Naturalized Epistemology and the Law of Evidence

Brian Leiter
Ronald J. Allen
NATURALIZED EPISTEMOLOGY AND THE LAW OF EVIDENCE

Ronald J. Allen* and Brian Leiter**

INTRODUCTION.................................................................................................1492
I. NATURALIZING EPISTEMOLOGY.................................................................1493
II. NATURALIZED EPISTEMOLOGY AND THE CONCEPTUAL FOUNDATIONS OF EVIDENCE..............................................................1499
III. NATURALIZED EPISTEMOLOGY AND THE EMPIRICAL ADEQUACY OF EXISTING THEORIES OF THE EVIDENTIARY PROCESS.................................................................................................................1503
   A. Expected Utility Theory ........................................................................1503
   B. Bayes' Theorem ...................................................................................1507
   C. The Economic Analysis of Evidence ..................................................1510
      1. The Implications of Rational Choice Theory.................................1516
      2. Bayes' Theorem and the Meaning of Relevance.............................1519
      3. Rootless Theorizing .......................................................................1521
   D. The Relative Plausibility Theory and Naturalized Epistemology........1527

* John Henry Wigmore Professor of Law, Northwestern University School of Law.
** Charles I. Francis Professor in Law, Professor of Philosophy, and Director of the Law & Philosophy Program, the University of Texas at Austin.

We are indebted to Craig Callen, Gilbert Harman, Susan Haack, Richard Posner, and to participants at the Northwestern University School of Law Faculty Workshop, the University of North Carolina Workshop on Law and Philosophy, and the University of Virginia School of Law “New Perspectives on Evidence” conference for helpful comments. We are also indebted to Ian Logan and William Rohner for their research assistance.
INTRODUCTION

THIS paper sets itself two tasks: (1) to introduce lawyers to important recent developments in epistemology; and (2) to show lawyers and philosophers how these developments provide a conceptual foundation for some familiar approaches to problems from the law of evidence. The developments in epistemology have not, to date, been much noted in legal scholarship, despite their importance in philosophy and their coincidence with some widely shared approaches to evidence scholarship. This may partly explain—or perhaps is partly explained by—the unfortunate fascination in some quarters of the legal academy with “postmodern” conceptions of knowledge and truth, conceptions notable for their superficiality and for the fact that almost no philosophers subscribe to them.¹ It may also partly explain—or be explained by—the even more common search by the legal professoriat for the algorithm that, a priori, provides answers to important legal questions. In the field of evidence, while there is some interest in post-modern epistemology,² more typical is either the search for the appropriate algorithm, such as Bayesian decision theory³ or more recently mi-

¹ Postmodernists are typically skeptical about the possibility of objective truth, as well as our capacity to find objective truth in the world. Such an outlook is remarkably useless for evidence law, as discussed in Mirjan Damška, Truth in Adjudication, 49 Hastings L.J. 289 (1998). Postmodern conceptions of truth and knowledge are patiently and clearly criticized in Alvin I. Goldman, Knowledge in a Social World 9-40 (1999).


croeconomics,⁴ or simply the complete neglect of epistemological matters. In our view, however, the naturalistic turn in epistemology of the past thirty years—and, in particular, that branch of naturalized epistemology known as social epistemology—provides the most appropriate theoretical framework for the study of evidence, as it does for virtually any intellectual enterprise concerned with the empirical adequacy of its theories and the truth-generating capacity of its methodologies. Evidence scholarship and law are concerned with both, and thus naturalized epistemology provides a fruitful way of understanding the limitations of some of the existing efforts to provide theoretical and philosophical foundations to evidence law. It also provides a way to conceptualize and evaluate specific rules of evidence. It has the additional virtue of explaining what most evidence scholars do, regardless of their explicit philosophical commitments. For the great bulk of evidentiary scholars, then, this paper merely solidifies the ground beneath their feet.

Part I will involve philosophical stage-setting, aimed at making recent developments in philosophy intelligible to lawyers. Part II will situate the naturalized epistemology approach briefly with respect to other “grand” attempts to provide conceptual foundations to evidence law. Part III will employ the naturalized epistemology approach to criticize existing theories of different evidentiary rules, including Bayesianism, expected utility theory, and Judge Richard A. Posner’s recent economic analysis of the law of evidence. Part III will conclude with a brief examination of another evidentiary theory—the relative plausibility theory—that better meets the demands of the naturalistic approach than the other theories considered. Part IV will show how this epistemological approach applies to specific rules of evidence and will sketch directions for further research.

