

UCLA

UCLA Previously Published Works

Title

Mistrial by Likelihood Ratio: Bayesian Analysis Meets the F-Word

Permalink

<https://escholarship.org/uc/item/2pv9927b>

Journal

Cardozo Law Review, 13

Authors

Bergman, Paul

Moore, Al

Publication Date

1991-11-01

Peer reviewed

MISTRIAL BY LIKELIHOOD RATIO: BAYESIAN ANALYSIS MEETS THE F-WORD

*Rabbi Paul Bergman and the Reverend Al Moore**

INTRODUCTION

Maria and David, two intrepid skydivers, are about to leap from the belly of an aircraft. David suddenly snaps his fingers, removes his parachute, and prepares to jump.

"Hey!" yells Maria. "What are you doing? You can't jump without a parachute—you'll be killed!"

"Nothing to worry about," David replies. "Didn't you read the paper this morning? A new study showed that people's life expectancies are longer than ever."

Inferential errors are not always as potentially lethal as David's. However, the thought that the Davids of this world sometimes wind up as jurors rather than skydivers has led many theorists to look for methods of minimizing what they regard as triers' tendencies to make inferential errors during the fact-finding process. Some of these theorists have suggested that we might be better off replacing the Davids with clones of the Reverend Thomas Bayes, who just over two centuries ago developed the mathematical formula known as Bayes' Theorem.¹

Over a period of approximately two decades, scholars have debated the wisdom of analyzing evidence in actual trials according to

* Please understand that we are neither reverend nor rabbi; we are plain Professors of Law at the UCLA School of Law. But in preparing to do battle with the followers of the Reverend Thomas Bayes, we were reminded of Clarence Darrow, who in the 1920s defended the teacher charged with unlawfully teaching the theory of evolution in the famous *Scopes* trial in Tennessee. *Scopes v. State*, 154 Tenn. 105, 289 S.W. 363 (1927). William Jennings Bryan, who acted as one of the prosecutors in the case, was addressed during the trial by the honorary title of "Colonel." Darrow, seeking equal stature before the jury, insisted on and was given the honorary title of "Colonel" for the duration of the trial. Hence, our good-natured use of the religious titles "Rabbi" and "the Reverend" is meant to recall this colorful bit of legal history, and to put both ourselves and our readers in the spirit of the Reverend Bayes, who can be largely held to account for this conference.

Without in any way suggesting that they agree with our thesis, we wish to thank Professors Ken Graham, Rick Sander, Judea Pearl, and David Kaye, from whose comments we have greatly benefitted. We also are grateful to the UCLA Summer Stipend Program, which helped fund our research; we hope it was money well spent.

¹ Bayes, *An Essay Toward Solving a Problem in the Doctrine of Chances*, in 53 *PHILOSOPHICAL TRANSACTIONS, GIVING SOME ACCOUNT OF THE PRESENT UNDERTAKINGS, STUDIES, AND LABOURS, OF THE INGENIOUS, IN MANY PARTS OF THE WORLD* 370 (1763).

Bayesian methodology.² Most scholars apparently agree that few, if any, triers of fact, be they judges or jurors, presently adhere to Bayesian precepts when resolving factual disputes.³ Thus, proponents of Bayesian analysis argue that theirs is a normative theory of how triers of fact *should* behave.⁴ As one adherent puts it simply, "the Bayesian equation describes the way the law's ideal juror evaluates new items of evidence."⁵

Nevertheless, most Bayesians⁶ disavow the use of Bayesian methodology by triers in actual trials.⁷ For example, David Schum has written, "I have never seen or heard advocates of any of these [mathematical] systems argue in favor of the overt incorporation and combination of numbers on the part of factfinders"⁸ Similarly, Ronald Allen has noted that it is "becoming increasingly obvious" that Bayesian analysis is best understood as a guide to rational thought, and not as a tool to be used in actual trials.⁹

The reluctance of Bayesians to advocate the use of their methodology in actual trials may be due to the force of a variety of practical criticisms. Bayesians recognize that even people who have extensive statistical skills (which of course excludes most judges and jurors) are likely to have extreme difficulty carrying out a Bayesian analysis.¹⁰ Moreover, most people would be uncomfortable attaching any kind of numerical labels to their degrees of uncertainty.¹¹ Finally, any at-

² One possible starting point of the debate over the utility of Bayesian methodology in trials is the article Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065 (1968). For many participants, however, the debate started, or at least accelerated rapidly, with the publication of Finkelstein & Fairley, *A Bayesian Approach to Identification Evidence*, 83 HARV. L. REV. 489 (1970), and the riposte which attempted to quell (but which in fact stoked) the fires of mathematical analysis, Tribe, *Trial By Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971).

³ See, e.g., Fienberg & Schervish, *The Relevance of Bayesian Inference for the Presentation of Statistical Evidence and for Legal Decisionmaking*, 66 B.U.L. REV. 771, 773 (1986); Edwards, *Comment*, 66 B.U.L. REV. 877, 880 (1986); Tribe, *supra* note 2, at 1358.

⁴ See Fienberg & Schervish, *supra* note 3, at 772-73.

⁵ Lempert, *Modeling Relevance*, 75 MICH. L. REV. 1021, 1023 (1977).

⁶ We recognize that the term *Bayesians* lumps together scholars whose views even about Bayesian methodology are often quite diverse. We do not use the term pejoratively, but to emphasize the common threads in Bayesian methodology.

⁷ A few Bayesian advocates do, however, promote the use of Bayes' Theorem in actual trials, even going so far as to suggest that judges and jurors be given instruction in formal probability theory. See, e.g., Fienberg & Schervish, *supra* note 3, at 795.

⁸ Schum, *Probability and Processes of Discovery, Proof, and Choice*, 66 B.U.L. REV. 825, 875 (1986).

⁹ Allen, *A Reconceptualization of Civil Trials*, 66 B.U.L. REV. 401, 402 (1986).

¹⁰ Fienberg & Schervish, *supra* note 3, at 773.

¹¹ Edwards, *supra* note 3, at 880; Cf. Tribe, *supra* note 2, at 1358. Even if jurors could convert their understanding of the evidence into numbers for Bayes' Theorem, if the formula produced a result contrary to what the jurors believed was correct, jurors would learn to "re-

tempt to institute Bayesian methodology would have to overcome a variety of policy concerns, such as Laurence Tribe's arguments that mathematical formulas give undue credence to "hard" numbers as compared to "soft variables," and that Bayesian analysis is inconsistent with the presumption of innocence.¹² Demurring to these and other "practical" criticisms, most Bayesians stop short of calling for triers to apply Bayes' Theorem in actual trials, even as they continue to insist that Bayesian analysis remains the theoretical ideal.¹³

In this essay, we will assume for purposes of analysis that the Bayesians could triumph over all these very cogent "practical" criticisms. The question then remaining would be whether the Bayesians occupy the theoretical high ground. Would ideal triers of fact be Bayesians? Our answer is no. Application of substantive legal principles relies on, and due process considerations perhaps require, that triers must make individualistic judgments¹⁴ about how they think a particular event (or series of events) occurred.¹⁵ In order to contest for normative-ideal honors, Bayesian methodology would have to produce individualistic judgments about historical events. This article contends that Bayesian methodology would not produce such judgments. Instead of determining what they think happened in a particular case, triers following Bayesian precepts would produce a

port the subprobabilities [that is, the numbers required by Bayes' Theorem] so that the ultimate verdict would be the one they had decided upon." Kaplan, *supra* note 2, at 1091.

¹² Tribe, *supra* note 2, at 1361-72. Recent empirical studies have, however, tended to indicate that at least in simulated jury deliberations, jurors do not give undue credence to "hard numbers." See, e.g., Faigman & Baglioni, *Bayes Theorem in the Trial Process—Instructing Jurors on the Value of Statistical Evidence*, 12 LAW & HUM. BEHAV. 1, 16 (1988).

¹³ Interestingly, at least some statisticians assume that while Bayesianism might not be as useful as other statistical methods for performing "objective" scientific tasks, Bayesianism is useful in legal settings:

Bayesian theory concentrates on inference The dominant Bayesian school . . . is the subjective Bayesianism of de Finetti and Savage. Now by definition one cannot argue with a subjectivist, so I will just state the obvious fact: though subjectivism is undoubtedly useful in situations involving personal decision making, for example, . . . legal decisions, it has failed to make much of a dent in scientific statistical practice.

Efron, *Why Isn't Everyone a Bayesian?*, AM. STATISTICIAN, Feb. 1986, at 1, 2-3. "Subjective Bayesianism" is commonly referred to as the "Ramsey-deFinetti-Savage theory of subjective probability." Fienberg & Schervish, *supra* note 3, at 774.

¹⁴ For a discussion of the reasons that individualistic judgments are inherent in the common-law adversary system, see Thompson, *Decision, Disciplined Inferences and The Adversary Process*, 13 CARDOZO L. REV. 725 (1991).

¹⁵ Implicit in such legal standards as preponderance of the evidence and beyond a reasonable doubt is the realization that jurors' individualistic judgments do not necessarily reflect "objective truth." Jurors' beliefs about what really happened are always influenced by cultural, economic, and political biases. Our point is simply that the Anglo-American system of justice is premised on jurors arriving at verdicts after determining how they believe particular events unfolded.

frequentist decision. That is, they would determine what is most likely to happen in "cases such as these." As Bayesian methodology does not produce individualistic judgments, it fails as a theoretical fact-finding model.¹⁶

I. FREQUENTIST ASSESSMENTS VS. BELIEVABILITY JUDGMENTS

To demonstrate the shortcomings of Bayesian methodology as a model of fact finding at trial, we begin by describing three ways of thinking about how past events might have occurred: (i) objective frequentist probabilities, (ii) subjective frequency judgments, and (iii) believability judgments.

An *objective frequentist probability* is "an objective fact about a repeatable event; it is the long-run frequency with which the event happens."¹⁷ For example, the objective frequentist probability that a fair coin flipped three times will come up heads each time is one-eighth.

