *Rationality and Relevancy: Conditional Relevancy and Constrained Resources*, 2003 MICH. ST. L. REV. 1243; Craig R. Callen, *Simpson, Fuhrman, Grice and Character Evidence*, 67 U. COLO. L. REV. 777 (1996); Craig R. Callen, *Hearsay and Informal Reasoning*, 47 VAND. L. REV. 43 (1994).

further information or explanations does not warrant reliance on the default rule by finding for the party that does not have the burden²¹⁹ are two points about this alternative that seem particularly worth making at this point.

A. *Optimization Is Not An Option*

Admission and evaluation of evidence does have costs and benefits, and the proposed explanation of the preponderance standard does refer to the disutility of entering a verdict in the absence of further information. It does not follow that courts or jurors can identify “the point at which marginal cost and marginal benefits are equated,”²²⁰ and so apply part (ii) of the preponderance test algorithmically. Nor will jurors know precisely whether the gravity of possible errors warrants reliance on the default rule. It is unlikely they, or we, can optimize in that fashion.

Certainly, benefits and costs are factors in the process of gathering and assessing evidence. Assume that the correct verdict creates a social benefit greater than any other response to the problem. Data may help one to gain that benefit, but not without costs. Accumulation and evaluation of data requires expenditure of material resources that the decision maker²²¹ might devote to other purposes, and also the expenditure of cognitive resources that the decision maker might use for other purposes, such as resolving other issues in the same dispute, resolving other disputes, or performing some other function that may have a greater social utility.²²² In addition, some items of evidence that simply confuse the fact-finder may have negative utility.²²³

In order to calculate the optimal amount (and nature) of evidence that one should gather for a decision, one would need to know the values and costs of gathering and evaluating individual items of evidence across a wide variety of disputes and factual contexts. One would also have to calculate the value of the opportunities lost if resources are devoted to evaluation of that evidence. Considering all evidence that might bear on any one problem²²⁴ (or trying to satisfy the prerequisites of Bayesian models) would require a superhuman effort. Then calculating the costs and benefits of all sorts of evidence, in all sorts of contexts, would require a dramatically greater exertion.²²⁵

²¹⁹ Craig R. Callen, *Kicking Rocks with Dr. Johnson: A Comment on Professor Allen's Theory*, 13 CARDOZO L. REV. 423, 431 (1991) (footnote omitted).

²²⁰ POSNER, *supra* note 35, at 339.

²²¹ And parties, to the extent the procedural system is adversarial. POSNER, *supra* note 35, at 346–48.

²²² See Gigerenzer & Todd, *supra* note 166, at 3, 11.

²²³ POSNER, *supra* note 35, at 386–89.

²²⁴ Callen, *Conditional Relevancy*, *supra* note 24, at 1260–64.

²²⁵ Gigerenzer & Todd, *supra* note 166, at 11. One might argue that the decision maker

The theory of bounded rationality, which “takes into account the cognitive limitations of the decision maker—limitations of both knowledge and computational capacity” is of critical importance.²²⁶ That theory helps to explain how decision makers can be successful despite the impossibility of optimizing under constrained material and cognitive resources. In order to make good decisions in light of their bounded rationality, humans need to recognize familiar situations and crucial data, search for new data in ways that reflect both benefits and costs of seeking data, and employ simple decision rules to help them make decisions based on the critical data.²²⁷ In other words, they use decision-making strategies, such as the story model.

Of course, reliance on cognitive strategies is not a guarantee of infallibility. Decision-makers may select the wrong strategy for a particular problem, misapply the correct strategy, or simply fail to understand the problem *ab initio*. Or one might fail to absorb the lessons of prior experience correctly, and so formulate a number of strategies that are ineffective.

In that light, advocates of approaches based on optimizing could argue that their theories, at the least, establish aspiration levels, ideals at which we should aim, even though we will never achieve them. Certainly, other things being equal, there is no reason why we should not try to improve our decisions and gain wisdom in the process. Yet, studies comparing decisions made with simple heuristics and those that follow multiple regression (a proxy for optimization) show that the heuristics may be more successful, at only a fraction of the cost of regression.²²⁸

Jurors (and probably judges) will tend to believe that the parties have put forward the best cases for their positions.²²⁹ On the other hand, jurors’ or judges’ experience may suggest that additional evidence should be available if one party’s position is accurate. Experience may even teach that the uncertainty remaining after the parties have presented their explanations

need only estimate the costs and benefits, but that argument fares no better. *See* SIMON, *supra* note 130, at 296 (noting that the difficulty of “[s]olving these estimation problems” may be as great, or greater, than the difficulty of trying to consider all data bearing on the matter in question).

²²⁶ SIMON, *supra* note 130, at 291.

²²⁷ Gigerenzer & Todd, *supra* note 166, at 12.