I. NATURALIZING EPISTEMOLOGY

If the twentieth century began with the “linguistic turn” in philosophy⁵—the idea that traditional philosophical problems were

---

⁵ For an overview, see The Linguistic Turn: Recent Essays in Philosophical Method (Richard Rorty ed., 1967).
best analyzed and conceptualized as problems about language and its relation to the world—its concluding quarter-century has been marked by "the naturalistic turn." "Naturalism" has acquired multiple meanings in recent years, but the core commitment of naturalists (at least for our purposes here) is methodological: Philosophy should be continuous with a posteriori inquiry in the empirical sciences; philosophy cannot be an exclusively a priori discipline. At one extreme, best exemplified by Professor W.V.O. Quine, this means the replacement of philosophy by empirical science. In a less extreme and more influential form, best exemplified by Professor Alvin I. Goldman, philosophical theorizing is constrained by empirical facts and often demands supplementation by empirical information. Thus, in the case of individual epistemology—that branch of the theory of knowledge which focuses "on mental operations of cognitive agents in isolation or abstraction from other persons"—we cannot craft epistemic norms (norms that would guide our acquisition of knowledge) without empirical information about how the human cognitive apparatus actually works. Since for Goldman a belief counts as knowledge if "caused by a generally reliable process," it follows that "[o]nly if (some of) our basic cognitive processes are... reliable... can we qualify as knowers. Therefore, whether we so qualify hinges, in part, on facts

---

6 For a review, with citations to much of the literature, and an attempt to sort out the different meanings, see Brian Leiter, Naturalism and Naturalized Jurisprudence, in Analyzing Law: New Essays in Legal Theory 79, 80-92 (Brian Bix ed., 1998).

7 Many naturalists also adopt a substantive view to the effect that the only things that exist are those countenanced by the natural sciences. For discussion and citations, see id. at 81-84. All substantive naturalists of this kind profess to be driven to the view by methodological naturalism, though it is hard to see how appeal to the empirical sciences would underwrite the most extreme forms of physicalism. See also Jerry Fodor, Look!, London Rev. Books, Oct. 29, 1998, at 3 (reviewing Edward O. Wilson, Consilience: The Unity of Knowledge (1998)) (noting that the proliferation of special sciences does not support the idea that all sciences will be reducible to physics).

8 E.g., W.V. Quine, Epistemology Naturalized, in Ontological Relativity and Other Essays 69 (1969).

9 E.g., Alvin I. Goldman, Epistemology and Cognition (1986). For an indication of the scope of the influence of Goldman's approach, see Philip Kitcher, The Naturalists Return, 101 Phil. Rev. 53 (1992), and the discussion and citations therein.

10 Goldman, supra note 1, at 4.

11 Goldman, supra note 9, at 51. The details of Goldman's externalist reliabilism do not matter here.
in [empirical] psychology's bailiwick.\textsuperscript{12} Notice that under Goldman's approach, the relevance of empirical psychology to epistemology results from a proffered conceptual analysis of a conventional philosophical kind, that is, the suggestion that a belief is knowledge if caused by a reliable process. It is \textit{that} conceptual analysis that makes empirical science relevant to assessing which beliefs count as knowledge and to crafting epistemic norms to regulate belief formation so that it yields knowledge.

But why naturalize epistemology or any branch of philosophy, in the sense of making it dependent upon empirical science? The motivations are various, and they sometimes depend on the part of philosophy that one considers. One important impetus for naturalizing philosophy was Quine's seminal attack on the distinction between "true in virtue of meaning" ("analytic" truths) versus "true in virtue of empirical fact" ("synthetic" truths).\textsuperscript{13} The former were thought, at least by some logical positivists, to constitute the distinctive domain of philosophical expertise, while the latter were the property of empirical science. Philosophers would analyze and clarify the meanings and concepts that define the framework in which empirical science operates. But if there is no distinctive domain of truths of meaning, as Quine argues, then there is nothing for philosophy to do: All the intellectual work falls to empirical science.\textsuperscript{14} Few philosophers have followed Quine this far, though the general moral—that philosophical claims are always vulnerable to the successes of a posteriori inquiry—has been hard to deny.\textsuperscript{15} A

