The Laws of Probability and the Law of the Land

David Kaye†

Lawyers are wordsmiths, not number crunchers. Thus, quantitative or mathematical evidence has long been a source of bewilderment to the profession.¹ Some years ago, a lawyer-statistician team, Michael Finkelstein and William Fairley, suggested a modest use of an elementary formula of probability theory, known as Bayes's formula, to aid jurors in assessing statistical identification evidence.² In an eloquent response, Laurence Tribe argued against the proposal on two general grounds.³ First, he offered a variety of practical reasons to suggest that probabilities calculated according to the formula would not be as accurate or useful as they might, at first blush, appear. Second, he pointed out how values other than the accuracy of factfinding might be undermined by the explicit use of probability calculations, especially in criminal cases.⁴ In the succeeding years, no one took the Finkelstein-Fairley proposal very seriously, although the question is not as clear-cut as most people assume.⁵

More recently, Finkelstein published a slightly revised version of his earlier proposal.⁶ This time, two law professors, Lea Brilmayer

† Professor of Law, Arizona State University. I am grateful to Dennis Karjala and David Schum for helpful discussions of the topics treated in this article. Ira Ellman also made many valuable contributions to this article, which grew out of our work on statistical evidence in paternity disputes. The text at notes 57-63 *infra* especially bears his imprint.

¹ See, e.g., *State v. Smiley*, 27 Ariz. App. 314, 554 P.2d 910 (1976); *State v. Sneed*, 76 N.M. 349, 414 P.2d 858 (1966); *Moore v. Leininger*, 299 Pa. 380, 149 A. 662 (1930); Brilmayer & Kornhauser, *Review: Quantitative Methods and Legal Decisions*, 46 U. CHI. L. REV. 116, 135-48 (1978); Jaffee, *Comment on the Judicial Use of HLA Paternity Test Results and Other Statistical Evidence: A Reply to Terasaki*, 17 J. FAM. L. 457 (1979).

² Finkelstein & Fairley, *A Bayesian Approach to Identification Evidence*, 83 HARV. L. REV. 489 (1970). For an earlier proposal along these lines, see I. GOOD, *PROBABILITY AND THE WEIGHING OF EVIDENCE* 66-67 (1950).

³ Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1350-1378 (1971).

⁴ Some implications of this observation are explored in Nesson, *Reasonable Doubt and Permissible Inferences: The Value of Complexity*, 92 HARV. L. REV. 1187 (1979).

⁵ A few courts have admitted Bayesian probability calculations into evidence. These cases are criticized in Ellman & Kaye, *Probabilities and Proof: Can HLA and Blood Testing Prove Paternity?*, 55 N.Y.U. L. REV. (forthcoming), in which a more modest use of Bayes's Theorem is defended.

⁶ M. FINKELSTEIN, *QUANTITATIVE METHODS IN LAW* 78-104 (1978). For further analysis of the advisability of using Bayes's formula to evaluate statistical identification evidence, see, for example, Fairley, *Probabilistic Analysis of Identification Evidence*, 2 J. LEGAL STUD. 493 (1973); Finkelstein & Fairley, *A Comment on "Trial by Mathematics"*, 84 HARV. L. REV. 1801 (1971); Gerjuoy, *The Relevance of Probability Theory to Problems of Relevance*, 18

and Lewis Kornhauser, entered into the rather academic debate. In an article published in this *Review*, they argued on epistemological grounds that the proposed use of probability theory might not be rational or meaningful.⁷ Their work is typical of a growing body of thought that holds that the type of probabilities pertaining to disputed facts in litigation do not obey the usual laws of probability.⁸ In reaching this seemingly paradoxical conclusion, Brilmayer and Kornhauser join a two-hundred-year-old debate among logicians, philosophers, and statisticians concerning the foundations and applications of probability theory.

In this article I hope to show that the new wave of skepticism about the application of probability theory to legal factfinding rests on fundamental misconceptions about the philosophical debate over the meaning of probability, the character of rational decisionmaking, and the values of the legal system. With respect to the specific controversy about the use of Bayes's Theorem at trial, I suggest that

JURIMETRICS J. 1 (1977); Lindley, *Probabilities and the Law*, in *UTILITY, PROBABILITY AND HUMAN DECISION MAKING* 223 (D. Wendt & C. Vlek eds. 1975); Tribe, *A Further Critique of Mathematical Proof*, 84 HARV. L. REV. 1810 (1971). Other writers have used probability theory heuristically, to explain, criticize, and justify various rules of evidence and procedure. See, e.g., Gelfand & Solomon, *Analyzing the Decision-Making Process of the American Jury*, 70 J. AM. STATISTICAL A. 305 (1975); Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065 (1968); Kaye, *Probability Theory Meets Res Ipsa Loquitur*, 77 MICH. L. REV. 1456 (1979); Kornstein, *A Bayesian Model of Harmless Error*, 5 J. LEGAL STUD. 121 (1976); Lempert, *Modeling Relevance*, 75 MICH. L. REV. 1021 (1977); Penrod & Hastie, *Models of Jury Decision Making: A Critical Review*, 86 PSYCH. BULL. 462 (1979).

⁷ Brilmayer & Kornhauser, *supra* note 1, at 135-48.

⁸ The most elaborate presentation of this thesis can be found in L. COHEN, *THE PROBABLE AND THE PROVABLE* (1977). Brilmayer and Kornhauser are impressed with, and reiterate, some of Cohen's analysis. Brilmayer & Kornhauser, *supra* note 1, at 145 & n.103, 146 n.106. So does Nesson, *supra* note 4, at 1199 n.27. Brilmayer and Kornhauser, and Nesson, also point to the work of a statistician, Glenn Shafer, who posits "belief functions" of which mathematical probability functions are but a special case. G. SHAFER, *A MATHEMATICAL THEORY OF EVIDENCE* (1976). (Despite its title, Shafer's work is not specifically directed to evidence in the legal sense. It is a purely mathematical analysis.) Interestingly, Shafer disputes the adequacy of Cohen's mathematical theory. *Id.* at 223-25.

Only Brilmayer and Kornhauser take as their explicit target the use at trial of Bayes's Theorem. Although they acknowledge the assistance of Shafer in preparing their article, it is not clear that Shafer, as a working statistician, would deny the appropriateness of using ordinary probability calculations to make judgments about facts, at least when the relevant probabilities can be estimated with reasonable precision. As Shafer remarks:

The chances [Shafer distinguishes between "mathematical chance" and "degree of belief"] governing an aleatory experiment may or may not coincide with our degrees of belief about the outcome of the experiment. *If we know the chances, then we will surely adopt them as our degrees of belief.* But if we do not know the chances, then it will be an extraordinary coincidence for our degrees of belief to be equal to them.

Id. at 16 (emphasis added). Cf. Dempster, *Foreword to id.* at vii ("Bayesian inference will always be a basic tool for practical, everyday statistics, if only because questions must be answered and decisions must be taken").

the only real question is the operationally oriented one whether the technique would reduce the number of errors in factfinding, and that the current criticism obscures rather than clarifies the answer to this question. I begin by analyzing some of the more general arguments against using probability axioms in the legal context. In Part I, I reject the contention that discrepancies in the way facts are proved in court and the way mathematical probabilities behave reveal some flaw in the mathematical theory. I argue in Part II that probability theory, subjectively interpreted, can be applied meaningfully to legal factfinding. I turn specifically to the Finkelstein-Fairley proposal in Part III, where I suggest that disputations about the "real" meaning of probability have little bearing on the logical propriety of using accepted probability formulae to explain the significance of legally admissible statistical evidence to a judge or jury. Finally, in Part IV, I consider the claim that there may be more "rational" methods to evaluate statistical evidence than those based on the mathematical theory of probability. I conclude that the epistemological brief against probabilistic models of legal factfinding is far from convincing.⁹

I. IS PROBABILITY THEORY WRONG?

One prong of the recent attack on the use of probability theory in legal factfinding and decisionmaking¹⁰ holds that the mathematical theory of probability¹¹ produces counterintuitive conclusions when applied to juridical proof.¹² The leading exponent of this view is L. Jonathan Cohen, a respected British philosopher of science. Cohen identifies six such "anomalies" which, he asserts, prove that the mathematical theory of probability is fundamentally inapposite to legal factfinding.¹³ Cohen finds, for example, the notion that the

⁹ It is not clear whether Brilmayer and Kornhauser mean to object to the use of probability theory for both heuristic and operational purposes or for just the latter. See note 6 *supra*. Their complaints about the axioms of probability theory, however, would seem to commit them to the view that even the heuristic models may be fundamentally inapposite to understanding the way probabilistic facts should be evaluated at trial. See note 59 *infra*. Nesson, on the other hand, concedes that the established theory of probability is valuable in legal analysis for conceptual purposes. Nesson, *supra* note 4, at 1199 n.27. See also G. SHAFER, *supra* note 8, at 22-25.