²²⁸ Gerd Gigerenzer, et al., *How Good Are Fast and Frugal Heuristics*, in DECISION SCIENCE AND TECHNOLOGY 81, 90–92 (James Shanteau, et al. eds., 1999) conclude that simple heuristics are superior to multiple regression when the experiments require the strategies to “learn” the value of data from one set of problems, and use what they have “learned” to solve a different, although similar, set of problems. Bayesian networks are much more computationally complex than simple heuristics or multiple regression, yet perform only slightly better than a simple heuristic when asked to “learn” values of data from one set of problems and employ those values to resolve a similar set. *Id.* at 98–100.

²²⁹ *See* ROBERT J. KLONOFF & PAUL L. COLBY, SPONSORSHIP STRATEGY: EVIDENTIARY TACTICS FOR WINNING JURY TRIALS 22 (1990) (jurors will tend to believe parties introduce their best evidence, and will not introduce evidence that is not important).

(and jurors have tried to formulate some on their own initiative) requires a decision in favor of the default. Even so, the limitations of jurors' experience and cognitive resources will generally not permit them to calculate (i) the optimal extent to which the parties should produce evidence, or (ii) the precise extent to which they should view under-producing parties' claims with skepticism. There is no argument here that jurors who rely on their own cognitive strategies will be infallible, just that judges and jurors will make better decisions if permitted to consider whether the evidence is sufficiently complete to justify a departure from the default rule.

B. *Employing 'Story' or 'Explanation' in Preponderance Instructions*

Instructing jurors in terms of explanations or stories, as Pardo and Allen suggest, would be unlikely to make their task any easier. The Pardo/Allen theory holds that the preponderance test requires that the jurors "select the best explanation of the evidence (or the most plausible version of the litigated events)," and that jurors should regard something as "proven by a preponderance of the evidence if it is part of their selected explanation."²³⁰

It seems quite likely that jurors do use the story model, or something very much like it, to understand and assess evidence at trial. Accordingly, telling them to reason in terms of stories seems unlikely to accomplish much if jurors understand it as intended. It might even confuse them, since we tend to assume that people communicating with us are striving to provide information we do not already have.²³¹ In that light, jurors might believe that they are supposed to formulate a form of story that they otherwise might not, or that the court was doing more than simply reassuring them that their normal decision-making methods would suffice. Moreover, asking jurors to select the best explanation of the evidence as a basis for their verdict would seem to cry out for further instructions. In other words, while jurors are likely to engage in some form of inference to the best explanation, they are unlikely to think of it as such. A theoretical restatement of that process in instructions might create confusion for some jurors that would outweigh any improvement in other jurors' decision making.

²³⁰ Pardo & Allen, *supra* note 3, at 266–67.

²³¹ See DAN SPERBER & DEIRDRE WILSON, RELEVANCE: COMMUNICATION AND COGNITION 149–50 (2d ed. 1995); Callen, *Conditional Relevancy*, *supra* note 24, at 1284.

CONCLUSION

This draft has focused on what seem to be flaws in the Pardo/Allen theory. In so doing, it has devoted relatively little attention to the Pardo/Allen theory's utility. The authors' achievement in developing a descriptive theory of inference at trial grounded in a philosophical account of induction is particularly noteworthy. Jurors may often behave as the theory predicts, but the arguments for requiring them to do so seem to fall short.

The primary purpose of this paper has been to set out the reasons why their comparative standard of proof in civil cases does not seem to be an adequate description or an appropriate norm for resolution of issues of fact. Empirical research on human decision making suggests that we take the adequacy of the available evidence in deciding whether to apply a default rule. Some analyses of inference to the best explanation agree. The comparative approach is inconsistent with current jurisprudence on summary judgment and judgment as a matter of law. In fact, the Pardo/Allen theory does not square with the decision-making processes that *Celotex* and *Reeves* require. It seems also to underestimate the social and cognitive utility of burdens of persuasion and production.

One of the recurring arguments in the theory of evidence is whether "ought" implies "can."²³² In other words, the issue is whether it makes sense to set up standards, such as Bayesian models, when research suggests that humans cannot adhere to them, and that striving to do so could well be counter-productive. Inference to the best explanation is an "ought" that reflects important aspects of the "can" that the story model and research on problem-solving suggest. It may be possible for us to set aside arguments that jurors must only compare stories in coming to a verdict—that they should not otherwise ask themselves whether evidence and explanations are adequate to support a verdict in favor of the party bearing the burden of persuasion. If we can do so, and implement inference to the best explanation in conjunction with the story model and research on human decision making, we will be well on our way to developing a model of inference at trial that is better than anything currently on offer. And we might even settle the meaning of "preponderance of the evidence."

²³² See Mike Redmayne, *Rationality, Naturalism and Evidence Law*, 2003 MICH. ST. L. REV. 849, 860–66; Ronald J. Allen & Brian Leiter, *Naturalized Epistemology and the Law of Evidence*, 87 VA. L. REV. 1491, 1554–55 (2001).