However, an assessment of the long-run frequency with which events happen need not always be objective. People typically assess long-run frequencies according to their subjective beliefs. *Subjective frequency judgments* grow out of unproven (and often unprovable) generalizations about the long-run behavior of people and objects.¹⁸ For example, you might have the subjective belief that people who flee the scene of a crime are usually guilty of committing the crime, or that drivers almost always stop at stop signs.¹⁹ Although these assessments are subjective, they are nonetheless frequentist. They are estimates of the frequency with which events happen when a class of cases is considered.²⁰

¹⁶ Bayesian analysis is, however, relevant to certain trial tasks. See *infra* pp. 614-15.

¹⁷ Shafer, *The Construction of Probability Arguments*, 66 B.U.L. REV. 799, 801 (1986).

¹⁸ See, e.g., 1A J.H. WIGMORE, EVIDENCE IN TRIALS AT COMMON LAW § 37.4, at 1033 (Tillers rev. 1983) [hereinafter WIGMORE ON EVIDENCE]; J.H. WIGMORE, THE SCIENCE OF JUDICIAL PROOF (3d ed. 1937); D.A. BINDER & P. BERGMAN, FACT INVESTIGATION: FROM HYPOTHESIS TO PROOF (1984).

For an interesting discussion of how a trier might use generalizations about the health of military commanders to make frequency assessments concerning Napoleon's battle strategy in the Battle of Waterloo, see W. TWINING, *Some Skepticism About Some Skepticism*, in RE-THINKING EVIDENCE: EXPLORATORY ESSAYS 92, 121 (1990).

¹⁹ Usually, as in the examples above, subjective beliefs grow out of generalizations drawn from everyday societal experience. However, generalizations might also be based on scientific laws (e.g., the law of gravity) as well as prejudices not founded on experience (e.g., that skydivers are more likely to be compulsive gamblers than other people, when one has no experience with skydivers).

²⁰ Unlike objective probabilities, people do not often state their subjective assessments numerically. However, if pushed to do so, they often would. For example, one might say that drivers stop at stop signs 95% of the time.

By contrast, a *believability judgment* is an assessment about how a particular series of events unfolded. For example, to make a believability judgment about whether a specific person running away from the scene of a crime was (or was not) guilty, or about whether a specific driver did (or did not) stop at a stop sign, you would do more than subjectively assess the likelihood that people fleeing the scenes of crimes are guilty, or that drivers do not stop at stop signs. In addition, you would consider whatever other evidence you thought relevant, and you would decide what you thought really happened in the specific case.²¹ Thus, you might decide that though your subjective frequency assessment is that most people who run away from the scene of a crime do so because they are guilty, the defendant in the particular case before you did so because the defendant feared for his safety.

The following example illustrates how a trier of fact might employ these three ways of thinking about how past events might have occurred. Assume that a nine-year-old child, Sherry, tells her parent, "I flipped a coin thrice and each time it came up heads." Sherry's seven-year-old brother, Orin, counters by saying, "It did not. It came up tails once." In the face of such a dispute, a trier's reasoning might go as follows:

(a) Adhering to an objective frequentist notion of probability, the trier might assess the probability that Sherry's story is accurate as one in eight.²²

²¹ Objective and subjective frequencies are often important to a trier's judgment of believability. Another way of saying this is that objective and subjective frequencies are circumstantial evidence of believability. Our point is not that frequency assessments are irrelevant, but that they are not sufficient for believability.

Note that in the absence of evidence that would allow a trier to make a believability judgment, objective frequency assessments standing alone are almost never sufficient to sustain a claim in court. For example, consider L. Jonathan Cohen's "rodeo gatecrasher" hypothetical. L.J. COHEN, *THE PROBABLE AND THE PROVABLE* 75 (1977). In this hypothetical, 1,000 people attend a rodeo. The rodeo operator brings suit against *A*, claiming that *A* did not pay the price of admission. The rodeo operator's only proof consists of evidence that *A* attended the rodeo, and that 501 of the 1,000 people who attended the rodeo did not pay the admission price. No evidence is offered as to whether *A* personally did or did not pay. If a trier were to make a frequency assessment, there is a probability of over .5 that *A* did not pay. Yet most commentators, Bayesian and non-Bayesian alike, agree that under these circumstances *A* should not be held liable. See, e.g., Fienberg & Schervish, *supra* note 3; Tribe, *supra* note 2; Lempert, *supra* note 5. The reason, we contend, that *A* should not be held liable, is the lack of evidence which would allow a trier to make a believability judgment. On very rare occasions, however, when a court is convinced that insisting on evidence on which a believability judgment could be made would work a serious injustice, verdicts based only on frequency assessments may be allowed. See *Kaminsky v. Hertz Corp.*, 95 Mich. App. 356, 288 N.W.2d 426 (Ct. App. 1979).

²² Though few actual cases revolve around coin flips, objective frequentist probabilities are often quite relevant in trials. Objective frequentist statistics underlie many forms of scientific

(b) Employing a subjective frequentist notion of probability, the trier might estimate how likely children are to be accurate when they make claims about the outcome of coin flips to their parents.²³ For example, the trier might estimate that in this class of case, juvenile coin flippers are likely to be accurate only 30% of the time, while nonflippers are likely to be accurate 70% of the time, and thus the trier might conclude that Orin is more likely to be a truth teller.

(c) Arriving at a believability judgment, the trier would have to consider whatever evidence it thinks relevant to determine whether, in this particular instance, Sherry's or Orin's statement is more likely to be accurate. For example, assume there was evidence that (1) Sherry had no reason to lie; (2) Sherry's manner of speech strangely resembles that of Elizabethans; and (3) Sherry and Orin had just had a big argument over who could eat the last cookie, with Sherry winning. Considering these items of evidence, as well as factors that typically relate to the credibility of the witnesses' competing stories (e.g., completeness, internal consistency, consistency with established facts),²⁴ the trier might conclude that Sherry is more likely to have told the truth on this occasion. This believability judgment may well be influenced by, but would not necessarily be the same as, the one in eight objective probability that a coin will come up heads three times in a row, and the trier's subjective frequentist assessment of how often juvenile flippers and nonflippers in this class of case are likely to be accurate when they make statements to their parents.

The above distinctions are vital, because triers are required to

evidence routinely admitted in trials. *See generally*, D. BARNES, STATISTICS AS PROOF (1983). For instance, in a paternity suit, an expert might testify that based upon the phenotype results of a blood test, the expert has concluded that the defendant is the father of a child. Underlying this testimony is the fact that large numbers of individuals from different populations have been tested, and the objective frequencies of different phenotypes have been constructed. *See Shafer, supra* note 17, at 810. Similarly, one factor in the current controversy over the admissibility of DNA evidence is the issue of whether population sample sizes are sufficient to allow meaningful objective comparisons between an individual's DNA pattern and the DNA patterns of some segment of the relevant population. *See Hoeffel, The Dark Side of DNA Profiling: Unreliable Scientific Evidence Meets the Criminal Defendant*, 42 STAN. L. REV. 465, 488-92 (1990).

²³ For purposes of illustrating subjective frequencies, we have defined the referent class as statements made by seven- and nine-year-old children about coin flips to their parents. But the class could be defined either more abstractly (e.g., statements made by young children to their parents) or more concretely (e.g., statements about coin flips made to their parents by seven- and nine-year-old children who have just had an argument over which child gets to eat the last cookie). Depending on how the class is defined, the probability estimate may vary. Our point is simply that to make a probability judgment based on subjective frequencies, the trier would have to treat a particular case as a member of a more abstract class.

²⁴ *See Pennington & Hastie, Cognitive Theory of Juror Decisionmaking: The Story Model*, 13 CARDOZO L. REV. 519 (1991).

make believability judgments at trial. Legal rights flow from triers' beliefs about what really happened.²⁵ Assume that a trier returned a verdict stating, "I find that in 80% of the cases similar to the one before me, defendants would be liable. 20% of the time, they would not be. However, I have not determined whether this case falls in the 80% group or the 20% group." Such a verdict would be inadequate.²⁶ Due process of law would require the trier to determine into which group the defendant falls *in this very case*. Without the believability judgment, there is no factual assessment to which legal principles can apply.²⁷

Few Bayesians would argue that outcomes at trial must satisfy objective frequentist notions of probability.²⁸ Instead, the Bayesian argument, as we understand it, is that trier use of Bayesian methodology would produce what we term believability judgments.²⁹ That is,

²⁵ "Successful projection of a legal rule depends on a court's ability to cast a verdict not as a statement about the evidence presented at trial, but as a statement about a past act—a statement about what happened." Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357, 1358 (1985). See also D.A. BINDER & P. BERGMAN, *supra* note 18, at 10 ("[A] court can apply the substantive law to determine the rights and obligations of the parties *only after* 'first it finds the facts. . .'" (quoting J. FRANK, COURTS ON TRIAL 3 (Athenum ed. 1970); A.A.S. ZUCKERMAN, THE PRINCIPLES OF CRIMINAL EVIDENCE 21-22 (1989)).

²⁶ Note that a determination that "in 80% of the cases like the one before me, defendants would be liable" is not the equivalent of a juror's statement, "I have determined what happened in this very case, and based upon that I find the defendant liable. I am not completely certain, however, that my determination of what happened is correct; my confidence level that my determination of what happened is accurate is only 80%." The latter statement reflects a judgment about what happened in a particular case; the former does not.

²⁷ Cf. Mosteller, *Legal Doctrines Governing the Admissibility of Expert Testimony Concerning Social Framework Evidence*, LAW & CONTEMP. PROBS., Autumn 1989, at 85.

Social framework testimony relevant to the credibility of an alleged child victim, for example, does not enable the jurors to determine whether an individual is telling the truth in that it does not generally give the jury guidance on how to decide whether a *particular* child is telling the truth. Rather, the framework testimony tells jurors only that children who make such allegations are generally highly credible.

Id. at 90-91.

²⁸ Cf. W. TWINING, *supra* note 18, at 122 ("[S]ome, but by no means all Pascalians, by setting very strict standards for rationality which are rarely attainable in practice . . . are analogous to disappointed absolutists].")