\textsuperscript{12} Id. at 53.
\textsuperscript{14} Of course, it is important for Quine's holism that there are no strictly "synthetic" truths either, since every claim can be maintained in the face of recalcitrant evidence, as long as we are willing to adjust other aspects of our worldview. See Quine, Two Dogmas, supra note 13, at 42–44.
\textsuperscript{15} This is not to say that it is universally heeded, as Professor Gilbert Harman correctly complains in Doubts About Conceptual Analysis, \textit{in} Philosophy in Mind 43 (Michaelis Michael & John O'Leary-Hawthorne eds., 1994).
priori conceptual analysis can continue to play a role in philosophy, but it offers no special insight into timeless truths and is always vulnerable to radical revision or elimination in the light of empirical progress.\textsuperscript{16}

Some of the reasons for naturalism are more particular to epistemology. Professor Edmund L. Gettier’s refutation of the then-prevailing analysis of the concept of knowledge as “justified true belief”\textsuperscript{17} was taken by many to show “that the epistemic status of a belief state depends on the etiology of the state and, consequently, on psychological facts about the subject.”\textsuperscript{18} What Gettier’s refutation meant, in other words, was that the \textit{actual} causal trajectory leading from evidentiary input to belief was crucial for establishing whether the resultant belief would count as “knowledge.”\textsuperscript{19} But the actual causal trajectory leading to belief fell within the domain of empirical science, not philosophy conceived as mere conceptual analysis. Quine, by contrast, thought the failure of Cartesian foundationalism rendered the normative project of philosophical epistemology futile and recommended its replacement with the empirical study of the causal relationship between sensory inputs and theoretical outputs.\textsuperscript{20}

This is not, to be sure, the place for a sustained defense of the naturalistic turn; we only want to locate our project within the existing philosophical landscape. Whether epistemology in general should be naturalized is, in any case, irrelevant to the question of whether naturalized epistemology provides a fruitful way of under-

\textsuperscript{16} This may understate the difficulties confronting conceptual analysis. For stronger criticisms, see Brian Leiter, The Naturalistic Turn in Legal Philosophy, 00 APA Newsl. on Law & Phil. 142 (2001).

\textsuperscript{17} Edmund L. Gettier, Is Justified True Belief Knowledge?, 23 Analysis 121 (1963).

\textsuperscript{18} Kitcher, supra note 9, at 60.

\textsuperscript{19} Goldman’s alternative conceptual analysis of knowledge (and its component elements like “justification”) was responsive to this problem, by introducing as an element of the concept the relevance of the actual trajectory leading to belief. See Goldman, supra note 9, at 42–57. That in turn made empirical science relevant to the project of epistemology. See supra text accompanying notes 11–12.

\textsuperscript{20} See Quine, supra note 8, at 83. There is some debate about the role of normative epistemology in Quine’s approach. Compare Jaegwon Kim, What is “Naturalized Epistemology”? in Supervenience and Mind (1993) (finding no such role), with Richard Foley, Quine and Naturalized Epistemology, 19 Midwest Stud. Phil. 243 (1994) (finding such a role). For Quine’s own views on the question (which support, in part, Foley), see W.V. Quine, Pursuit of Truth 19–21 (1990).
standing evidence law. The latter is our central contention and the one this paper as a whole is meant to vindicate.

At the most general level, then, naturalizing epistemology means viewing philosophical theorizing about knowledge as more than an a priori armchair exercise, but rather as continuous with and dependent upon empirical science. Insofar as we follow Goldman in retaining the distinctively normative element of epistemology—the regulation of our cognitive activities so that they result in knowledge—then the rationale for naturalization is “that one cannot give the best [normative] advice about intellectual operations without detailed information about mental processes” and how they really work. But information about “mental processes” suffices only for individual epistemology. As Goldman points out, “[t]he bulk of an adult’s world-view is deeply indebted to her social world. It can largely be traced to social interactions, to influences exerted by other knowers . . . . It is imperative, then, for epistemology to have a social dimension.” Social epistemology is simply that branch of naturalized epistemology concerned not with individual knowers but with the social processes and practices that inculcate belief.

While naturalized individual epistemology depends primarily on the empirical sciences of the human cognitive apparatus, naturalized social epistemology must consider the range of empirical sciences that examine the social mechanisms of belief-inculcation. In what follows, we shall often speak of “naturalized epistemology” and “social epistemology” interchangeably.