¹⁰ Factfinding can be treated as a species of decisionmaking. See, e.g., Lempert, *supra* note 6. See also V. BARNETT, *COMPARATIVE STATISTICAL INFERENCE* 201 (1973).

¹¹ By the "mathematical theory of probability," I mean the axiomatized theory presented in any standard textbook on probability or statistics. See note 28 *infra*.

¹² Brilmayer and Kornhauser similarly suggest that the applicability of the probability axioms is "particularly doubtful with regard to legal reasoning." Brilmayer & Kornhauser, *supra* note 1, at 155.

¹³ "There are certainly at least six anomalies involved in any attempt to construe Anglo-

probability of a defendant's guilt (or liability) plus the probability of his innocence add up to one ($P(G) + P(\text{not-}G) = 1$) to be "paradoxical" when applied to civil cases.¹⁴ To resolve such "paradoxes," Cohen proposes a detailed mathematical framework for calculating the very different kind of "probability" he sees at work in the legal realm.¹⁵ In his system, the probability that an event will happen and the probability that it will not happen would not necessarily add up to one, and neither Bayes's Theorem nor the other features of the conventional theory would necessarily hold.¹⁶

American standards of juridical proof in terms of mathematical probabilities." L. COHEN, *supra* note 8, at 116. For a description of these "anomalies," see *id.* at 49-115.

¹⁴ Cohen's objection is to a probabilistic interpretation of the more-probable-than-not standard generally applicable to civil liability:

Suppose the threshold of proof in civil cases were judicially interpreted as being at the level of a mathematical probability of .501. Would not judges thereby imply acceptance of a system in which the mathematical probability that the unsuccessful litigant deserved to succeed might sometimes be as high as .499? This hardly seems the right spirit in which to administer justice.

Id. at 75.

¹⁵ Cohen's "polycriterial" theory of inductive "support functions" is quite general and is meant to apply in a number of other domains. As explained more fully in note 16 *infra*, this conception of inductive support "has nothing to do with mathematical probability." L. COHEN, *supra* note 8, at 198.

¹⁶ The details of Cohen's "inductive probabilities" are concisely outlined in Schum, *A Review of A Case Against Blaise Pascal and His Heirs*, 77 MICH. L. REV. 446 (1979). In general terms:

Inductive probabilities, although they rest upon developments no less carefully reasoned than mathematical probabilities, have simpler, more primitive, properties. . . . [T]hey cannot be added, subtracted, multiplied, or divided. Although evidence can change the probability of an event, we cannot say, for example, that event A is twice as likely as event B, given relevant evidence. In fact, we can only make what are called *ordinal* or *ordering* relations among inductive probabilities. I can say that, on the evidence, A is more probable than B, but I can neither say how much more probable nor how many times more probable.

Id. at 449-50 (emphasis in original).

To arrive at these inductive probabilities, Cohen defines "support functions" of the form $s(H, E)$, which denote the degree of support for the hypothesis H given certain evidence E . Mathematically speaking (and omitting certain refinements), a "support function" maps ordered pairs of propositions (H, E) on to the first $n+1$ integers. In this nomenclature, $s(H, E) = 0$ means that the hypothesis H is disproved by the evidence acquired from one test that distinguishes between H and some competing hypotheses—that is, which can disprove at least one of the hypotheses invoked to explain the phenomenon in question. Similarly, $s(H, E) = n/n$ means that H is not disproved or eliminated by the evidence obtained from n tests designed to decide among a set of competing hypotheses that includes H . Thus $s(H, E) = n/n$ implies that H has the highest level of inductive support of all the hypotheses under consideration. Between the extremes $s(H, E) = 0$ and $s(H, E) = n/n$, the degree of support for H is $s(H, E) = i/n$, where i is an integer between 0 and n ($0 < i < n$). If $s(H, E) = i/n$, then the evidence consists of the fact that H survived the first i tests but was disproved by the $(i+1)$ th. L. COHEN, *supra* note 8, at 133.

The "inductive probabilities" that ultimately emerge from these support functions be-

Now, there are a number of ways to respond to this sort of argument. One could simply dispute the paradoxical quality of the conundrums Cohen creates. Where, for instance, Cohen insists that to refuse recovery to a civil litigant who had proved that the probability the defendant is liable is .499 "hardly seems the right spirit in which to administer justice,"¹⁷ one might reasonably suggest that such an outcome is perfectly appropriate in the uncertain and imperfect world of litigation.¹⁸ There is, in addition, a different and perhaps more decisive rejoinder to Cohen's ambitious argument. Whatever puzzlement is produced by Cohen's "anomalies" is, I think, attributable to an underlying misconception about the role of probability theory in legal proof. Conventional probability theory does not purport to describe the structure of current legal doctrine. Rather, it prescribes how probabilistic evidence should be handled if one is to make the fewest mistakes in predicting the outcome of uncertain events.¹⁹ Few people would argue, for example, that there

have differently than the familiar mathematical probabilities. For a summary of their distinctive properties, see *id.* at 231-40.

Shafer's "belief functions," like Cohen's "inductive probabilities," do not exhibit the familiar properties of mathematical probabilities, such as the "negation rule" that $P(A) + P(\text{not-}A) = 1$. In essence, this is because Shafer reasons that one can have a certain degree of belief in a proposition A (designated $Bel(A)$), another degree of belief in the negation of A (denoted $Bel(\text{not-}A)$), and a residual quantity of nonbelief. To the extent that such nonbelief is present,

$$Bel(A) + Bel(\text{not-}A) < 1.$$

See G. SHAFER, *supra* note 8, at 43.

¹⁷ See note 14 *supra*.

¹⁸ Cohen's arguments as to the difficulty of quantifying the burden of persuasion in civil cases are considered in more detail in Kaye, *The Paradox of the Gatecrasher and Other Stories*, 1979 ARIZ. ST. L.J. 101. The analysis in Lempert, *supra* note 6, implicitly helps resolve some of the other "anomalies" that trouble Cohen and Brilmayer and Kornhauser, and two of these paradoxes are treated specifically in Wagner, Book Review, 1979 DUKE L.J. 1071. In addition, all the putative paradoxes are examined in Schum, *supra* note 16, at 463-71, and Williams, *The Mathematics of Proof* (pts. 1 & 2), 1979 CRIM. L. REV. 297, 340.

¹⁹ Some theorists have expressed this idea by treating the axioms of probability theory as normative. See, e.g., Carnap, *Inductive Logic and Rational Decisions*, in 1 STUDIES IN INDUCTIVE LOGIC AND PROBABILITY 5, 7-9 (R. Carnap & R. Jeffrey eds. 1971). Of course, this characterization of the theory merely moves the debate to another plane. Since the truth of a theory cannot be a matter of definition, it must be shown that the axioms that state how probabilities should be handled are preferable to other formulae. This can be done by constructing a pragmatic or positivist argument for the existing mathematical theory. See text and notes at notes 71-72 *infra*. Alternatively, one can make an intuitionist argument. An interesting version of the latter approach is to appeal to the axioms of rational choice familiar to decision theorists. These axioms, in effect, provide a formal definition of rational behavior. One such axiom, for instance, requires that preferences be transitive. It states that if outcome A is preferred to outcome B , and B is preferred to C , then A is also preferred to C . See, e.g., C. COOMBS, R. DAWES & A. TVERSKY, *MATHEMATICAL PSYCHOLOGY* 122-26 (1970); J. VON NEUMANN & O. MORGENTERN, *THEORY OF GAMES AND ECONOMIC BEHAVIOR* 26 (3d ed. 1953). From such intuitively attractive postulates, the usual axioms of probability theory can be deduced.

is some more accurate way to predict whether the first two cards dealt from a well shuffled deck will be red than to calculate the probability as $(\frac{1}{2}) \times (\frac{1}{2}) = \frac{1}{4}$.²⁰ The same theory should be no less accurate for making "postdictions," or findings of fact. If we wish to decide whether two cards which have already been dealt face down are red, the most accurate basis for coming to a conclusion (without peeking) is to rely on the same probability calculation: $(\frac{1}{2}) \times (\frac{1}{2}) = \frac{1}{4}$.²¹

That the present rules of evidence and procedure do not invariably appear to promote this ideal of accuracy does not impugn the accuracy of probability theory as a tool in legal factfinding. In some cases, the law may exclude mathematical probability calculations precisely to encourage the production of more individualized and revealing evidence—to permit peeking at the cards, as it were.²²

See, e.g., T. FERGUSON, *MATHEMATICAL STATISTICS: A DECISION THEORETIC APPROACH* (1967); L. SAVAGE, *THE FOUNDATIONS OF STATISTICS* (2d ed. 1972); notes 31 & 32 *infra*.