²⁹ Bayesians would undoubtedly substitute the phrase *probability assessment* for our term, *believability judgment*. That is, whereas we might ask a trier, "What do you believe happened?" a Bayesian might ask, "What probably happened?" But if Bayes' Theorem permitted triers to take into account the same evidence they do now in arriving at a determination of what happened, and compelled them to weigh it in a mathematically responsible way, our theoretical objections would largely disappear. Whether a trier's reasoning is termed a probability assessment or a believability judgment would be only a semantic difference. Our point is that Bayesian methodology compels a trier to *ignore* certain evidence and thus produces a subjective frequency assessment.

they may concede that Bayesian methodology requires people to make estimates based on subjective frequencies.³⁰ However, Bayesians would contend that in making a frequency assessment, a trier can take into account all of the particular circumstances of an individual case. For example, in determining the probabilities that Sherry and Orin were accurate, Bayesians would allow the trier to take into account evidence that Sherry had no motive to lie and spoke like an Elizabethan, and that Orin was angry because his sister had gotten the last cookie.³¹ Hence, Bayesians would conclude, their methodology produces an assessment of how particular events probably occurred.³² The Bayesian argument has plausible surface appeal. However, as we will demonstrate, triers could not apply Bayes' Theorem to disputed questions of fact at trial without substantially changing the nature of their inquiry. Bayesian methodology would compel triers to make frequency assessments, not believability judgments.

The next section briefly explains the mechanics of Bayes' Theorem. Ensuing sections explain why those mechanics require triers to ignore some evidence introduced at trial and to make unwarranted inferences from a particular kind of circumstantial evidence, and thus produce frequency assessments. We then conclude that the Reverend Bayes would be no more suitable a juror than David Skydiver.

II. A THUMBNAILED OVERVIEW OF BAYES' THEOREM

To explain the constraints that use of Bayes' Theorem would

³⁰ A number of authors, Bayesians included, have noted that Bayesian methodology is based on frequency assessments. For example, Peter Tillers notes that to apply the Bayesian calculus to the effect of a defendant's escape attempt on a determination of whether the defendant is guilty, "[w]e should estimate the relative frequency with which (we believe) guilty and innocent people escape." 1A WIGMORE ON EVIDENCE, *supra* note 18, § 37.6, at 1059. Similarly, Glenn Shafer remarks that "[m]ost Bayesian statisticians compromise on the meaning of probability; they agree that their goal is to estimate objective probabilities from frequency data, but they advocate using subjective prior probabilities to improve the estimates." G. Shafer, *The Meaning of Probability*, in READINGS IN UNCERTAIN REASONING 9 (1990). Finally, Finkelstein and Fairley, whose article did much to stimulate the inferential debate, state that, "[b]oth subjective probability [i.e., probability assessed by Bayesian methods] and probability as classically defined reflect frequencies of events." Finkelstein & Fairley, *supra* note 2, at 504.

³¹ Note that Bayesian scholars do not, as a rule, make the empirical claim that use of Bayes' Theorem will necessarily enhance the accuracy of fact determination. See Schum, *supra* note 8, at 875.

³² Bayesians might rationally respond to the distinction between subjective frequency and believability judgments by conceding that Bayes' Theorem produces frequentist assessments, and arguing that triers *ought* to make only frequency assessments. Such an argument would rest on an empirical claim that over the long haul, frequency assessments are likely to yield a greater number of accurate verdicts than are believability judgments. While we might resist this claim, we think it sufficient to point out here that as presently constituted, our legal system requires triers to make believability judgments.

place on triers, we begin with a brief explanation of the theorem. We will not impose upon any readers by being the umpteenth writers to set forth the mathematical innards of the theorem.³³ Doing so would imply an understanding of the theorem's complexities that we simply do not possess. Moreover, knowing how to make Bayesian calculations is not necessary for understanding how Bayesian methodology would alter the nature of fact finding. Thus, what follows is our verbal description of realities that are often described in abstract mathematical symbolism.³⁴

Essentially, Bayes' Theorem provides a mechanism for revising one's estimate of probabilities in the light of new information.³⁵ It provides "a quantitative description of the ordinary process of weighing evidence."³⁶ Its use depends on two critical pieces of data. First, one must have a data base, or a *prior probability*—that is, an already-existing estimate of probability that a proposition is true or false. In an experimental setting devoid of disputed versions of historical events, arriving at a prior probability is relatively easy. For example, if one has a sack of one hundred marbles, fifty of which are red, the prior probability is .50 that the first marble removed from the sack will be red.

In actual trials, unfortunately, few cases involve sacks containing known quantities of different colored marbles. Instead, a trier having to make a believability judgment about disputed events confronts what Laurence Tribe has termed "soft variables,"³⁷ such as the credibility of witnesses and the probative value of evidence. Just how or when a Bayesian judge or juror would decide upon the prior probability of a defendant's guilt (in a criminal case) or liability (in a civil case) is unclear. Tribe suggests that in criminal cases, formulation of a prior probability would violate the presumption of innocence.³⁸ Those who believe that the Bayesians would surely have an easy riposte to Tribe might be put off by the Bayesian scholars Ste-

³³ Readers interested in the mathematical underpinnings of Bayes' Theorem might consult, 1A WIGMORE ON EVIDENCE, *supra* note 18, at 1056 n.12; Tribe, *supra* note 2, at 1351-52; J. PEARL, PROBABILISTIC REASONING IN INTELLIGENT SYSTEMS: NETWORKS OF PLAUSIBLE INFERENCE 29-70 (1988); *Probability and Inference in the Law of Evidence*, 66 B.U.L. REV. 377 (1986) (symposium issue).

³⁴ See, e.g., Friedman, *A Diagrammatic Approach to Evidence*, 66 B.U.L. REV. 571, 584-88 (1986); Fienberg & Schervish, *supra* note 3, at 773-76.

³⁵ See Martin, *Comment*, 66 B.U.L. REV. 709, 710 (1986) ("Bayes's Theorem . . . is a prescription for revising opinions about the relative likelihood of some hypothesis based on observable evidence.").

³⁶ I. GOOD, PROBABILITY AND THE WEIGHING OF EVIDENCE 62 (1950).

³⁷ Tribe, *supra* note 2, at 1361.

³⁸ *Id.* at 1368. Some prior probability is essential because, in the Bayesian system, a prior probability of zero can never be adjusted upward.

phen Fienberg and Mark Schervish, who despite being bothered by "the somewhat uncomfortable presumption that each juror has some prior probability of finding guilt," rather remarkably recommend that such a prior probability be based on "general information about the nature of charges, as well as external information such as the percentage of defendants found to be guilty when charged with similar crimes."³⁹ Of course, other possibilities have been suggested.⁴⁰ At some point, however, Bayesian methodology would require a trier to formulate or be instructed about a prior probability that is consistent with constitutional guarantees.⁴¹

Second, a Bayesian revises a prior probability to reflect the impact of new information. In Bayesian terms, the impact is encapsulated in the second critical piece of data, a *likelihood ratio*. To arrive at a likelihood ratio, one determines how likely the new information is on the assumption that a proposition is accurate, and how likely the new information is on the assumption that the proposition is false. For example, a trier might reason that if a person were guilty of an assault, there is a .9 probability that he would flee from the scene of the assault, and that if a person were innocent of an assault there might be a .5 probability that he would flee from the scene. These numbers produce a likelihood ratio, which in this case would indicate the trier's subjective estimate of how much more likely it is that a guilty person will flee the scene of an assault than an innocent person.⁴² Bayes' Theorem multiplies this likelihood ratio by the prior

³⁹ Fienberg & Schervish, *supra* note 3, at 780. See also Fienberg & Kadane, *The Presentation of Bayesian Statistical Analyses in Legal Proceedings*, 32 STATISTICIAN 88, 92 (1983). Apart from noting that the prior probability must be less than the ultimate probability that the trier would associate with guilt, the authors are unclear about just what the trier should use for its prior probability.

⁴⁰ For example, a trier might not be asked to formulate a prior probability until almost all of the evidence is submitted to it. Then, the trier might deliberate, much as it does now, and arrive at a prior probability. This prior probability might then be revised in accordance with Bayes' Theorem based on additional "mathematical" or "statistical" evidence. See, e.g., Faigman & Baglioni, *supra* note 12. However, this suggestion is flawed. As will be discussed, *infra* part V, a juror's decision about what happened in a case reflects a degree of confidence, not a prior probability.

⁴¹ Although we believe that Bayesians would have serious difficulty establishing prior probabilities in many cases, recall that we are assuming for purposes of argument that these difficulties could be overcome.

⁴² These numbers may confuse some readers, since a basic axiom of probability theory is that probability of an event occurring and of its not occurring must sum to one. See, e.g., J. PEARL, *supra* note 33, at 30. As the example in the text indicates, when considering a class of cases (e.g., people fleeing from the scenes of crimes), likelihood ratios need not add up to one, because it may be quite likely that both innocent and guilty people will flee the scene of a crime. When considering only a particular case, however, we believe that it is not possible to obtain a meaningful likelihood ratio. See discussion, *infra* pp. 607-13.

probability, thus allowing one to calculate with seeming mathematical precision the *posterior probability*—that is, the probability that a proposition is accurate after receipt of the new information. For example, assume that before calculating the impact of the evidence of “flight,” a trier has calculated the prior probability to be .50 that the defendant charged with assault was guilty. This prior probability and the likelihood ratio of .9 and .5 would be combined according to the mathematical dictates of Bayes’ Theorem, yielding a posterior probability of the defendant’s guilt. That is, the .50 prior probability of guilt that existed before the flight information was accounted for would be updated with the flight information to a posterior probability. Of course, when numerous pieces of new information must be accounted for, each posterior probability that is generated becomes, in its turn, the prior probability with which one multiplies a new likelihood ratio.

One of the most common illustrations of how likelihood ratios and prior probabilities work together in a Bayesian world involves the calculation of the probability that a blue cab was involved in an accident. In this illustration, one knows only that one hundred cabs operate in the locale where the accident occurred, and of these, eighty-five are green and fifteen are blue. With only this limited knowledge in hand, the “prior probability” would be .85 that the cab involved in the accident was green.