It does bear noting, however, that there is another, rather different sense of “social epistemology” current in the literature, with which our approach should not be confused. Some of those who talk about “social” epistemology mean an approach which tries to

23 More precisely, as Goldman says in his recent book, the “social” element of social epistemology is threefold: (1) “It focuses on social paths or routes to knowledge;” (2) it “does not restrict itself to believers taken singly. It often focuses on some sort of group entity—a team of co-workers, a set of voters in a political jurisdiction, or an entire society—and examines the spread of information or misinformation across that group’s membership;” and (3) “instead of restricting knowers to individuals, social epistemology may consider collective or corporate entities, such as juries or legislatures, as potential knowing agents.” Goldman, supra note 1, at 4–5.
explain away what "passes" for knowledge in some community as simply the product of social factors rather than epistemic considerations of any kind. We may call this approach debunking social epistemology since it means to unmask the pretense of putative knowledge claims by showing them to reflect social interests and circumstances. Of course, debunking social epistemology might be the upshot of a naturalized approach to epistemology. It could turn out, as an a posteriori matter, that the best explanation of the claims we call "knowledge" only makes reference to non-epistemic social factors. We need not rule out that possibility, and, as naturalists, we cannot rule it out a priori. But it remains an open empirical question whether debunking social epistemology is accurate in its portrayal of knowledge.

Social epistemology, in the sense we adopt, is normative or regulative in its ambitions. We want to ask, as Goldman puts it in his recent important book on the subject, "Which [social] practices have a comparatively favorable impact on knowledge as contrasted with error and ignorance?" Social epistemology is, in this respect, veritistic (to borrow Goldman's term): It is concerned with the production of knowledge, meaning (in part) true belief. The normative naturalized epistemologist embraces as his goal the promulgation of norms by which to regulate our epistemic practices so that they yield knowledge. In the case of individual epistemology, this means the norms governing how individuals should acquire and weigh evidence as well as, ultimately, form beliefs. In the case of social epistemology, it means the norms governing the social mechanisms and practices that inculcate belief.

The rules of evidence are a prime case of the latter, for these rules structure the epistemic process by which jurors arrive at beliefs about disputed matters of fact at trials. As such, the rules of evidence are a natural candidate for investigation by social epistemologists. We may ask of any particular rule: Does it increase the likelihood that jurors will reach true beliefs about disputed matters of fact? Of course, it does not make sense to ask that of every rule

24 E.g., Steve Fuller, Social Epistemology (1988).
25 See the discussion of "radical naturalism" in Kitcher, supra note 9, at 96.
26 Goldman, supra note 1, at 5.
27 Id. See generally id. at 79–100 for some of the details and complications involved in assessing practices along veritistic dimensions.
since some rules of evidence—for example, Federal Rules of Evidence 407–11—are not meant to facilitate the discovery of truth but to carry out various policy objectives such as reducing accidents and avoiding litigation. We return to this issue in Part II below.

Social epistemology, as a branch of naturalized epistemology, must honor two particularly important constraints. First, as Goldman remarks, “advice in matters intellectual, as in other matters, should take account of the agent's capacities. . . . As in the ethical sphere, 'ought' implies 'can'.” In other words, normative epistemology, like normative ethics, cannot require of agents actions (mental or physical) that they cannot perform. Second, naturalized epistemology “assumes that cognitive operations should be assessed instrumentally: given a choice of cognitive procedures, those which would produce the best set of consequences should be selected.” In other words, normative epistemology must always ask about the actual consequences of alternative sets of epistemic norms, that is, which are the most effective means for producing knowledge. Given these constraints, naturalized epistemology must then be continuous with empirical science in two quite particular senses: (1) We need to know what epistemic norms in fact lead to the acquisition of knowledge; and (2) we need to identify epistemic norms that are actually usable by creatures like us. This rules out epistemic norms which require of cognizers belief-formation practices (individual or social) beyond their ken. Naturalized epistemology, in short, emphasizes the instrumental character of normative theorizing in epistemology, but the only way to assess instrumental claims is to do so empirically, that is, by finding out what means really bring about what ends.