To be sure, the definition accepted in decision theory may not provide a philosophically satisfactory definition of rationality. *See, e.g.,* Kaye, *Playing Games with Justice: Rawls and the Maximin Rule*, 6 *SOC. THEORY & PRAC.* (forthcoming). Nevertheless, there seems to be no reason in principle why ideal legal factfinding cannot be captured by a model in which the decisionmaker seeks to maximize expected utility, where the utility function is sufficiently multidimensional, and the familiar sorts of probabilities are used. *See generally* R. KEENEY & H. RAIFFA, *DECISIONS WITH MULTIPLE OBJECTIVES: PREFERENCES AND VALUE TRADEOFFS* (1977); *UTILITY, PROBABILITY AND HUMAN DECISION MAKING* (D. Wendt & C. Vlek eds. 1975). In this regard, Brilmayer and Kornhauser's suggestion that there may be "rational" strategies available to jurors other than "maximizing expected value," Brilmayer & Kornhauser, *supra* note 1, at 142, is difficult to understand in the absence of some explanation of what they mean by "value." Presumably, they have in mind something like the "satisficing" principle as a replacement for utility maximization. *See, e.g.,* *DECISION AND ORGANIZATION* (C. McGuire & R. Radner eds. 1972); H. SIMON, *MODELS OF MAN* 246-60, 261-73 (1957). Still, it is hard to see why "bounded rationality," to use the economist's jargon, should cause us to alter the mathematical structure of existing probability theory.

²⁰ This calculation assumes that the first card drawn is reinserted and the deck reshuffled. Without such replacement, the probability of two red cards being dealt would be $(\frac{1}{2}) (25/51) = .245$

²¹ Courtroom decisions as to what has occurred are analogous to guessing what happened when the cards were dealt face down, and accuracy in making predictions thus translates into accuracy in factfinding. It might seem that a logical positivist would be troubled by this description, and would prefer to say that when the cards are dealt face down, one is not making a statement about what has already occurred but rather about what will occur when the cards are turned over. *Cf.,* A. RAPOPORT, *OPERATIONAL PHILOSOPHY* 61 (1953) (predictivist perspective on propositions of historical fact). For present purposes, this is more a matter of phraseology than substance. Nevertheless, it points up the fact that after the jury decides, we do not test its "postdictions." In principle, we could. *See* text and note at note 41 *infra*. Hence, to state the analogy more precisely, we can imagine that the cards have been dealt, that one is constrained by reasons of social policy from turning them over, and that one is compelled to decide whether they are both red. In these circumstances, if all we desire is the more accurate guess, we should not hesitate to follow the dictates of probability theory (or to advise whoever is making the decision of the result of our probability calculation).

²² Brilmayer and Kornhauser overlook this point when they note that "a probability

Consider, for instance, Tribe's argument that the facts that the plaintiff was negligently run over by a blue bus and that the defendant operates four-fifths of all the blue busses in town should not, without more, support a verdict for the plaintiff.²³ A rule of law denying the plaintiff recovery in such a case would not establish that probabilistic reasoning is inapplicable or that jurors would not be well advised to find facts according to their best estimates of the relevant probabilities. It would merely reflect the policy that where individualized evidence is likely to be available—evidence which would typically permit better estimates of the probabilities than can be had from the background statistics alone—plaintiffs should be forced to produce it. In the long run, fewer mistaken verdicts should result under this rule of law.

In other cases, the fact that legal doctrine does not seem to conform to probability calculations may be a cause for concern and reform.²⁴ Or it may be a consequence of socially important policies other than accuracy in factfinding. Why is it, for example, that we make no effort to explain the meaning of "proof beyond a reasonable doubt" in precise quantitative terms? Surely it is not some defect in probability theory that restrains us from instructing jurors that they should convict so long as they are, say, at least ninety-five percent certain that the defendant is guilty. A much simpler explanation is that we would prefer not to advertise the fact that we are willing to sacrifice one innocent person in order to secure the conviction of nineteen guilty ones. Some things, it has been argued, are best left unsaid.²⁵ Although Cohen recognizes that the legal process is concerned with more than accuracy in factfinding, he insists that the mathematical theory of probability must take into account these other values in describing how the probabilities of disputed

figure that reflects the extent to which a given body of evidence supports an inference, but not the sufficiency of the evidence itself, cannot adequately capture all the elements of legal proof." Brilmayer & Kornhauser, *supra* note 1, at 145. For a more detailed analysis of the "sufficiency of the evidence" problem using the accepted mathematical theory of probability, see, for example, Kaye, Book Review, 89 YALE L.J. 601 (1980). See also Ellman & Kaye, *supra* note 5.

²³ Tribe, *supra* note 3, at 1340, 1341 & n.37.

²⁴ See, e.g., Wagner, *supra* note 18, at 1076. Cohen is aware of this response to his analysis but dismisses it as implausible. He writes:

This objection shows a certain contempt for the natural intelligence of lawyers. Their logicomathematical intuitions must be surprisingly widely fallacious if the objection is correct. Nor is it at all clear why their pattern of error should be so self-consistent and uniform.

L. COHEN, *supra* note 8, at 119-20.

²⁵ See Tribe, *supra* note 3, at 1372-73. See generally G. CALABRESI & P. BOBBITT, TRAGIC CHOICES (1977).

facts should be calculated. It seems more sensible (though less original) to say that the probabilities behave as prescribed by the conventional probability axioms, but that the law accepts imprecision and inaccuracy because other values are at work.²⁶ The putative anomalies Cohen depicts thus do not compel us to abandon the established theory of probability at the courthouse door.

II. ARE THE SUBJECTIVISTS RIGHT?

A second prong of the criticism of probability theory in the context of legal factfinding does not claim that there is something special about legal proof that makes the axioms of probability theory inapplicable. Rather, it asserts that the axioms break down when applied to degrees of belief about the likelihood of uncertain events. This argument, or some variant of it, is the dominant point raised by Brilmayer and Kornhauser in their attack on the use of Bayes's formula as an explicit tool in legal factfinding.²⁷

To see why this view is not persuasive, however, one must understand the historically competing interpretations of probability and their connection with the probability axioms. To a mathematician, probabilities are numbers that obey a few axioms.²⁸ Where the

²⁶ See, e.g., Underwood, *Law and the Crystal Ball: Predicting Behavior with Statistical Inference and Individualized Judgment*, 88 YALE L.J. 1408 (1979); Tribe, *supra* note 3, at 1376.

²⁷ Brilmayer & Kornhauser, *supra* note 1, at 138, 141-45. The authors argue particularly that the axiom of additivity, *see* note 28 *infra*, does not apply to degrees of belief. Brilmayer & Kornhauser, *supra* note 1, at 141-42, 141 n.91. The argument seems to carry the most weight as applied to unique events, those for which probability estimates cannot be made by observing repeated outcomes.

²⁸ The following tongue-in-cheek comment expresses this viewpoint cogently: "[T]here is no problem about probability: it is simply a non-negative, additive set function, whose maximum value is unity." H. KYBURG & H. SMOKLER, *STUDIES IN SUBJECTIVE PROBABILITY* 3 (1964). The axiomatic formulation of probability, first presented rigorously in the 1930s, proceeds by assigning to each event *A* in a simple space *S* of all possible outcomes a number *P*(*A*), which indicates the probability that *A* will occur. To qualify as a probability, this number must satisfy only three axioms. The first axiom states that the probability of every event must be nonnegative: *P*(*A*) ≥ 0. The second axiom states that if an event is certain to occur, then the probability of that event is one: *P*(*S*) = 1. The third axiom provides that if events have nothing in common with one another (they are disjoint, or mutually exclusive), then the probability that at least one of these events will occur is the sum of the probabilities

that each will occur: $P(\bigcup_{i=1}^{\infty} A_i) = P(A_1) + P(A_2) + \dots$. See, e.g., M. DEGROOT, *PROBABILITY*

AND STATISTICS 12-13 (1975). A weaker form of the third axiom, stating that the probabilities of a finite set of disjoint events are additive, is also sufficient, but makes for less elegant derivations of many probability theorems. For more careful, but progressively less accessible presentations of the probability axioms, see, for example, J. KINGMAN & S. TAYLOR, *INTRODUCTION TO MEASURE AND PROBABILITY* 261-84 (1966); J. NEVEU, *MATHEMATICAL FOUNDATIONS OF THE CALCULUS OF PROBABILITY* 2-25 (A. Feinstein trans. 1965); Jeffrey, *Probability Measures and Integrals*, in 1 *STUDIES IN INDUCTIVE LOGIC AND PROBABILITY*, *supra* note 19, at 167-221. From

numbers come from is of no mathematical concern. The mathematical theory only prescribes what can be done in the way of combining and manipulating these probabilities, once the numbers are somehow obtained. Some of the mathematicians who made pioneering contributions to probability and statistics thought of probabilities as objectively determined by the number of mutually exclusive, collectively exhaustive, and "equally likely" outcomes associated with an experiment. For example, according to Laplace's "principle of insufficient reason," if we know only that a coin having two sides is about to be tossed, we should assign to the outcome of heads (or for that matter, tails) the probability one-half.²⁹ Most contemporary theorists, however, reject this a priori, or "logical" conception of probability, and many favor an empirical or "frequentist" interpretation.³⁰ Under this view, the probability associated with a given outcome is defined in terms of the relative frequency with which such an outcome will result on repeated trials.