Then, along comes new information, in the form of a witness’ testimony, “The cab involved in the accident was blue.” The trier is told that in experiments under similar conditions, the witness correctly identifies the color of the cab 80% of the time. That is, assuming the cab were blue, the likelihood that the witness would say the cab was blue is .8, and that, assuming the cab were green, the likelihood is .2 that the witness would say it was blue.⁴³ These probabilities become a likelihood ratio, which when combined with the prior probability according to Bayes’ Theorem results in a posterior probability of .41 that the cab was blue.⁴⁴

⁴³ In the actual source from which this example is drawn, the authors arrive at an 80% figure by having the witness view cabs under similar circumstances over and over, and determining that the witness is correct 80% of the time. In trial settings that sort of data base is unavailable, and would be most closely approximated by asking triers to attach a number (here, for consistency’s sake, .80) to the likelihood that the witness’s observation is accurate. In either situation, the Bayesian calculation would remain the same. See J. PEARL, *supra* note 33, at 44-45.

⁴⁴ For an explanation of how Bayes’ Theorem leads to the .41 probability in the blue cab case, see Tversky & Kahneman, *Evidential Impact of the Base Rates*, in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 155-58 (A. Kahneman, D. Slovic & A. Tversky eds. 1982). So far as we can tell, every Bayesian accepts the accuracy of this illustration. However, we believe that the Bayesians’ analysis is flawed. See *infra* pp. 616-18.

Assume that one is in a trial setting in which the plaintiff is suing the Blue Cab Company. The .41 probability that the cab was blue should result in a defense verdict, as plaintiff's case would not have been proved by a preponderance of the evidence.⁴⁵ Yet, in experiments in which hypothetical jurors are given the above information, they routinely convert the .8 likelihood that the witness is correct into an 80% probability that the cab was blue. In a trial setting, then, these jurors (apparently David Skydivers all) would probably return a verdict for the plaintiff. Such seemingly inaccurate outcomes fuel the Bayesian charge that fact finding at trial is untrustworthy.⁴⁶

The blue cab hypothetical exemplifies how Bayes' Theorem seemingly accounts for two types of data which empirical studies indicate triers tend to ignore or undervalue.⁴⁷ First, the prior probability prevents triers from ignoring the already-existing chances (odds) that a proposition is true. Second, the likelihood ratio ensures that a prior probability is adjusted not only according to the probability that a proposition would be correct in the light of a given item of evidence, but also according to the probability that a proposition would be incorrect given that same item of evidence.

Hail Bayes Theorem you might be tempted to say at this point. First, its use seems to compel triers to take relevant data into account that they might otherwise overlook. Second, because a trier can seemingly consider all of the evidence in a case when assigning a likelihood ratio to any item of evidence, its use appears to produce a believability judgment (or what Bayesians would prefer to term a probability judgment) about how a particular series of events unfolded.

However, the conclusion that Bayes' Theorem produces a believability judgment is incorrect. Like magicians who use long sleeves

⁴⁵ Of course, the .41 probability is a posterior probability which in an actual trial could be altered by likelihood ratios pertaining to other items of evidence.

⁴⁶ Demonstrating why Bayesian analysis of the blue cab case is problematic in the context of actual trials takes a bit of doing. See *infra* pp. 616-18. But that something is wrong with the Bayesian analysis seems clear if you consider one bizarre result that flows from it. Assume that a client enters Lawyer Smith's office and tells Smith the tale of having been struck by a cab of indeterminate color. After some investigation, Smith uncovers the statistics about the number of blue and green cabs, as well as a witness who is prepared to testify, "The cab that struck your client was blue." Smith questions the witness and determines that a trier would conclude that the probability that the witness is correct is .8. What entity should Lawyer Smith sue? The Green Cab Company! Educated in Bayesian realities, Smith would call the witness to testify that the cab was blue, offer evidence of the statistics, argue that the witness is credible to the tune .8, and then assert, "Bayes' theorem tells us that the likelihood that the cab was blue is .41. That leaves a .59 probability that the cab that struck my client was green. Pay up." In a case where a sole and credible witness testifies that the cab was blue, Bayes' Theorem mandates a verdict against the Green Cab Company!

⁴⁷ See, e.g., Tversky & Kahneman, *supra* note 44, at 155-58.

and clever patter to fool an audience, Bayesian mathematical calculations conceal the transformation of believability judgments into frequency assessments. As we will demonstrate, Bayesian sleight-of-hand has two components:

1. Bayes' Theorem assigns a fixed weight to a prior probability. For example, in the blue cab case, the prior probability always makes it more likely that the cab will be green. In reality, however, a prior probability is simply a form of circumstantial evidence. Like any other item of circumstantial evidence, the significance of a prior probability depends on the circumstances of a particular case. Indeed, in the blue cab case a trier might rationally infer from the larger number of green cabs that the cab involved in the accident was green.⁴⁸
2. To formulate likelihood ratios, triers would have to ignore the everyday, common sense reasoning processes by which they draw inferences and recall and evaluate the credibility of conflicting stories as a part of determining how a particular series of events unfolded.⁴⁹

Hence, neither aspect of the theorem's mathematical framework responds to the critical question triers must answer: What do you believe really happened in this case?⁵⁰ Use of Bayes' Theorem would produce frequentist assessments, not believability judgments.

III. LIKELIHOOD RATIOS PRODUCE FREQUENTIST ASSESSMENTS OF PROBABILITY

A. *Decision-Making Heuristics*

Understanding why likelihood ratios yield frequentist outcomes begins with the decision-making process by which triers typically resolve factual questions.⁵¹ Triers typically rely on heuristics, which are intuitive, common-sense decision-making strategies. Triers use heuristics to formulate stories from trial testimony.⁵² Stories in turn

⁴⁸ See *infra* pp.617-18.

⁴⁹ See *infra* pp. 601-13.

⁵⁰ Though he criticizes Bayesian methodology on numerous grounds, Tribe appears to accept its ability to produce what we refer to as believability judgments. To this extent we disagree with Tribe. See Tribe, *supra* note 2, at 1346-48.

⁵¹ We assume that the trier is genuinely trying to ascertain what really happened, and is not intending to act on an overt bias such as: Building contractors are evil and deserve to be punished at every opportunity.

⁵² See Moore, *Trial By Schema: Cognitive Filters in the Courtroom*, 37 UCLA L. REV. 273, 284-86 (1989); Pennington & Hastie, *supra* note 24; Saks & Kidd, *Human Information Processing and Adjudication: Trial By Heuristics*, 15 LAW & SOC'Y REV. 123 (1980-81).

allow triers to resolve factual disputes.⁵³ Stories are simply chronological accounts of interrelated events.⁵⁴ At trial, stories are a primary factor in opening statements, direct examinations, and closing arguments.⁵⁵ Not surprisingly, then, story construction is the basis on which triers decide whether a defendant is guilty or innocent, liable or not liable.⁵⁶ As a well-known group of jury researchers put it, "The Story Model has three processing stages. . . . [T]he story construction stage specifies the processing during which evidence is comprehended and organized into one or more plausible accounts describing 'what happened' at the time of events testified to during the trial."⁵⁷

Triers' need to organize evidence into specific stories is understandable. Evidence at trial typically consists of myriads of events and details. Moreover, events often are capable of at least two different interpretations—a plaintiff's version, and a defendant's version. Because of memory limitations, however, triers cannot store all of these potential stories in their heads.⁵⁸ They simplify their cognitive tasks by resolving how events took place, and discarding alternate versions. Of course, a juror might change his or her interpretation of an event after talking to other jurors. Nevertheless, the point is that in the real world a story has unfolded in one and only one way, and triers will naturally be attracted to the interpretation of events that they think best reflects the real-world unfolding. A trier's final story usually can stand only a single interpretation of each event.⁵⁹

Consider the process of story formation in the context of a mundane mishap involving a pedestrian, Potter, being injured when struck by a pickup truck driven by Schrackle. Potter claims that Schrackle's careless driving caused him to strike her. In particular, Potter asserts that Schrackle was preoccupied with business concerns and was late

⁵³ See, e.g., Pennington & Hastie, *supra* note 24; W. BENNETT & M. FELDMAN, *RECONSTRUCTING REALITY IN THE COURTROOM* (1981). Twining summarizes Bennett and Feldman's views as follows: "Above all, (stories) provide essential *frames of reference* for organizing, evaluating and interpreting evidence." W. TWINING, *supra* note 18, at 240 (emphasis in original).

⁵⁴ See D.A. BINDER & P. BERGMAN, *supra* note 18, at 11.

⁵⁵ See T. MAUET, *FUNDAMENTALS OF TRIAL TECHNIQUES* 50, 76, 287 (2d ed. 1988); P. BERGMAN, *TRIAL ADVOCACY IN A NUTSHELL* 63, 234-37, 287 (2d ed. 1989).

⁵⁶ Pennington & Hastie, *supra* note 24.

⁵⁷ R. HASTIE, S. PENROD & N. PENNINGTON, *INSIDE THE JURY* 22 (1983).

⁵⁸ See G. HARMAN, *CHANGE IN VIEW* 25-27 (1986), for the suggestion that to keep track of the possible inferences pertaining to 30 items of evidence, a Bayesian would have to record a billion probabilities. See also, Moore, *supra* note 52.

⁵⁹ Of course, this is not always true. Sometimes the trier may be unable to choose between two versions of events, but both versions result in a verdict in favor of the same party. Alternatively, the trier may be unable to decide how they believe events took place, and therefore may find against the party with the burden of proof.

for a meeting, and so made an unsafe left turn at an excessive rate of speed and struck her in the crosswalk.

Schrackle, not surprisingly, offers a different version of events. He asserts that he was driving carefully, and struck Potter only because she suddenly ran out from between two parked cars about fifty feet away from the crosswalk. The meeting that Schrackle was late for was not a critical one, so Schrackle had no cause to speed to it. Moreover, he had expensive kitchen cabinets in the back of his truck, and was driving extra slowly so as not to damage them.