II. NATURALIZED EPISTEMOLOGY AND THE CONCEPTUAL FOUNDATIONS OF EVIDENCE

Professor James B. Thayer, in his famous treatise, wrote, “[w]hen men speak of historical evidence and scientific evidence, and the

---

28 Goldman, supra note 21, at 510.
29 Id. at 520. For a similar approach, see Larry Laudan, Normative Naturalism, 57 Phil. Sci. 44, 46 (1990).
30 See, e.g., Goldman, supra note 21, at 512–13.
evidences of Christianity, they are talking about a different sort of thing. The law of evidence has to do with the furnishing to a court of matter of fact, for use in a judicial investigation.^^31 Social epistemology, as a naturalistic approach to the law of evidence, is premised on a rejection of Thayer's position, at least as it is understood as a claim about differences in kind. In fact, social epistemology might properly consider history, science, and religion from the veritistic standpoint, trying to ascertain the extent to which the constitutive practices of each arena reliably produce knowledge. For the social epistemologist, then, the law of evidence is not a "different sort of thing" from any other practice that has as one of its elements the production of knowledge.

In another sense, though, the law of evidence is different. It operates within a distinctive social institution (the trial and the adversarial system more generally), rather than within the laboratory or the library. It employs a distinctive division of epistemic labor, with one set of actors (judges) first determining the evidential base upon which another set of actors (jurors) will rely in forming beliefs about disputed matters of fact.^^32 Finally, considered as a whole, the law of evidence does not have as its only aim the production of true belief. It is true that Federal Rule 102 defines the "purpose" of the rules as "that the truth may be ascertained,"^^33 but some of the rules themselves have no veritistic dimension,^^34 while others mix veritistic and non-veritistic concerns.^^35

---


^^32 The division is actually more complex than this since judges make their determination within an evidential base created not by them but by the advocates for each party to a dispute. In theory, at least, the adversarial system should produce a very large evidential base, one that might even match or exceed in scope the evidential base that the scientist or historian might consider for his distinctive purposes.

^^33 Fed. R. Evid. 102.

^^34 E.g., Fed. R. Evid. 407-11 (excluding evidence to achieve various non-truth-seeking policy objectives). The common-law privileges also match this description.

^^35 E.g., Fed. R. Evid. 413-15 (concerning sex offense cases). These recent lamentable amendments to the Federal Rules—which admit the sexual background of sexual predators apparently regardless of its relevancy or prejudicial effect—are unique in authorizing admission of evidence on "policy" grounds. The cases, however, seem to be concluding that Federal Rules 413-15 do not dispense with Federal Rule 403. E.g., United States v. Meacham, 115 F.3d 1488, 1495 (10th Cir. 1997).
Still, considered as a whole, it is striking and important that the vast majority of the rules of evidence have as their primary rationale their (alleged) truth-conducive virtues. Competency of witnesses, authentication of evidence, relevancy, expert testimony, and hearsay (including the exceptions) all rest on the thought that inclusion and exclusion of evidence in line with these rules will increase the frequency with which truth is ascertained. Social epistemology, as a framework for thinking about evidence, is committed to an investigation of the extent to which the conceptual foundations of evidence law rest upon the aim of true belief and the extent to which evidence law succeeds in that aim.

This approach permits conceptual neutrality over some of the debates about evidence law familiar from the secondary literature. Thus, for example, we have no reason to take sides between those who advocate the "jury control principle" (the idea "that the organizing principle of Evidence law [is] a fear that lay jurors [will] misuse certain types of evidence") and those who advocate the "best evidence principle" (the idea that "[t]he best evidence must be given of which the nature of the case permits"). Both are quite obviously predicated on an interest in promoting true belief, and thus both are compatible with the social epistemology framework. Thus, the worry that "jurors [will] misuse certain types of evidence" is precisely the worry that they will misuse them by drawing inferences that lead to false beliefs. So too, the underlying impetus for the best evidence principle is precisely the idea that "the rules of evidence with respect to trial in a court of law are, and should be, focused primarily on ascertaining the truth about controverted issues of fact within appropriate resource constraints." Of course, the empirical adequacy of either account deserves investigation,

---

36 Hearsay is also in part driven by policy considerations, in particular in the criminal arena where there is a deeply-held belief in the importance of confrontation.
39 Nance, supra note 38, at 294.
with respect to both the accuracy of the description of the litigation process and the veritistic consequences of different approaches.