It bears repeating, however, that neither of these "objective" definitions is presupposed by the axiomatized mathematical theory. In fact, there is every reason to believe that numbers arrived at by other processes will also qualify as probabilities and obey all the theorems and postulates of the mathematical theory. Various related approaches to defining such "subjective" probabilities have been exhibited over the last half-century.³¹ All rely on the idea of a

the three axioms, all the rules (including Bayes's Theorem) presented in all textbooks on probability and statistics inexorably follow. Brilmayer & Kornhauser, *supra* note 1, at 141, question whether the third axiom holds for "subjective" probabilities.

²⁹ For descriptions of pre-twentieth century views as to the meaning of probability, see, for example, the authorities cited in Brilmayer & Kornhauser, *supra* note 1, at 137 n.75. The description of the competing interpretations given in the text is, for the sake of brevity, badly oversimplified, but a more careful summary is available in V. BARNETT, *supra* note 10, at 62-89. For a still more elaborate presentation, see, for example, T. FINE, *THEORIES OF PROBABILITY* (1973). See also Hawkins, *Probability, Information and Inference*, in *NEW DEVELOPMENTS IN THE APPLICATIONS OF BAYESIAN METHODS* 11 (A. Aykac & C. Brumat eds. 1977).

³⁰ See Carnap, *The Two Concepts of Probability*, in *READINGS IN THE PHILOSOPHY OF SCIENCE* 438, 450 (H. Feigl & M. Brodbeck eds. 1953).

³¹ Informal treatments of subjective probability date back to J. BERNOULLI, *ARS CONJECTANDI* (Basel 1713). Formal theories were advanced independently and almost simultaneously by Frank Ramsey and Bruno de Finetti in the period 1925-1935. For example, in 1926 Ramsey used "ethically neutral propositions" to show how cardinal scales of probability and utility could simultaneously be derived from a set of coherent preference rankings. See F. RAMSEY, *Truth and Probability*, in *THE FOUNDATIONS OF MATHEMATICS AND OTHER LOGICAL ESSAYS* 156 (1931). Ramsey's derivation is described in a simplified way in R. JEFFREY, *THE LOGIC OF DECISION* 31-43 (1965), and his method for defining utilities was rediscovered in the 1940s by J. VON NEUMANN & O. MORGENTHAU, *supra* note 19, at 15-31. For other developments along these lines, see, for example, J. PRATT, H. RAIFFA & R. SCHLAIFER, *INTRODUCTION TO STATISTICAL DECISION THEORY* (1965); L. SAVAGE, *supra* note 19, at 6-104; Anscombe & Aumann, *A Definition of Subjective Probability*, 34 *ANNALS MATHEMATICAL STATISTICS* 199

coherent ordering of preferences in choosing among uncertain outcomes. Where subjective probabilities are defined in terms of the odds a person would be willing to accept in betting on the outcome of an event, they can be shown to obey all the rules of probability theory,³² at least where the judgments satisfy certain conditions of consistency.³³ For instance, a person who is indifferent between taking either side of a bet that the defendant is the murderer when offered two-to-one odds has a subjective probability of one-third for the likelihood of the defendant's guilt. Alternatively, subjective probability can be defined in terms of lotteries rather than betting odds. Under this analysis, having the subjective probability of one-third for the likelihood of defendant's guilt is tantamount to being indifferent between two lotteries: (a) one that offers a prize of, say,

(1963); de Finetti, *Foresight: Its Logical Laws, Its Subjective Sources*, in *STUDIES IN SUBJECTIVE PROBABILITY* 93 (H. Kyburg & H. Smokler eds. 1964); Giles, *A Logic for Subjective Belief*, in *1 FOUNDATIONS OF PROBABILITY THEORY, STATISTICAL INFERENCE, AND STATISTICAL THEORIES OF SCIENCE* 41 (W. Harper & C. Hooker eds. 1976); Harper, *Rational Belief Change, Popper Functions and Counterfactuals*, in *id.* at 73; May & Harper, *Toward An Optimization Procedure for Applying Minimum Change Principles in Probability Kinematics*, in *id.* at 137; Shimony, *Coherence and the Axioms of Confirmation*, 20 *J. SYMBOLIC LOGIC* 1 (1955); Smith, *Personal Probability and Statistical Analysis*, 128 *J. ROYAL STATISTICAL SOC'Y (SERIES A)* 469 (1965). An engaging and compact sketch of the historical development of some subjective probability theories can be found in H. RAIFFA, *DECISION ANALYSIS* 273-78 (1968).

³² See, e.g., Freedman & Purves, *Bayes Method for Bookies*, 40 *ANNALS MATHEMATICAL STATISTICS* 1177 (1969); Kemeny, *Fair Bets and Inductive Probabilities*, 20 *J. SYMBOLIC LOGIC* 263 (1955); Lehman, *On Confirmation and Rational Betting*, 20 *J. SYMBOLIC LOGIC* 251 (1955); Smith, *Consistency in Statistical Inference and Decision*, 23 *J. ROYAL STATISTICAL SOC'Y (SERIES B)* 1 (1961). Although rigorous proofs of this result are mathematically involved, the underlying logic is easily illustrated. Consider the plight of someone whose subjective probabilities do not have the requisite mathematical properties. Suppose, for instance, that they do not obey the familiar rule that $P(X) + P(\text{not-}X) = 1$. For concreteness, let X stand for the outcome that tomorrow will be a rainy day, and let us suppose that the individual asserts that at odds of three-to-one in favor of rain he is equally willing to take either side of a \$10 bet on this outcome, X . His subjective probability that tomorrow will be rainy is then $1/(3+1) = 1/4$. He further asserts that he believes that the same odds are fair in betting \$10 on $\text{not-}X$, the outcome that it will not rain tomorrow. This produces a subjective probability of $1/4$ in favor of $\text{not-}X$, in violation of the rule that these probabilities must sum to one. It also puts him in the position of being certain to lose \$20. This is so because he is indifferent between either side of each bet. We therefore ask him to bet \$10 against rain at the three-to-one odds in favor of rain (the short side of the bet on X). We also ask him to bet another \$10 that it will not rain at these same odds (the long side of the bet on $\text{not-}X$). If it rains, he loses \$30 on the first bet and wins only \$10 on the second bet (since $\text{not-}X$ has not occurred). Similarly, if it does not rain, he wins \$10 on the first bet but loses \$30 on the second (since $\text{not-}X$ has occurred). Unless this person wishes to serve as a "money pump," he had best revise his odds in such a way that the subjective probabilities they generate for X and $\text{not-}X$ add up to one. Brilmayer and Kornhauser recognize that such logic supports the use of the conventional probability axioms with subjective probabilities "if one conceives of the decisionmaking process as a gamble." Brilmayer & Kornhauser, *supra* note 1, at 141-42 & n.93.

³³ See, e.g., V. BARNETT, *supra* note 10, at 82; R. JEFFREY, *supra* note 31, at 42; note 19 *supra*.

\$100 with probability one-third and, say, \$10 with probability two-thirds, and (b) one that gives the \$100 prize if the defendant is guilty and the \$10 prize if the defendant is not guilty.³⁴

Recognizing that subjectively interpreted probabilities are logically compatible with the mathematical structure of probability theory,³⁵ those who challenge their use are reduced to arguing that such probabilities cannot be meaningfully estimated in some circumstances and, in particular, in litigation. In part, they are right. One can construct a few intractable cases in which the choice of a subjective probability is either impossible or must be totally arbitrary. A newborn child, although not entirely without information about his environment, would have difficulty estimating probabilities. A person deprived of all knowledge of physics, chemistry, biology, and astronomy would be puzzled if asked the probability that there are living beings on a planet orbiting the star Sirius.³⁶ But these fanciful examples put forward by Brilmayer and Kornhauser³⁷

³⁴ Tribe, *supra* note 3, at 1346-48, describes this approach in more detail. The size of the prizes is unimportant, as long as one prize is preferred to the other. (Where one is indifferent between the prizes, one is also indifferent as between the two lotteries regardless of the probabilities involved.) Yet another approach to defining subjective probability relies on the device of bidding for a "reference contract." See R. WINKLER, INTRODUCTION TO BAYESIAN INFERENCE AND DECISION 23 (1972).

³⁵ Most "subjectivists" believe that the "objective" interpretations of probability are subsumed within the subjective conception. Thus, in their view, one can (and in some cases should) accept a relative frequency number as the appropriate "subjective" probability. See, e.g., Y. CHOU, STATISTICAL ANALYSIS 417-18 (2d ed. 1975); H. PAZER & H. SWANSON, MODERN METHODS FOR STATISTICAL ANALYSIS 17-19 (1972). Finkelstein seems to share this view, although his language is ambiguous. See M. FINKELSTEIN, *supra* note 6, at 64. When he makes this point, however, Brilmayer and Kornhauser accuse him of "unfamiliarity with, or disregard of, an enormous body of literature on the subject." Brilmayer & Kornhauser, *supra* note 1, at 140.