Even this simple dispute demonstrates why story formation is necessary to simplify a trier's decision-making burden. For example, focus on one minor piece of circumstantial evidence which Schrackle might provide. Assume that Schrackle testifies that as he made the left turn, he noticed three children on the corner waiting to cross the street. From this, Schrackle asks the trier to infer that he was paying close attention to driving conditions. Potter, in contrast, might ask the trier to infer that Schrackle was distracted by the children, and was therefore not paying attention to the crosswalk. In addition, a trier might come up with other interpretations. For instance, a trier might not believe that any children were on the corner. Or, a trier might believe that Schrackle noticed the children only after he struck Potter, and testified that he noticed them beforehand in an effort to strengthen his case. These four possible interpretations of this single item of evidence, when multiplied by the number of evidentiary items in a trial as routine as *Potter v. Schrackle*, demonstrates a trier's need to make sense of information by organizing it into a story.

In arriving at its final story of "what really happened," a trier may simply choose between the parties' competing versions of reality. In the above hypothetical, for instance, a trier may be impressed by Potter's testimony and believe that it accurately depicts what really happened. Often, however, a trier picks and chooses between parties' stories and formulates its own version of reality, a version which does not exactly coincide with any told in court. For example, a trier may disbelieve Schrackle's testimony that having expensive cabinets in the back of his truck led him to drive extra slowly, yet credit his testimony that Potter ran into the path of his truck from between two parked cars and find him not negligent. Similarly, in a jury trial, a juror may alter a story after listening to the opinions of other jurors. At the end of the day, however, a trier typically develops a single story of what really happened, and based on that story determines the

parties' legal rights.⁶⁰

If use of Bayes' Theorem required wholesale changes in the use of heuristics and common sense by triers, its implementation in actual trials would be inconceivable. So ingrained are these reasoning methods that it is difficult to imagine how we might substitute an entire way of thinking for the present system and still preserve what we think of today as the right to trial by jury. What is facially intriguing, then, about the Bayesian argument is that it allows triers to rely on the same factors they do now.⁶¹ For example, in all the "New Evidence Scholarship"⁶² and its concern for the inferential process there is not one new idea about the factors triers should consider when evaluating testimony. Apparently, the old standbys (e.g., inconsistency, bias, demeanor, and implausibility) are good enough for our David Skydivers, and would be good enough for the Reverend Bayes.

*B. Leakage: How It Occurs, and How Likelihood Ratios
Try To Stop It*

As described above in subsection A, in both a Bayesian and non-Bayesian world, triers would be permitted to continue to rely on everyday reasoning processes. Under current practices, triers use these reasoning processes to formulate stories, apply pertinent substantive law to those stories, and produce verdicts. However, Bayesians contend that in the process of constructing stories triers overlook important data.⁶³ To recapture the probative weight of data that "leaks out"⁶⁴ because of story construction, Bayesians would have triers calculate likelihood ratios for individual items of evidence.⁶⁵

To understand how likelihood ratios might capture lost weight, return again to *Potter v. Schrackle*. Assume that a trier is evaluating Schrackle's claim that he drove slowly because he had expensive kitchen cabinets in the back of his truck. The average, non-Bayesian

⁶⁰ "The story classification stage specifies that the juror's decision takes the form of a classification in which the best match between story features and verdict category features is determined." R. HASTIE, S. PENROD & N. PENNINGTON, *supra* note 57, at 22. See also Pennington & Hastie, *supra* note 24.

⁶¹ See, e.g., Fienberg & Schervish, *supra* note 3, at 816.

⁶² See, e.g., Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof*, 66 B.U.L. REV. 439 (1986).

⁶³ See *supra* p. 600.

⁶⁴ The leakage metaphor is used to describe this loss of probative value in Schum & Martin, *Formal and Empirical Research on Cascaded Inference in Jurisprudence*, 17 LAW & SOC'Y REV. 105 (1982).

⁶⁵ Even apart from our contention that likelihood ratios produce frequentist assessments of probability, compelling triers to analyze evidence on an item-by-item basis rather than as an organic whole might itself distort the fact finding process. See Abu Hareira, *An Early Holistic Conception of Judicial Fact-finding*, 1986 JURID. REV. 79.

trier would evaluate Schrackle's testimony in the context of the construction of an overall story. And, since both Bayesians and non-Bayesians agree that real world events could have taken place in only one way, the trier's final story would probably be either: (i) Schrackle was speeding, despite the fact that he had expensive kitchen cabinets in the back of his truck; or (ii) Schrackle was not speeding, perhaps in part because he had expensive kitchen cabinets in the back of his truck.⁶⁶

The story must tell of Schrackle speeding or not speeding; it cannot coherently tell both.⁶⁷ As a consequence, whatever version of an event a trier's final story contains, the possibility that another version might be true is lost.⁶⁸

Thus, if the trier decides that the cabinets did lead Schrackle not to speed, the trier incorporates this version into its story and ignores the possibility, which probably any trier would admit to, that having expensive cabinets in the back of his truck might *not* have affected his driving.

Conversely, if the trier decides that having expensive cabinets in the back of the truck did not affect how Schrackle drove, in organizing its story the trier will no longer consider the possibility, which again almost any trier would concede must exist, that the cabinets *did* affect Schrackle's driving.

In other words, whichever story a trier ultimately formulates, the possibility that the alternate version might be correct leaks out and is eliminated from consideration.⁶⁹ The ability to capture this leaked

⁶⁶ In this example, we assume that both parties agree that Schrackle in fact had kitchen cabinets in the back of his truck. Of course, in reality the parties might dispute that fact as well. This would require a trier to resolve an additional issue, but would have no effect on our analysis.

⁶⁷ Bayesians certainly accept this assertion. In Bayesian analysis, the probability that a fact is true and that it is not true must always add up to one. See Fienberg & Schervish, *supra* note 3, at 774-75.

⁶⁸ Of course, prior to formulating a final story, a trier might waver back and forth between which version of an event to accept. The final story, however, can hold no more than one version.

⁶⁹ A different form of leakage which is of concern to some Bayesian analysts concerns the trier's tendency to forget about or ignore items of evidence during the decision-making process. Thus, when mock jurors are forced to "decompose" the factfinding process by attaching likelihood ratios to individual items of evidence, they are less likely to rely on particularly salient, and therefore more memorable, testimony to the exclusion of other items of evidence. Schum & Martin, *supra* note 64, at 134-37. But see R. HASTIE, S. PENROD & N. PENNINGTON, *supra* note 57, at 230, where the authors of an impressive empirical study conclude that the abilities of jurors as a group to reconstruct testimony "occur(s) with thoroughness and precision."

However, while compelling triers to adhere to Bayesian analysis undoubtedly tends to ensure that triers at least think about almost all items of evidence, we have both theoretical and practical concerns about this supposed benefit of Bayesian analysis. On the theoretical

probative value is the essence of likelihood ratios,⁷⁰ and is what fuels the thrust of Bayesian attacks on the current fact-finding process.

How does the use of likelihood ratios seemingly capture leaked probative value? Recall that under Bayes' Theorem a prior probability is altered by the ratio of the likelihood that a given piece of information would be found if a certain proposition were true and the likelihood that same piece of information would be found if that proposition were not true. Thus, likelihood ratios appear to assign some weight to all items of evidence a trier finds relevant. By contrast, when a trier formulates a story about what happened, relevant evidence and inferences that would support competing stories are typically given zero weight.

For example, assume that in *Potter v. Schrackle*, Potter's attorney argues that one factor that makes her testimony credible is that she looked the trier squarely in the eye when she testified. Under Bayesian rules, the trier would have to place a value both on the likelihood that a witness would look the trier in the eye if she were telling the truth, and on the likelihood that she would look the trier in the eye if she were not. Whereas in the current system the trier's final story would consist only of belief or disbelief of Potter, Bayes' Theorem, through the device of the likelihood ratio, seems to take into account both possibilities.⁷¹

level, "forgetting" may be necessary if triers are to perform their inferential tasks. Trial is an essentially oral milieu, and in an oral culture people need to employ a "system of elimination" from memory if they are to attend to what is important. Goody & Watt, *The Consequences of Literacy*, in *LITERACY IN TRADITIONAL SOCIETIES* 56-57 (J. Goody ed. 1968). Bayesians can assure that triers attend to every individual item of evidence only by giving them written lists of evidence. So doing may change the entire character of fact finding. Our response to written words tends to be abstract and logical, whereas in oral dialogue we rely more on intuition and emotion. See Bergman, *The War Between the States (of Mind): Oral Versus Textual Reasoning*, 40 *ARK. L. REV.* 505 (1987). Some may favor such a change; our only point is that Bayesian procedures may go beyond simply assuring that each item of evidence is attended to during decision making.

Our practical concerns stem from differences between actual trials and the Bayesian experiments. Unlike the experimental situations, in actual trials attorneys are able to structure their presentation of evidence during testimony and argument in such a way as to stress the most significant evidence. See, e.g., P. BERGMAN, *supra* note 55, at 17-23. Thus, although this sort of leakage is potentially a serious problem, it can be mitigated to some extent by an effective presentation of the case by counsel.

⁷⁰ "A likelihood ratio, appropriately formulated, allows you to incorporate every source of uncertainty you recognize and also prescribes how these uncertainties ought to be combined in the process of assessing the probative weight of the evidence setting up the reasoning chain." Schum, *supra* note 8, at 862.

⁷¹ Some limitations of the concept of likelihood ratios may be immediately evident, and in any event have already been the subject of prior criticism. For example, most people are not terribly good at attaching specific numbers to levels of confidence. It is one thing for a trier to say, "I didn't place much faith in Jones' assertion that the cab was green." It is quite another

C. *Likelihood Ratios Convert Believability Judgments
to Frequency Assessments*

Whatever its shortcomings, story formation does enable triers to determine what they believe really happened: "[T]he trier is likely to understand what happened in the past in story terms, and to evaluate legal principles in terms of that understanding."⁷² If Bayes' Theorem would enable triers to decide how they believed unique historical events actually took place, and did so without the leakage that accompanies story formation, the theorem's claim to the theoretical high ground would be extremely persuasive.

As we understand the Bayesian position, its claim to producing believability judgments would run along these lines: Bayesian analysis is subjective, not frequentist. Indeed, we are powerful,⁷³ even emotional,⁷⁴ competitors with the frequentists. Subjective Bayesianism is not frequentist because when forming likelihood ratios triers evaluate all of the evidence in a specific case according to their subjective beliefs about how the world operates.⁷⁵ Thus, likelihood ratios reflect a person's subjective degree of confidence that a proposition is true.⁷⁶ Accordingly, Subjective Bayesianism allows triers to reflect their personal degrees of belief,⁷⁷ and thereby to make believability judgments; that is, to determine what they believe really happened "*in this case*."⁷⁸

We illustrate what we understand to be the Bayesian position through the following example from *Potter v. Schrackle*. To help de-

to expect the juror to label that assertion with a number: My level of belief in Jones' assertion that the cab was green is .3." The distinctions among .1, .2, .3, .4, and so on are unlikely to be little more than ragged estimates. Yet such estimates, when plugged into Bayes' theorem, may determine a trial's outcome. This will be particularly true when, as in most trials, triers must attach numerical labels to a large number of items of evidence. Verdicts, it seems, should not depend on which number a trier haphazardly happens to choose.

Though we share these concerns, for purposes of this essay, we put them aside. We assume that triers can make realistic numerical assessments of likelihood.

⁷² P. BERGMAN, *supra* note 55, at 11; *see also* Pennington & Hastie, *supra* note 24.

⁷³ *See* Efron, *supra* note 13.

⁷⁴ "Every statistician would be a Bayesian if he took the trouble to read the literature thoroughly and was honest enough to admit that he might have been wrong." Lindley, *Comment*, 40 AM. STATISTICIAN 6, 7 (1986).

⁷⁵ A subjective probability is simply a person's opinion. *See* Shafer, *supra* note 17, at 801. Subjective Bayesianism "was developed originally in terms of gambles and betting odds." Fienberg & Schervish, *supra* note 3, at 780.

⁷⁶ *See* Dant, *Gambling on the Truth: The Use of Purely Statistical Evidence as a Basis for Civil Liability*, 22 COL. J.L. & SOC. PROBS. 31, 38 (1988); Finkelstein & Fairley, *supra* note 2, at 504; Tribe, *supra* note 2, at 1347.

⁷⁷ Fienberg & Schervish, *supra* note 3, at 783.

⁷⁸ Tribe, *supra* note 2, at 1347 (emphasis in original). *See also*, W. TWINING, *supra* note 18, at 121.

termine whether Schrackle was speeding, a trier is asked to formulate a likelihood ratio for the evidence (testimony) that Schrackle had expensive cabinets in the back of his truck. The trier could consider all of the circumstances of the *Potter v. Schrackle* case when determining the likelihood that if Schrackle had been speeding he would have had expensive cabinets in his truck, and the likelihood that if he had not been speeding he would have had them.⁷⁹ If likelihood ratios for all the other evidence were calculated and plugged into Bayes' Theorem, the outcome, a Bayesian would argue, would be the trier's believability judgment (or probability assessment) about what really happened in *Potter v. Schrackle*.

But likelihood ratios cannot produce believability judgments. To see why this is so, assume that trial has concluded in the case of *Potter v. Schrackle*. Rather than ask the trier to deliberate in the usual way, we will ask the trier to produce likelihood ratios. We ask for a likelihood ratio based on the evidence of expensive cabinets.⁸⁰ The trier thus has to assess: (1) How likely is it that Schrackle would have

⁷⁹ In order to arrive at a believability judgment, we think that Bayesian methodology would have to permit the trier to consider all of the evidence introduced in the case when forming a likelihood ratio for any individual item of evidence. If the trier were required to form likelihood ratios for individual items of evidence in a vacuum (*i.e.*, without taking into account all the other testimony in the case), the result would clearly be frequentist. For Schrackle's testimonial evidence about the expensive cabinets, for example, the trier would be forming a likelihood ratio based on the abstract issue: Assuming a truck driver is speeding (or not speeding), how likely is it that he would have expensive cabinets in the truck? Addressing the issue at this level of abstraction requires the trier to ignore all the rebuttal evidence (including all the evidence relating to witness credibility) which might explain why in this particular case Schrackle might have been speeding despite having expensive cabinets in his truck.

There is a possible middle ground between permitting the trier to consider all of the evidence when forming likelihood ratios and requiring that likelihood ratios be formed in a vacuum. The trier might be instructed to consider some items of evidence and ignore others when forming likelihood ratios. *See, e.g.*, J. PEARL, *supra* note 33, at 44-46 (giving an example of ignoring the testimony of witness one when forming a likelihood ratio for the testimony of witness two). Not only will this approach still produce a frequentist assessment (because the trier must ignore some of the evidence in a particular case), but also there seems no possible way to determine how to instruct triers as to the evidence they can and cannot consider when forming likelihood ratios. Finally, this middle ground is based on a questionable assumption: "Bayesian practitioners claim that people are capable of retracing the origins of their beliefs and of entertaining hypothetical questions such as [what would you believe about witness A's testimony if you pretend that you had never heard witness B's testimony?]" *Id.* at 46. However, a trier's ability to do this is dubious. *See* 1A WIGMORE ON EVIDENCE, *supra* note 18, at 1062 n.23 ("[I]n a complex trial with many separate items of evidence, it may be very difficult for jurors to disentangle [their] prior estimates of probability for their intuitive posterior estimates of probability in order to perform in the proper fashion the computations required by Bayes' theorem."); *see also infra* pp. 610-11.

⁸⁰ We eliminate a credibility issue by assuming that both parties agree that Schrackle did in fact have expensive cabinets in the back of his truck. This assumption simplifies the analysis without affecting our argument.

expensive cabinets in his truck assuming he was speeding?⁸¹ and (2) How likely is it that Schrackle would have expensive cabinets in his truck assuming he was not speeding?⁸²

A Bayesian request to a trier for a likelihood ratio might produce the following conversation:

BAYESIAN: I'd like you to give me a likelihood ratio for the cabinets that Schrackle had in the back of his truck.

TRIER: Well, as you know, I've already made sense of the evidence by constructing a story. In my story, Schrackle was speeding and everyone concedes he had the cabinets in his truck. Therefore, *in this particular case*, the likelihood that Schrackle would have expensive cabinets in his truck while speeding is one; the likelihood that he would have cabinets and not speed is zero.

BAYESIAN: What you've given me is not a likelihood ratio.⁸³ You've simply identified which hypothesis (or story) you've chosen to believe. You need to tell me more than just what you think happened in this case. I want you to assume for the moment that Schrackle was speeding. Then, taking into account all the evidence at trial, tell me the likelihood that he would have had expensive cabinets in his truck.

TRIER: When you say I should take into account all the evidence in this case, you must realize that "evidence" consists of more than just the testimony at trial. I had to comprehend and remember the testimony as it was introduced. To do so, I constructed a story about what I think really happened.⁸⁴ As I constructed the story, I tentatively assessed the credibility of the evidence by taking into account such factors as consistency with common experience, internal consistency of the conflicting stories told by the witnesses, adequacy of detail, and the witnesses' demeanors.⁸⁵ Can I assume that the

⁸¹ Those readers who find the wording of likelihood ratios difficult may want to restate them as conditional probabilities. For example, the statement, "How likely is it that Schrackle would have expensive cabinets in his truck assuming he was speeding," can be converted into, "If Schrackle had expensive cabinets in his truck how likely is it he was speeding?" For our purposes, there is no significant difference between these two formulations.

⁸² It is by no means troubling to subjective Bayesian methodology that different triers might make different likelihood assessments. See Fienberg & Schervish, *supra* note 3, at 773-74. Quite the opposite is true: a supposed value of subjective Bayesianism is that it allows each trier to react according to her or his subjective state of mind. Of course, were Bayesian methodology employed in actual jury trials, some method of cumulating or averaging the different jurors' likelihood ratios would be necessary.

⁸³ This is not a likelihood ratio because the ratio would be the same for *every* undisputed item of evidence in the case. Furthermore, if this were a true likelihood ratio it would mean that the trier has concluded that drivers with expensive cabinets in their trucks always speed.

⁸⁴ See Moore, *supra* note 52; Pennington & Hastie, *supra* note 24.

⁸⁵ Moore, *supra* note 52; Pennington & Hastie, *supra* note 24.

conclusions and inferences I have drawn and incorporated into my story based on these common sense reasoning tools are part of the evidence I can take into account?

BAYESIAN: Yes, you can take into account what your common sense tells you happened in this case when forming likelihood ratios.⁸⁶

TRIER: Well then, we're right back where we started from. My story is that on this particular occasion Schrackle was speeding. So the likelihood that he had cabinets and would speed is one in this case. The likelihood is zero that he would not. Of course, I'm not 100% sure that my story is true. I'm about .8 certain that my story is true now, and I might be more or less certain that my story is true after I deliberate with the other jurors and we argue about the plausibility of the various witnesses' stories, the witnesses' demeanors, et cetera. But regardless of how confident I may be that my story reflects what actually happened in this case, it doesn't seem that my level of confidence in my story is what you're interested in when you ask for a likelihood ratio. You're looking for something different, aren't you?

BAYESIAN: You're right. I'm not just asking you to tell me how confident you are that your story is correct. Let's try another approach.⁸⁷ Continue to keep in mind the stories told by Potter and Schrackle, excluding the statement about the cabinets. So you have before you all the evidence except for the evidence about the cabinets. Given everything you know about Schrackle and the events on the day of the accident, is it more likely that he would drive fast if he had the expensive cabinets in the truck than if he didn't?

TRIER: Let me get this straight. You want me to: (1) imagine what story I would have constructed in order to comprehend and remember the testimony in this case if the trial had been just as it was except for the testimony about the cabinets; and (2) to then tell you how much more likely it would be that Schrackle was speeding if I added in the testimony about the cabinets?