³⁶ For reasons explained in, for example, Carnap, *supra* note 30, at 444; J. KEYNES, A TREATISE ON PROBABILITY 42-51 (1921); G. SHAFER, *supra* note 8, at 23-24; E. NAGEL, PRINCIPLES OF THE THEORY OF PROBABILITY 46-47 (1939), this example may expose a logical flaw in the Laplacean principle of insufficient reason. As such, it may undermine the classical, objective, a priori interpretation of probability. See text and note at note 29 *supra*. Unless there are realistic situations in which one must resort to the Laplacean principle in making subjective estimates, however, it is no argument against subjective probabilities.

³⁷ Brilmayer & Kornhauser, *supra* note 1, at 143. The authors do not actually speak of infants making probability estimates. Instead, they state:

It seems more plausible that there are some questions for which we have no empirical data. Unless the subjectivists would argue that we are born with information bearing on every question that can be phrased in the English language, there must be a point at which we received our first piece of information about any given proposition. Before that point we had none.

Id. at 143 n.96. It is, however, difficult to conceive of any question that can be answered by scientific investigation for which a person of more than a few years of age who has not been reared in a state of total sensory deprivation would have no "empirical data." The data may be extremely limited, highly peripheral, or grossly faulty, but there will almost always be

have little bearing on the estimation of probabilities in the trial process. Jurors are not neonates, and ample evidence should ordinarily be available to justify meaningful, if imperfect, probability estimates.³⁸ If such minimal evidence is not present, then the plaintiff has not met his burden of production, and the case should be resolved against him.

It is true, of course, that jurors do not make book on disputed factual issues. But those who deny that subjective probabilities are meaningful in modeling or aiding legal factfinding must show that the various methods of defining subjective probabilities are not available, even in principle, in the trial context. For instance, one can imagine quizzing a juror as to the odds governing any particular issue. Thus, to revert to an earlier illustration, one can conceive of demanding that the juror state the odds at which he would be willing to take either side of a bet that the defendant is the murderer, as alleged. Nevertheless, both Cohen and Brilmayer and Kornhauser contend that this is not enough.³⁹ Cohen, whose thinking on this point is better developed, condemns such a procedure as "grossly fallacious" for two reasons. First, he asserts, "bets must be settleable. In each case the outcome must be knowable otherwise than from the data on which the odds themselves are based. When the horse-race is finally run, the winner is photographed as he passes the winning-post."⁴⁰ There is nothing outrageous, however, about considering what odds would be acceptable even if one has no intention of carrying through on the bets. And, bets about facts contested at trial could, in theory at least, be settled afterwards if society desired to expend the resources or use sufficiently brutal methods to gather further evidence.⁴¹

some information on which to base a subjective probability.

This conclusion is not affected by de Finetti's remark, cited by Brilmayer and Kornhauser, that "[a]ccording to the subjective view, no problem can be correctly stated in statistics without an evaluation of the initial probabilities." *Id.* (quoting B. DE FINETTI, *PROBABILITY, INDUCTION AND STATISTICS* 143 (1972)). We are concerned at this point not with the question whether Bayesian techniques of inference are suitable for solving every statistical problem. That is indeed a controversial claim. The issue, rather, is whether subjective probabilities suitable for manipulation according to Bayes's formula and other rules of mathematical probability can be stated. As the works cited in notes 31 & 32 *supra* show, there can be little serious dispute over this contention. Perhaps Brilmayer and Kornhauser have been led astray by the fact that it is common to speak of subjective probabilities as "Bayesian" probabilities. See note 57 *infra*.

³⁸ But see Gerjuoy, *supra* note 6, at 23.

³⁹ L. COHEN, *supra* note 8, at 90; Brilmayer & Kornhauser, *supra* note 1, at 142 & n.93.

⁴⁰ L. COHEN, *supra* note 8, at 90.

⁴¹ To be sure, even after administrations of sodium pentothal, polygraph tests, torture, or what have you, there might still be some residual uncertainty about a defendant's guilt. But the same could be said of a horse race. The picture of the supposed winner may have

Cohen's second point is that the odds a juror would accept depend in part on the magnitude of the bet.⁴² At first glance, this observation seems troublesome because it suggests that persons who are not risk neutral will be unable to settle on a single probability characterizing their degree of belief in the outcome of a specific event.⁴³ The problem, however, is not as serious as it might seem. To begin with, it pertains to but one of the several successful formulations of subjective probability. Probabilities ascertained via the lottery method, for example, are immune from this attack, for they are independent of the size of the prizes involved.⁴⁴ Moreover, even under the betting-odds approach, the probabilities are very nearly constant so long as the bets are small compared to the decision-maker's total wealth.⁴⁵ In addition, I would conjecture that dependence of the odds on the magnitude of the bet can be circumvented by leaving the amount of the wager indeterminate. In other words, we elicit a person's subjective probability of an outcome by requiring him to state the odds he considers fair. To make certain that the resulting assessment is sincere and fair, we inform him that once he provides these odds, he will have to take whatever side of the bet we select for him, and that, furthermore, he will have to wager whatever amount we prescribe.

been taken from an improper angle, it may have been doctored, or the laws of optics may have been suspended by a sinister force during the exposure. It is all a matter of degree.

⁴² L. COHEN, *supra* note 8, at 90.

⁴³ In economic theory an individual is said to be risk neutral if, roughly speaking, he values a gain of X dollars exactly as much as he abhors a loss of this magnitude. For such an individual, the utility of money is linear, and maximizing utility is equivalent to maximizing income. See, e.g., R. WINKLER, *supra* note 34, at 255. A decisionmaker who is not risk neutral, however, cannot maximize expected utility by maximizing expected monetary return. If he is risk averse, for instance, and the amount he must wager is large, he will select more favorable odds (to overcome the risk of a large loss in income) than if the bet is trifling because he finds the loss of a large amount of money disproportionately more discomforting than the loss of a small amount. See generally B. LINDGREN, *ELEMENTS OF DECISION THEORY* 56-60 (1971); D. LINDLEY, *MAKING DECISIONS* 81-87 (1971); W. NICHOLSON, *MICROECONOMIC THEORY* 153-56 (1972).

⁴⁴ See text and note at note 34 *supra*.

⁴⁵ Within this region, any continuous utility function is approximately linear. Maximization of expected utility is then equivalent to maximization of expected monetary return. To avoid an expected negative return, the decisionmaker will select the one set of odds that, he believes, makes the bet statistically fair. E.g., R. WINKLER, *supra* note 34, at 24. A statistically fair bet is one whose expected payoff in dollar terms is zero. E.g., L. CHAO, *STATISTICS* 118 (1974); B. LINDGREN, *supra* note 43, at 54. For example, betting X dollars at odds of three-to-one that a heart will be drawn from a well-shuffled deck of playing cards is fair, since the expected winnings are given by (a) the probability that the card will be a heart ($p = 1/4$) times the winnings if a heart appears ($3X$) plus (b) the probability that the card will not be a heart ($1 - p = 3/4$) times the winnings if a heart is not turned up ($-X$):

$$(p)(3X) + (1 - p)(-X) = 1/4(3X) - 3/4(X) = 0.$$

Of course, the assumption that the decisionmaker must take either side of the gamble—that he must act as a bookmaker—can be questioned. Indeed, this is a standard objection in the literature on subjective probability.⁴⁶ Brilmayer and Kornhauser believe this argument against subjective probabilities is especially powerful in the trial setting because the decisionmaker may “consider gathering more information.”⁴⁷ Yet the opposite seems to be the case. The involuntary-bookmaker logic is particularly apt in the context of trials. Disputed facts must be resolved then and there on the basis of incomplete and imperfect information, and the conscientious decisionmaker should be able to take either side of a bet at the odds corresponding to his probability estimate that he has made the correct decision. A juror who takes his job seriously can hardly follow the Brilmayer-Kornhauser strategy and “consider gathering more information about the likelihood of *A* and *not-A*, or make a decision based on some criterion distinct from maximizing expected value.”⁴⁸ The numbers jurors might supply if asked to quantify their personal beliefs can therefore be viewed as approximations of the theoretically satisfactory probabilities. In this way, the probabilities at work in legal factfinding can be given a conceptually meaningful subjective interpretation.