BAYESIAN: You've got it!

TRIER: Well, if I've got it I want to get rid of it. I can't do what you're asking. First of all, I can't tell you what sort of story I would have constructed about this case in the absence of the testimony about

⁸⁶ Anne Martin has noted that Bayesians evaluate by initially assigning only "provisional likelihoods for individual pieces of evidence." Martin, *supra* note 35, at 714. Only after all the evidence has been received are likelihood ratios finally formulated. *See id.* This delay, of course, means that a trier will first evaluate the story, and then assign likelihood ratios with that determination already in mind.

⁸⁷ The approach that follows in the text was suggested by David Kaye in response to the previous draft of our article presented at the conference.

the cabinets. What I remember about this case is defined in large part by the story I constructed as the testimony unfolded.⁸⁸ I don't have a verbatim transcript of the trial from which I can simply delete out the testimony about the cabinets and then reread. And even if you gave me such a transcript, my initial story, which took into account the cabinets, would influence the story I would construct as I read the transcript. And I assume you'll want me to repeat this process for every item of evidence in the case, so I'd be required to construct and then forget hundreds or thousands of stories to resolve this case. I don't think it can be done.⁸⁹

But even assuming that I could surmount these cognitive limitations, I don't see how I could give you the likelihood ratio you're looking for.

BAYESIAN: Why not?

TRIER: When I constructed my story I focused on interrelationships between various items of evidence. For example, I thought about the adequacy of detail and internal consistency in Potter's story. I needed to assess the credibility of the competing versions of events, and these are some of the factors that seemed most important to my determination of what happened. Now, it might well be that the cabinets are unrelated to these interrelationships. That is, if I'm strongly persuaded by the consistency and plausibility of Potter's story in this case, evidence about the cabinets won't change my view of what happened. In effect, I would assign zero weight to the cabinets evidence in this particular case, even though I know that it is generally relevant in these sorts of cases, and that in another case where the plaintiff told a less believable story, I might factor in the cabinets evidence.

BAYESIAN: You've made it too easy. Suppose the evidence about the cabinets is related to the interrelationships that you have focused on to construct your initial story. Then you can give me a likelihood ratio, can't you?

TRIER: No, I can't. The cabinets evidence may increase or decrease my confidence in my initial story, or as you Bayesians like to say, the probability that my story is true. The cabinets evidence may even cause me to change my story. But neither of those situations will permit me to give you a likelihood ratio.

BAYESIAN: This can't be so hard. You mean you can't tell me if

⁸⁸ See Moore, *supra* note 52.

⁸⁹ For further discussion of this point, see *supra* note 79.

it is more likely that Schrackle would drive fast if he had expensive cabinets in his truck than if he didn't?

TRIER: If you're asking me about what happened in this particular case, that's right, I can't tell you that. Assume that I find the cabinets evidence, in the light of the other evidence, extremely important. The evidence is so important that it causes me to disbelieve Potter's version of events. Before, I was impressed by the fact that she was able to recall part of the license plate of the car that drove past her just before Schrackle hit her; now I think that maybe Potter just made that up. What the evidence causes me to do, then, is to construct a whole new story in which Schrackle is not speeding. The cabinets evidence does not just make it less likely that Schrackle would drive fast; it affects my evaluation of the competing versions of events. So I just can't give you a likelihood ratio when I'm trying to evaluate the particular events in *Potter v. Schrackle*.

BAYESIAN: What about the other situation that you mentioned, in which the new evidence affects the probability that your initial story is accurate, but doesn't cause you to change that initial story?

TRIER: You Bayesians are dogged to the end—I give you credit for that. Let's say that I think that the cabinets evidence does have some evidentiary significance. It casts some doubt on Potter's story—after all, it does tend to give Schrackle an incentive to drive slowly. So I rethink my initial story. At the end of the day, though, I decide that Schrackle's overall story still isn't internally consistent. Thus, I stand by the initial story, though I may be a little less confident in it. So, in this particular case, the likelihood that Schrackle would have expensive cabinets and speed remains one; the likelihood that he would have them and not speed remains zero. As we've agreed, that's not a likelihood ratio.

BAYESIAN: I'm very suspicious of this story business. You seem close minded. You compose a story as you listen to the evidence. Then I ask you to formulate a likelihood ratio, which would enable you to account for weight you might have ignored by being too quick with your story. You tell me you can't do it because you already have your story. If you ask me, this seems to be a head in the sand attitude.

TRIER: I know it seems like that to you. But as I mentioned earlier, I'm more than willing to alter my story if necessary. A story is simply an organizing tool. I can change a story if it fails to account for an important piece of information, or if another juror develops a more satisfactory one. But without that organizing tool, I cannot make sense of evidence.

BAYESIAN: So you can never give me a likelihood ratio?

TRIER: I can, but only based on subjective frequencies. If I'm not limited to whether Schrackle was speeding on this particular occasion, but can think about how often truck drivers like Schrackle speed if they have cabinets, I can answer your question.⁹⁰ I think that the likelihood that truck drivers like Schrackle speed when they have expensive cabinets in their trucks is .3; the likelihood that they speed when they don't have expensive cabinets is .4. But remember, these likelihoods are not tied to the case of *Potter v. Schrackle*. I've had to ignore the factors which I would otherwise use to determine the credibility of competing versions of history.

BAYESIAN: I'm a bit disappointed in you. When I agreed to come into this article, it was because I thought you were an ideal juror.

TRIER: More Bayesian sleight-of-hand! What you claim to be ideal is Bayes' Theorem, not me. Even if you gave me some training in statistical analysis, I'm still flesh and blood and I'd still use my everyday reasoning skills. Maybe if I were a supercomputer I could keep track of tens of thousands of stories, each one simultaneously ignoring a different item of evidence and adjusting every other story according to the likelihood ratios. To tell you the truth, I don't look forward to that day.

As the dialogue illustrates, the Bayesians can coax a likelihood ratio out of a trier, but the ratio is divorced from believability. The trier's response is a frequentist assessment. The response indicates what the trier believes is likely to happen in "these sorts of cases," not what the trier believes actually happened in *Potter v. Schrackle*. Formulated for each item of evidence and combined with a prior probability as mandated by Bayes' Theorem, likelihood ratios can only tell us how probable the trier thinks it is that truck drivers speed in these sorts of cases. But the ratios cannot determine believability.

Thus, our concerns about applying Bayes' Theorem in actual decision-making situations are neither semantic nor based on such practical difficulties as triers having to place numerical labels on uncertain feelings. Rather, our concern is that formulating likelihood ratios requires triers to ignore how they believe events actually took place in a given case. If triers respond based on their beliefs about how events

⁹⁰ This is not the only type of frequentist interpretation a trier might give in this situation. Instead of thinking about other truck drivers, a trier might interpret the request for a likelihood ratio as calling for an estimate of how often Schrackle might speed with expensive cabinets in his truck under circumstances that were similar but in undefined ways different from those that existed in *Potter v. Schrackle*. This interpretation might produce a different number, but it would remain a frequentist assessment.

actually occurred, they give statements of belief, not likelihood ratios.⁹¹

Interestingly, though our present system of trial is inconsistent with Bayes' Theorem, one could imagine trials organized along Bayesian lines. One would have to eliminate testimony in story form. In its place, lawyers could write out individual items of evidence which they want the trier to consider. These individual items could be randomly scrambled by a computer to eliminate any vestige of a story element. The individual items would then be presented to triers, who would assign them probabilities. The triers, of course, would never be told, and would be told not to guess at, what particular series of events gave rise to the evidence. Each item would be assigned a likelihood ratio, the ratios would be combined with an initial probability as Bayes' Theorem dictates, and a final posterior probability produced.

Such a procedure would allow a trier to make subjective probability assessments, but those assessments would be frequentist. We might learn, for example, that six times out of ten a person in the situation of a particular criminal defendant would be guilty, or that a building contractor would be negligent, or that a corporation would commit an antitrust violation. But such frequency assessments are not enough.⁹² The assessment we want triers to make is whether *this particular case* is one of the six, or one of the four. That assessment requires a trier to make a believability judgment.

D. *Potential Uses of Bayes' Theorem*

Must one conclude from this essay that Bayes' Theorem is irrelevant in the trial context? Not at all. The likelihood ratio concept embodied in the theorem is well suited to helping lawyers gauge the probative value of evidence, and to helping judges trying to assess relevance. When engaged in these tasks, judges and lawyers are not making credibility assessments about what happened. Rather, when determining whether evidence is relevant, a judge assesses only whether an item of evidence makes an issue in dispute more or less

⁹¹ Our conclusion that likelihood ratios cannot convert frequency assessments into believability judgments leaves open the interesting question of how triers usually transform subjective frequency assessments into believability judgments. In other words, how do subjective frequencies become stories? This question is beyond the scope of the present essay. In general, however, we suggest that triers (often subconsciously) commonly organize their beliefs about how the world operates into "schemas." Their beliefs about how events took place are often determined by how closely witnesses' stories match their schemas. See Moore, *supra* note 52.

⁹² "[I]t is meaningless to speak of the probability of the defendant's guilt in terms of the number of times he would be guilty in an infinite number of exactly similar cases because . . . even if there were many identical cases the court must reach a verdict, not a ratio, in the case at bar." Kaplan, *supra* note 2, at 1066.

likely. And a lawyer, when determining how (if at all) an item of evidence fits into his or her own version of events, typically assesses what inference a trier might draw from it. Since in neither circumstance does the judge or lawyer make a credibility assessment, the frequency analysis which the likelihood ratio provides may well be helpful.