III. IS SUBJECTIVISM NECESSARY?

So far I have discussed the applicability of probability theory to the trial process in a very general way. I have tried to show that

⁴⁶ As one well-known critic of subjective probability puts it,

That I will not allow a book to be made against me is entailed by my preference ranking for money (that I'd prefer not to suffer a certain loss), and has nothing to do with my degree of belief. Having accepted a bet at odds of $(1-p):p$ on *S*, of course I won't accept a bet at odds of $(1-q):q$ on *not-S* [where $p \neq (1-q)$]. This is not to say that if I had *not* accepted the first bet, I wouldn't accept the second bet. The hidden assumption required is that the behavioral counterpart of a distribution of beliefs over an algebra of sentences is a willingness to make book on them: to accept bets of *any* size, on *either* side, with respect to *any* combination of sentences. . . . [This] puts the person whose beliefs we are discussing in the position of the bookmaker; . . . a bookmaker had better be jolly sure that the odds he posts satisfy the axioms of the probability calculus.

H. KYBURG, *THE LOGICAL FOUNDATIONS OF STATISTICAL INFERENCE* 95-96 (1974).

⁴⁷ Brilmayer & Kornhauser, *supra* note 1, at 142.

⁴⁸ *Id.* Admittedly, there are some unusual cases in which a court may choose correctly not to let a case reach the jury even where the probability given the evidence seems, at first blush, to exceed one-half or whatever number would justify a finding in favor of the proponent of the evidence. This “rational strategy” of “insufficient evidence,” *id.*, does not indicate any defect in probability theory. On the contrary, it can be explained by a Bayesian model, *see* Kaye, *supra* note 6, at 1475-81; Kaye, *supra* note 18, at 106-08, and merely reflects a legal policy designed to minimize errors in the long run by encouraging plaintiffs to produce better evidence about the matter in dispute. *See* text and note at note 22 *supra*.

there is nothing paradoxical or counterintuitive in thinking of legal factfinding in terms of the mathematics of probability and that the subjective definition of probability provides a meaningful interpretation of the probabilities involved in the legal process. It is helpful, however, to consider the Finkelstein-Fairley proposal for using probability theory in litigation in more detail. I do so not to argue in favor of actually implementing the proposal, but rather to ask whether the Finkelstein-Fairley suggestion requires us to interpret the probabilities involved in a subjective way. Finkelstein, who is a subjectivist, assumes that it does, and, it does not occur to Brilmayer and Kornhauser to question this assumption.⁴⁹ In this section, however, I shall suggest, with some diffidence, that recourse to the subjective theory is not essential, and that therefore the long-standing and esoteric debate among logicians, philosophers, and statisticians about subjective probabilities and the probability axioms is, in the end, irrelevant to evaluating the Finkelstein-Fairley proposal.

The problem Finkelstein and Fairley seek to solve with the machinery of Bayes's formula is as much psychological as mathematical. It is a problem that occurs in every case in which one party seeks to introduce into evidence a quantified statement of some crucial probability. For instance, a robbery has been committed by a couple, and the description of the robbers is clear and definite: they are said to be a young, white woman with blond hair and a ponytail, and a black man with a beard and a moustache who drives a yellow convertible. A couple meeting that description is apprehended, and the prosecutor offers a carefully conducted survey to show that no more than one couple in a million fits the given description.⁵⁰ Or, to take a more common example, in a paternity action the plaintiff offers serologic test results to establish that the defendant and the child both have a set of genes found in only one out of a hundred males in the general population.⁵¹ Contrary to what one might think at first glance, these numbers do not imply that the probability that the couple apprehended committed the robbery is $1 - 10^{-6} = .999999$ or that the probability that the defendant is the father is $1 - 10^{-2} = .99$.⁵² But what, then, should a jury make of these

⁴⁹ See note 35 *supra*.

⁵⁰ Cf. *People v. Collins*, 68 Cal. 2d 319, 438 P.2d 33, 66 Cal. Rptr. 497 (1968) (similar probability calculation made without any survey—careful or otherwise).

⁵¹ Cf. *Cramer v. Morrison*, 88 Cal. App. 3d 873, 153 Cal. Rptr. 865 (1979) (similar statistic derived from HLA haplotyping).

⁵² See, e.g., *People v. Collins*, 68 Cal. 2d 319, 438 P.2d 33, 66 Cal. Rptr. 497 (1968); Tribe, *supra* note 3, at 1336 & n.23; Ellman & Kaye, *supra* note 5; Fairley & Mosteller, *A Conversa-*

relative-frequency figures?

Finkelstein and Fairley suggest one answer: use Bayes's formula to convey to the jurors the probative force of the quantitative evidence. They illustrate their proposal by a graphic example.⁵³ Suppose a woman's body is found in a ditch. There is evidence that the deceased had a violent quarrel with her boyfriend the night before. He is known to have struck her on other occasions. Investigators find the murder weapon, a knife, whose handle bears a latent palm print similar to the defendant's. The information on the print is limited so that an expert can say only that such prints appear in no more than one case in a thousand. Finkelstein and Fairley believe that the jurors will be aided in accurately assessing the total probative value of the evidence if they are first shown a chart depicting, in numerical terms, how much the probability that the defendant wielded the murder weapon is enhanced by the discovery of the prints. The mathematical tool for devising such a chart is Bayes's formula. Let X denote the event that the defendant stabbed the deceased; let $P(X)$ stand for the probability of this event without regard to the palm print evidence; let $P(X/E)$ represent the probability that the defendant stabbed the deceased, taking into account the statistical evidence; and let $P(E/X)$ be the probability that the palm print would have matched, given that the defendant had actually stabbed the deceased with the knife. Then Bayes's formula relates $P(X/E)$ (the "posterior probability") to $P(X)$ (the "prior probability") as follows:⁵⁴

$$P(X/E) = \frac{1}{(1-f) + f/P(X)} \quad (1)$$

where

$$f = \frac{P(E/\text{not-}X)}{P(E/X)} . \quad (2)$$

tion About Collins, 41 U. CHI. L. REV. 242 (1974); Charrow & Smith, *A Conversation About "A Conversation About Collins,"* 64 GEO. L.J. 669 (1976).

⁵³ Finkelstein & Fairley, *supra* note 2, at 496.

⁵⁴ Any introductory text on mathematical probability includes a derivation of Bayes's Theorem. Derivations can also be found in, for example, M. FINKELSTEIN, *supra* note 6, at 87-98; Gerjuoy, *supra* note 6, at 9-12; Kaye, *supra* note 6; Lempert, *supra* note 6, at 1022-23, 1023 & n.12; Tribe, *supra* note 3, at 1351-54. For the more general formulation of Bayes's formula, see, for example, G. BOX & G. TIAO, *BAYESIAN INFERENCE IN STATISTICAL ANALYSIS* (1972); Y. CHOU, *supra* note 35, at 414; R. WINKLER, *supra* note 34, at 76, 144 (1972).

It is more common to write Bayes's formula in a slightly different form:

$$P(X/E) = \frac{P(X)P(E/X)}{P(X)P(E/X) + [1-P(X)]P(E/\text{not-}X)} .$$

These equations thus show how the prior probability should be modified or "conditioned" by the new item of evidence E . For example, if there is no doubt that the prints would match if the defendant used the knife, then $P(E/X) = 1$. At the same time, the probability that the prints would match given that someone else used the knife is (neglecting the possibility of a frameup) given by $P(E/\text{not-}X) = 1/1000$. Thus, the ratio f defined by equation (2) is $1/1000$, and equation (1) reduces to:

$$P(X/E) = \frac{1}{.999 + .001/P(X)}. \quad (3)$$

If a juror believed, on the basis of all the nonquantified evidence—the violent argument, the defendant's murderous propensities, and so forth—that the prior probability, $P(X)$, is one-half, then the equations reveal that the numerical palm print evidence has raised the probability that the defendant used the knife to the posterior probability

$$P(X/E) = \frac{1}{.999 + .001/.5} = .999.$$

By applying equation (3) in this fashion, the effect of the statistical evidence on any given prior probability can be calculated, and a chart displaying the impact of the quantitative data across a broad range of prior probabilities can be prepared. This modest use of Bayes's formula would not require a juror to commit himself to any specific prior probability. He could scan the chart to see how powerful the quantitative evidence is. In short, Finkelstein and Fairley suggest using the Bayesian chart as a pedagogical device, which would give the jurors some guidance in assessing the significance of the statistical evidence.

As I remarked at the outset, whether implementation of this modest chart approach would be workable and effective in practice is highly debatable.⁵⁵ But the more abstract epistemological objection raised by Brilmayer and Kornhauser has little bearing on this practical question. Contrary to what Brilmayer and Kornhauser seem to imply,⁵⁶ the objection to using subjective prior probabilities

Dividing numerator and denominator of the right hand side by $P(X)P(E/X)$ gives the version presented here. The f defined by equation (2) is thus the reciprocal of the "likelihood ratio" employed by Lempert, *supra* note 6, at 1025.

⁵⁵ See text and notes at notes 3-5 *supra*.