For example, return to the issue of the likelihood that if Schrackle were driving slowly, he would have expensive kitchen cabinets in the back of his truck. Assume (though admittedly the event is unlikely) that Potter's counsel were to object to the relevance of the evidence when Schrackle offered it. To rule on the objection, a judge should consider only the likelihood, in situations such as Potter's and Schrackle's in general, that someone with expensive kitchen cabinets would drive slowly, as compared to the likelihood that one without expensive cabinets would drive slowly. If these two events are equally likely, the evidence is irrelevant. If the former event is somewhat more likely, the judge should admit the evidence (subject to a claim that the probative value is outweighed by prejudicial effect).⁹³ However, the judge should not rule on whether in this particular case the cabinets did in fact induce Schrackle to drive slowly. Hence, the frequentist assessment is appropriate; frequency assessments are circumstantial evidence of believability.⁹⁴

Likelihood ratios might also help counsel evaluate the probative value of evidence. This time, consider the probative value of the evidence that Potter looked the trier squarely in the eye when she testified. Each counsel typically attempts to persuade the trier to compose a story that favors her or his client. When deciding what story to formulate, a trier relies heavily on how events typically occur in daily life. Initially, then, each lawyer makes a frequentist appeal. Potter's counsel would ask a trier to believe based on everyday experience that looking someone in the eye is highly correlated with truth telling; Schrackle's counsel would ask the trier to use its experience to recognize that people often look others in the eye even when they are not telling the truth. Which version a trier accepts is a believability judgment that requires a trier to move from how events are likely to occur to how they did occur. This is a nonfrequentist assessment that Bayes' Theorem cannot control. However, the likelihood ratios can help counsel recognize how circumstantial evidence can cut in multiple directions, and the ratios thereby assist counsel to gauge and argue the probative value of the evidence.

⁹³ See, e.g., FED. R. EVID. 403.

⁹⁴ See *supra* note 21.

V. PRIOR PROBABILITIES AND BELIEVABILITY JUDGMENTS

As is the case with likelihood ratios, Bayesians believe that prior probabilities can capture evidential weight typically lost during story construction. Once again, our response is that the effort to capture this lost weight produces a frequency assessment rather than a believability judgment. The reason is that when triers make believability judgments, prior probabilities are not necessarily entitled to the weight they are assigned by Bayes' Theorem.

Reconsider the infamous blue cab hypothetical. As you may recall, a cab was involved in a hit and run accident at night. Two cab companies, Green Cab Company and Blue Cab Company, operate in the city. You are given the following data: (a) 85% of the cabs in the city are green and 15% are blue; (b) after a witness testified that the cab was blue, the court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time.

The plaintiff is suing the Blue Cab Company. The question is, what is the probability that the cab involved in the accident was blue rather than green?⁹⁵

According to Bayes' Theorem, .41 is the correct answer to this question.⁹⁶ Despite the witness' testimony that the cab was blue, the Bayesian probability is that it was green. The reason why the Bayesian probability inculcates Green Cab Company is that the Green Cab Company's "base rate [the prior probability] is more extreme than the witness is credible."⁹⁷ But most people presented with this hypothetical "incorrectly" estimate the probability of the cab being blue at .80. They make this "error" by relying entirely on the evidence of the witness' accuracy and ignoring the prior probability.⁹⁸ That is, the weight assigned by Bayes' Theorem to the prior probability "leaks out."

We unashamedly confess that our initial reaction too was that 80% was the correct answer. Our thinking went something like this. If the witness is correct, the cab is blue. The witness is correct 80% of the time. No evidence suggests that the color of a cab in any way affects the witness' ability to identify it. Thus, there is no reason for

⁹⁵ The hypothetical is taken from Tversky & Kahneman, *supra* note 47, at 156-58.

⁹⁶ *Id.*

⁹⁷ *Id.* at 157.

⁹⁸ The prior probability in this case is .85 because there is no additional information which might affect the prior probability—for example, driving records of the employees of the cab companies, how many green cabs are on duty at night, et cetera—in evidence.

thinking the witness should be correct less than 80% of the time when he sees a blue cab. Therefore, the probability that the cab is blue is 80% and the prior probability is irrelevant.

Having spent part of our remaining precious youth on this hypothetical, we now realize that the "correctness" of the .41 outcome results from two crucial assumptions hidden within the hypothetical.⁹⁹ First, one must assume that the witness knows that all of the cabs in the city are either green or blue, and thus will never incorrectly identify a cab as black, yellow, or some other color. Second, the witness will never claim inability to identify or recall the color of a cab.¹⁰⁰ Consequently, the witness can give but two answers: blue or green.

With these assumptions in place, one can tease out the Bayesian reasoning. Over the course of 100 trials, a witness who is correct 80% of the time will give twenty wrong answers. Assume that the errors are randomly distributed. More than fifteen of the twenty times the witness is wrong, the witness will say, the cab was blue. Therefore, if one considers all the "cab was blue" statements in the 100 trials, more of these statements will be incorrect than correct. That is, there are fifteen blue cabs, but seventeen incorrect "the cab was blue" statements. Ergo, the Bayes' Theorem .41 probability that the cab was blue. Therefore, under Bayesian rules, in a suit against Blue Cab Company, plaintiff will probably lose.¹⁰¹

But should a rational trier faced with the blue cab case weigh the prior probability the same way that Bayes' Theorem does? Not necessarily. Recall that a prior probability is but circumstantial evidence of believability. Like other items of circumstantial evidence, its weight (or probative value) depends on the circumstances of a particular case. In the blue cab case, for example, the significance of the prior probability depends on whether the witness knows that most of the cabs in the locale are green. But if the witness is aware of this, the prior probability might cut in exactly the opposite direction that Bayes' Theorem mandates. Consider why this is so.

The witness testifies, "the cab was blue." In determining

⁹⁹ To our knowledge, these assumptions are nowhere stated, which makes it difficult for anyone other than statisticians to analyze them.

¹⁰⁰ In an actual case, of course, this second assumption would never be true. A witness can always say, I couldn't tell what color that cab was.

¹⁰¹ The assumptions hidden in the blue cab hypothetical are crucial. For example, assume that all of the 20 wrong answers in the 100 trials were, "The cab was yellow," or, "I couldn't tell what color that cab was." Then, all the "cab was blue" statements would be correct, and Bayes' Theorem could not automatically assign a .41 posterior probability to the likelihood that the cab was blue.

whether the witness was accurate on this particular occasion, a trier might reason as follows:

Clearly, the witness would be accurate if the witness had gotten a very good look at the color of the cab. But what testimony would the witness have given if the witness were somewhat unsure of the cab's color? Assuming that the witness knows that most of the cabs in the locale are green, the witness would probably fill in his uncertainty by saying, green.¹⁰² The witness' willingness to say blue, therefore, indicates that the witness did not simply fill in this information, and therefore the most believable scenario is that the witness got a really good look at the cab. It must really have been blue.

In short, a trier might regard the prior probability that the cab was green, valued so heavily by Bayes' Theorem in the direction of the green outcome as circumstantial evidence that the cab was blue.

This reasoning seems reasonable to us, and shows that in an individual case the prior probabilities could point in the opposite direction than they would utilizing Bayes' Theorem.¹⁰³ We think this results from the trier's construction of a story about what really happened in a particular case. For the abstract class, accidents involving cabs, Bayes' Theorem correctly assigns a fixed weight to a prior probability. But triers deal with concrete cases where the weight of a prior probability varies from case to case.¹⁰⁴ The believability of a particular scenario is not always influenced in the same way by the prior probability for the class of events of which that scenario is a member.

CONCLUSION

Return to Maria and David, our two intrepid skydivers, who

¹⁰² That such reasoning might well be appropriate is supported by studies of how expectancy affects eyewitness identification. Such studies tend to show that when a white eyewitness observes an interracial scene in which a white person is the aggressor, the witness may remember the black person as the aggressor. See Johnson, *Cross-Racial Identification Errors in Criminal Cases*, 69 CORNELL L. REV. 934, 950-51 (1984). In a similar manner, a trier might logically infer that because an uncertain witness who expects to see a green cab will identify a cab as green, the testimony that the cab was blue is especially likely to be valid.

¹⁰³ Bayesians might argue that the analysis we put forth could be taken into account by simply raising the estimate of the credibility of the witness and then combining that number with the prior probability. This seems illogical to us. If the witness is correct, the cab is blue. If the trier accepts the argument that prior probabilities tend to prove that the witness was more likely correct in this particular case, we see no reason why the same evidence should require the trier to decrease his or her estimate of the probability that the witness was correct. Counsel can of course make such an argument. But like any other argument, the trier could find it unpersuasive.

¹⁰⁴ In the blue cab hypothetical, for example, the significance of the prior probabilities depends on whether the witness knows that green cabs predominate.

have now become estranged lovers while remaining skydiving buddies. They are about to leap, for a second time, from the belly of an aircraft. David again suddenly snaps his fingers, removes his parachute, and prepares to jump.

"Hey!" yells Maria. "Are you at this again? You can't jump without a parachute—you'll be killed!"

"Nothing to worry about," David replies. "I'm a vegetarian. My doctor told me that vegetarians live much longer than flesh eaters."

David jumps and dies. David's wife sues Maria for wrongful death, alleging that Maria, in a jealous rage, pushed him from the plane before he could don his parachute. Maria tells the foregoing story in her defense.

Viewed from a subjective frequentist standpoint, Maria's story seems improbable. Events in this sort of case do not usually occur in the way she described. If a trier were limited to prior probabilities and likelihood ratios derived from what usually happens in this class of cases, the Marias of the world would always lose. If Maria and all other litigants who tell unusual or statistically improbable stories are to have their day in court, it must be possible for a trier to conclude that a statistically improbable story is nonetheless believable. To do so, a trier must be able to assign weight to evidence other than solely on the basis of what usually happens. As a result, the potential probative value of certain items of evidence may well leak out. However, some leakage is a necessary by-product of story construction and believability judgments.

Condoning leakage does not make us apologists for business as usual at trial. Leakage is potentially a serious problem, and should be minimized. Bayesian analysis can help in that regard by helping lawyers to assess thoroughly the potential probative value of evidence. In turn, lawyers may be better able to argue to a trier about the probative value of evidence, thus making the trier's believability judgment more fully informed.

The authors thank and excuse juror number three, the Reverend Thomas Bayes; and resuscitate, thank, and excuse juror number eight, David Skydiver.