⁵⁶ Brilmayer and Kornhauser seem impressed with the fact that "Bayesians" have not yet written many elementary textbooks. Brilmayer & Kornhauser, *supra* note 1, at 135 n.68, 148 n.115. But see L. PHILLIPS, *BAYESIAN STATISTICS FOR SOCIAL SCIENTISTS* (1973); S. SCHMITT, *MEASURING UNCERTAINTY: AN ELEMENTARY INTRODUCTION TO BAYESIAN STATISTICS* (1969); R.

does not deny that Bayes's Theorem is mathematically correct.⁵⁷ It

WINKLER, *supra* note 34. They go so far as to quote, in italics, Nagel's 1939 observation that "few writers today take [Bayes's Theorem] seriously as a means for determining the probability of a given hypothesis on the basis of given evidence." *Id.* at 148 n.115 (quoting E. NAGEL, *PRINCIPLES OF THE THEORY OF PROBABILITY* 31 (1939)). Surely a truer picture is provided by G. Box & G. TIAO, *supra* note 54, at 1:

Opinion as to the value of Bayes' Theorem as a basis for statistical inference has swung between acceptance and rejection since its publication in 1763. During periods when it was thought that alternative arguments supplied a satisfactory foundation for statistical inference Bayesian results were viewed, sometimes condescendingly, as an interesting but mistaken attempt to solve an important problem. When subsequently it was found that initially unsuspected difficulties accompanied the alternatives, interest was rekindled. Bayes' mode of reasoning, finally buried on so many occasions, has recently risen again with astonishing vigor.

Furthermore, as Schum observes in his review of Cohen's book:

I cannot agree that the subjective interpretation of probability appeals to few researchers in the natural and social sciences. I find it hard to believe that the author is innocent of the prevalence of the subjective interpretation of probability in the decision-theoretic areas of economics, business, medicine, and psychology. In these and other areas many researchers would prefer to use Bayesian statistical methods for analysis and interpretation of data but do not because they also wish to have their papers accepted by journal editors who, all too frequently, adhere to classical statistical approaches based on a relative-frequency interpretation of probability and reject other approaches, not always for informed reasons.

Schum, *supra* note 16, at 478.

⁵⁷ Bayes's Theorem is part and parcel of the theory of probability presented in every elementary text on the subject. It is all but implicit in the very definition of conditional probability. It can be used with either objectively or subjectively determined probabilities. See, e.g., G. Box & G. TIAO, *supra* note 54, at 12-20; Y. CHOU, *supra* note 35, at 412.

The dispute about "Bayesian methods," Brilmayer & Kornhauser, *supra* note 1, at 135 n.68, is quite different. It pertains to the sticky problem of statistical inference, or the formulation of generalizations (or statements about population parameters) on the strength of sample data. The "classical" or "orthodox" school, originating with R.A. Fisher, J. Neyman, E.S. Pearson, and others, relies on the techniques of point (or interval) estimation and hypothesis testing. Roughly speaking, classical statistics utilizes sample data as its only source of relevant information. Bayesian inference, on the other hand, demands the processing of prior information as well as sample data. The prior information is modified by the sample data through the repeated use of Bayes's Theorem. See 1 D. LINDLEY, *INTRODUCTION TO PROBABILITY AND STATISTICS* xi (1965). For an overview of the different approaches to this and related problems, see H. RAIFFA, *supra* note 31, at 278-88.

If Brilmayer and Kornhauser, and others who challenge the use of the probability axioms in the domain of law, were right, then classical as well as Bayesian statistics in litigation would be suspect. Whether or not Brilmayer and Kornhauser recognize this is unclear. They imply that there are "schools of statistics" that do not rest on the axioms given at note 28 *supra*, Brilmayer & Kornhauser, *supra* note 1, at 136-37. They take examples from L. COHEN, *supra* note 8, that, if valid, undermine both classical and Bayesian inference; they then act as if only "Bayesian assumptions" and "Bayesian proof schemes" are implicated. Brilmayer & Kornhauser, *supra* note 1, at 145 & n.103.

It is unfortunate that the use of the term "Bayesian" in the literature is so unstructured as to encourage such errors. Cf. note 37 *supra* (confusion regarding subjective and "Bayesian" probabilities). To quote the tongue-in-cheek remark of one statistician, "while non-Bayesians should make it clear in their writing whether they are *non-Bayesian Orthodox* or *non-Bayesian Fisherian*, Bayesians should also take care to distinguish their various denomina-

asserts only that the prior probability one proposes to use in equation (1) is somehow not a "true" probability and that the resulting posterior probability is therefore also not a "real" probability. That is, it contends that even in theory jurors cannot produce by subjective methods a number that qualifies as a "real" probability suitable for manipulation by valid mathematical theorems.

Two replies to this criticism are available. First, the objection misses the point of the modest chart approach. That approach does not ask the jurors to produce *any* number, let alone one that can qualify as a probability. It merely shows them how a "true" prior probability would be altered, if one were in fact available. It thus supplies the jurors with as precise and accurate an illustration of the probative force of the quantitative data as the mathematical theory of probability can provide. Such a chart, it can be maintained, should have pedagogical value to the juror who evaluates the entire package of evidence solely by intuitive methods, and who does not himself attempt to assign a probability to the "soft" evidence.

The other, more fundamental response is that there appears to be no reason in principle why a juror could not generate a prior probability that could be described in terms of the objective, relative-frequency sort of probability. One could characterize the juror's prior probability as an *estimate* of the proportion of cases in which a defendant confronted with the same pattern of nonquantitative evidence as exists in the case at bar would in fact turn out to have stabbed the deceased.⁵⁸ As Brilmayer and Kornhauser empha-

tions of *Bayesian Epistemologist*, *Bayesian Orthodox*, and *Bayesian Savages*." Bartlett, *Discussion on Professor Pratt's Paper*, 27 J. ROYAL STATISTICAL SOC'Y (SERIES B) 192, 197 (1965).

⁵⁸ The probability that the defendant used the knife to kill the deceased, conditioned on all the evidence presented to the jury, can be given a similar, relative-frequency interpretation. Imagine 100 homicide cases in which identical evidence is presented. The same facts are testified to by persons with the same demeanor and the same grounds for their testimony. To assert that the probability that the state's allegation (that the defendant wielded the knife) is true given all this information in a particular case is, say, $\frac{1}{2}$, means that in 50 of the 100 cases the defendants actually stabbed the victims and that in the remaining 50 they did not. If the jurors had more evidence in some of these 100 cases, they might be able to decide into which group a particular case would be more likely to fall. But, by hypothesis, they do not have such information, and each of the 100 cases looks identical, so the relative-frequency calculation must be based on all 100 cases.

An analogy can be made to the problem of stating (without peeking) the probability that a given coin, out of 100 coins flipped at once, has come up heads. To say that this probability is $\frac{1}{2}$ is to assert, in relative-frequency terms, that 50 of the coins showed heads and that the remaining 50 turned up tails. If one had more information about some of the coins' equations of motion, a different probability might be calculated. But, by hypothesis, one has identical evidence in each case and knows only that each coin was flipped in the same general way, at the same time, and onto the same surface.

size, the number so conceived is only an approximation of the "true" proportion or probability,⁵⁹ and there is no workable, independent method by which the accuracy of the juror's estimate could be measured.⁶⁰ But this practical difficulty does not undercut the conceptual point. Even a more direct measure of relative frequency can be faulty. An "objectivist" who estimates the probability that a woman will bear a male child by sampling to find the proportion of newborns who are male may also make mistakes. He may misidentify the sex, miscount, or misrecord the data. Naturally, to the extent the probabilities inserted in Bayes's formula are erroneous, so too will be the posterior probability it generates.⁶¹ But this is an operational concern.⁶² As important as it may be in coming to a decision about the desirability of implementing the procedure, it does not suggest any conceptual defect in the use of the formula. The juror can have the same confidence in the number produced by the formula as he does in the prior probability he estimates. And presumably, jurors already do make such estimates, even if not in numerical terms.⁶³

IV. WHAT IS RATIONAL?

Still, Brilmayer and Kornhauser ask, why is it "rational" to believe that the formula works correctly?⁶⁴ Might not a formula derived from another set of axioms provide a more rational solution to the problem of modifying prior probabilities? If Brilmayer and Kornhauser are using the term "rational" in a conventional way,⁶⁵

⁵⁹ Brilmayer & Kornhauser, *supra* note 1, at 140-41.

⁶⁰ There are some conceivable but socially undesirable methods of verification. See note 41 *supra*. This observation serves to counter the contention that a juror's estimate is untestable even in principle and is therefore meaningless.

⁶¹ Mathematically, the uncertainty in the estimation of the prior probability can be expressed, more or less, by using a "spread out" prior-probability mass or density function, $f(X)$, instead of a single number $P(X)$, to produce, according to Bayes's Theorem, a posterior-probability mass or density function, $g(X/E)$, instead of a single posterior-probability number, $P(X/E)$. See, e.g., G. Box & G. Tiao, *supra* note 54, at 10-11.

⁶² It is, as Tribe, *supra* note 3, at 1358-59, observes, one of the reasons that implementation of the Finkelstein-Fairley proposal might convey a false sense of precision to the jurors. Cf. Gerjuoy, *supra* note 6, at 39 ("a sensible quantitative measure of relevance in terms of probabilities can be formulated, but . . . Bayesian-computed probabilities will very rarely . . . be useful for deciding on the relevance of evidence proffered during actual litigation").

⁶³ See, e.g., Schwartz, *Decision Analysis: A Look at the Chief Complaints*, 300 NEW ENGLAND J. MED. 556, 557 (1979).

⁶⁴ Brilmayer & Kornhauser, *supra* note 1, at 135-36.

⁶⁵ On the conventional meaning of the word, see J. BENNETT, RATIONALITY (1964); C. HEMPEL, ASPECTS OF SCIENTIFIC EXPLANATION 463-86 (1965); H. LEIBENSTEIN, BEYOND ECONOMIC MAN: A NEW FOUNDATION FOR MICROECONOMICS 71-94 (1976); J. VON NEUMANN & O. MORGENTHAU, *supra* note 19, at 8; J. RAWLS, A THEORY OF JUSTICE 142-50 (1971); A. SEN,

there can be no direct mathematical or empirical demonstration that using probability theory to combine the statistical evidence with the juror's estimate of the prior probability is more "rational" than deducing the posterior probability by methods that reject the usual principles of statistical reasoning. That is because rationality is a normative concept not amenable to empirical measurement.⁶⁶ A theory of rational behavior does not "predict" anything empirically measurable.⁶⁷ It only prescribes how one should behave if one wishes to conform to that theory's conception of rationality.⁶⁸ The "rational man" of economic theory, for example, always acts to maximize utility, and his utility function is monotonic and quasi-concave.⁶⁹ Observations of individuals whose behavior is not

COLLECTIVE CHOICE AND SOCIAL WELFARE 34 (1970); Davidson, *Actions, Reasons and Causes*, 60 J. PHILOSOPHY 685 (1963); Gibson, *Rationality*, 6 PHILOSOPHY & PUB. AFF. 193 (1977); Mabbott, *Reason and Desire*, 28 PHILOSOPHY 113 (1953); Sen, *Rational Fools: A Critique of the Behavioral Foundations of the Economic Theory*, 6 PHILOSOPHY & PUB. AFF. 317 (1977); note 19 *supra*.

⁶⁶ Brilmayer and Kornhauser seem to think otherwise. They cite the example of poker-playing computer programs to support their surprising view that "theories about what constitute a rational approach" can be "validated" by "empirical testing of the theory's predictions." Brilmayer & Kornhauser, *supra* note 1, at 135. Perhaps their use of the word "rational" is imprecise, and what they meant to say is that one can test whether one decision-making strategy works better than another in meeting specified criteria such as beating one's opponent in poker. Certainly, one can "validate" the "accuracy" of a particular poker-playing algorithm in this way. Cf. note 73 *infra* (citing to studies of the relative accuracy of formal statistical methods and intuitive judgments). But one cannot demonstrate empirically that it is "rational" to beat one's opponent in poker. See generally, Oppenheim, *Rational Decisions and Intrinsic Valuations*, in NOMOS VII—RATIONAL DECISION 217 (C. Friedrich ed. 1964).

⁶⁷ A theory of rational behavior can, of course, be used descriptively, to model actual behavior. See, e.g., C. COOMBS, R. DAWES & A. TVERSKY, *supra* note 19, at 122-29; Hogarth, *Cognitive Processes and the Assessment of Subjective Probability Distributions*, 70 J. AM. STATISTICAL A. 271 (1975); Slovic, Fischhoff & Lichtenstein, *Behavioral Decision Theory*, 28 ANN. REV. PSYCH. 1, 13-17 (1977). Indeed, the National Science Foundation is funding an investigation by David Schum, a mathematical psychologist at Rice University, of whether probability judgments involving the kind of evidence that might be used in a trial setting are better described by the calculus proposed by L. COHEN, *supra* note 8, or by ordinary probability theory.

The proposal to use Bayes's formula in legal factfinding, however, does not purport to describe how jurors *do* behave. It is concerned with how they *should* behave. Thus it does not rely on the assumption that the processing of probabilistic information by real jurors reflects that prescribed by Bayes's formula. Rather, it assumes that more accurate factfinding will occur if the technique is implemented. See note 73 *infra*. Cf. de Finetti, *supra* note 31, at 111 n.e ("probability theory is not an attempt to describe actual behavior; its subject is coherent behavior, and the fact that people are only more or less coherent is inessential").

⁶⁸ See, e.g., C. COOMBS, R. DAWES & A. TVERSKY, *supra* note 19, at 122; A. RAPOPORT, *STRATEGY AND CONSCIENCE* (1964); Lempert, *supra* note 6, at 1023. Recognizing the normative quality of theories of rational behavior, a number of authors have attempted to develop a "metaethics" of rational choice. See, e.g., Anderson, *The Place of Principles in Policy Analysis*, 73 AM. POL. SCI. REV. 711, 715-23 (1979).

⁶⁹ See generally W. NICHOLSON, *supra* note 43, at 43-53.

"predicted" by this model—for example, a Buddhist monk who shuns worldly possessions or a Kung San tribesman who wants no more than an ample supply of mongongo nuts—cannot disprove the economist's belief that a "rational man" will behave as his theory demands.⁷⁰

Of course, Brilmayer and Kornhauser may be using "rational" in some special but unstated way. It would appear that they conceive of "rationality" in a limited "means-end" or instrumental sense, and that their criticism can be restated more clearly as follows: Why should we accept the axioms of probability theory as generating the most accurate solution to the problem of modifying prior probabilities—especially those that are ascertained by subjective methods? One answer is frankly intuitionist. One can simply insist that the axioms (or some higher-level postulates) are intuitively satisfying.⁷¹ Needless to say, it is notoriously difficult to convince determined skeptics by this retort. Fortunately, a pragmatic response is also available.⁷² The equations of the axiomatized theory of probability—like the rules of logic and arithmetic—work admirably in other contexts. Physicists, engineers, economists, geneticists, businessmen, actuaries, bookmakers, and casino operators prove, day in and day out, that the mathematical theory provides more useful and more accurate predictions of important phenomena than any alternative methods. Surely the probability axioms work sufficiently well for objectively estimated probabilities. Why should they not serve as well when applied to thoughtful, subjective estimates? Perhaps the laws of probability really are suspended in the courtroom, but the burden of proof should fall on those who claim that other ways of reasoning about probabilities may be more accurate.⁷³

⁷⁰ See, e.g., Leff, *Economic Analysis of Law: Some Realism about Nominalism*, 60 VA. L. REV. 451, 457 (1974); W. NICHOLSON, *supra* note 43, at 155-56.

⁷¹ See note 19 *supra*.

⁷² One student of probability theory expressed the pragmatic position clearly when he observed that "a bookmaker had better be jolly sure that the odds he posts satisfy the axioms of the probability calculus." H. KYBURG, *supra* note 46, at 96. See also text and notes at notes 19-20, 32 *supra*.

⁷³ There is a sizeable body of experimental work suggesting that more accurate predictions about complex matters are made by formal statistical methods than by intuitively inspired judgments. See Underwood, *supra* note 26, at 1423-24. There is also an interesting body of literature attesting to the value of using Bayes's formula (often with probabilities obtained from relative-frequency data) in medical diagnoses. See, e.g., Diamond & Forrester, *Analysis of Probability as an Aid in the Clinical Diagnosis of Coronary-Artery Disease*, 300 NEW ENGLAND J. MED. 1350 (1979); Gustafson, *Evaluation of Probabilistic Information Processing in Medical Decision Making*, 4 ORGANIZATIONAL BEHAVIOR & HUMAN PERFORMANCE 20 (1969); Rifkin & Hood, *Bayesian Analysis of Electrocardiographic Exercise Stress Testing*, 297 NEW ENGLAND J. MED. 681 (1977). Furthermore, empirical Bayes techniques have proved

The critics who oppose using the laws of probability in the courtroom and those who have formulated theories of alternative mathematical structures for calculating probabilities have not yet met this burden.

effective in predicting the performance of applicants to law school. See Rubin, *Using Empirical Bayes Techniques in the Law School Validity Studies*, 75 J. AM. STATISTICAL A. (forthcoming).

In any event, if the point Brilmayer and Kornhauser mean to raise by their discussion of rational models is the question whether formal statistical techniques like Bayesian analysis would improve the performance of lay persons in evaluating probabilistic information correctly, their skepticism about the probability axioms remains mysterious. Even if such procedures did not work well under courtroom conditions, the fault, to paraphrase Cassius's remark to Brutus, might not lie with the probability axioms but with those persons trying to act as the axioms prescribe. Cf. de Finetti, *supra* note 31, at 111 n.e ("It is . . . natural that mistakes are common in the . . . complex realm of probability; nevertheless . . . , fundamentally, people behave according to the rules of coherence even though they frequently violate them (just as [they depart from the rules of] arithmetic and logic).").