The jury control principle does, however, highlight an interesting feature of our evidentiary rules, namely, their epistemic paternalism:

Paternalism in any domain of legal regulation supposes that rules should substitute the rulemaker's judgment about what is best for agents for the agents' own judgments. Epistemic paternalism substitutes the rulemaker's judgment about what is epistemically best for agents for their own judgment. Assuming that the primary epistemic value is truth, epistemic paternalism entails designing rules of evidence that are epistemically best for jurors, i.e., that lead them to form true beliefs about disputed matters of fact. Doing so requires, of course, taking into account both the epistemic frailties of jurors, and the epistemic limits of the rule-appliers [the "gatekeepers"], namely judges.\(^{40}\)

Indeed, it is useful in analyzing the law of evidence to distinguish primary from secondary epistemic rules:

Primary epistemic rules take into account the epistemic shortcomings of jurors, such as their susceptibility to confusion and prejudice or their generally modest level of intellectual ability. Secondary epistemic rules take into account the epistemic shortcomings of judges, such as their general lack of expertise in scientific matters. The rule of evidence that excludes unscientific evidence is a primary epistemic rule in the sense that it is predicated on the assumption that jurors must be "protected" from junk science in forming beliefs about disputed matters of fact. The rule of evidence requiring judges to exclude unscientific evidence is a secondary epistemic rule in the sense that it requires judges to make an epistemic judgment about whether some piece of evidence is scientific or not.\(^{41}\)

We can evaluate either a primary or secondary epistemic rule along veritistic dimensions. With respect to the former, we ask whether evidence excluded and included in accordance with the rule will reliably lead jurors to form true beliefs about disputed


\(^{41}\) Id. at 815.
matters of fact. With respect to the latter, we ask whether the rule is such that judges can reliably apply it. Of course, a primary epistemic rule that fails along its veritistic dimension precludes any need to inquire about how it fares *qua* secondary epistemic rule.

III. NATURALIZED EPISTEMOLOGY AND THE EMPIRICAL ADEQUACY OF EXISTING THEORIES OF THE EVIDENTIARY PROCESS

Existing attempts to make theoretical sense of the evidentiary process have inadequately attended to their empirical adequacy and, in addition, have fairly systematically run afoul of the two constraints of naturalized epistemology: "ought implies can" and the instrumental character of normative advice. Formal models of legal decisionmaking often ask actors to do what they cannot do and (unsurprisingly) fail to inquire as to whether the formalized models at issue will *in fact* increase the veritistic reliability of the process. We shall critique three models of the evidentiary process, two of which fall prey to these problems and for which there is a substantial literature: expected utility theory as an explanation of burdens of persuasion, and Bayes' theorem as a theory of inference and relevance. The third model we consider is the economic analysis of evidence law represented by Posner's recent foray into the field. The economic model is less directly concerned with veritism than its competitors; our main interest in it is that it exemplifies the limits of formal, a priori reasoning about the evidentiary process. We conclude with a discussion of a fourth model of the evidentiary process, the relative plausibility theory, that we think better captures the relevant empirical phenomena and does not demand of decisionmakers that they engage in tasks beyond their ken.

A. Expected Utility Theory

Burdens of persuasion appear to lend themselves quite readily to decision-theoretic explanations. Actual truth is rarely known in any particular case. Otherwise, there would be little need for a trial. The large number of cases needing decision indicates that actual truth is rarely known. It certainly violates no canon of common sense to view these circumstances as calling for maximizing expected utility. Moreover, simple calculation generates the expected utility maxi-
mizing rules. In a case involving a binary choice where the disutili-
ties of wrongful verdicts are equal, the decision should be for
whomever the probabilities favor. This is the 0.5 rule of civil litiga-
tion. In cases involving more than two possible explanations, the
decision should be for the most probable (and here we see the first
problem, for this is not the law). If disutilities of wrongful decisions
are not equal, as in criminal cases where a wrongful conviction is
considerably worse than a wrongful acquittal, the decision rule is
adjusted to accommodate the difference. At least, this is what the
expected utility theorists argue.43

Naturalized epistemology should alert us to the fact, however,
that the relationship between the world of mathematics and the
world of human affairs may not be simple. Here there are two sub-
stantial difficulties. First, in its present version, expected utility
theory does not in fact describe the law although it may be used to
criticize it. The law applies burdens of persuasion to elements, not
to causes of action as a whole. Expected utility theory conflates the
to as though the question asked at trial were liability or no liabil-
ity. In a sense, that is in fact the question but only after burdens of
persuasion are applied to individual elements. Applying burdens of
persuasion understood as probability measures to elements yields
the well-known paradoxes of proof. If, for example, two stochasti-
cally independent issues are each established to a 0.6 probability,
the probability of them both being true is 0.36.44 Returning a ver-
dict for a plaintiff in such cases is not going to maximize expected
utility.

The puzzles press more deeply, however. The expected utility
theorist may respond by criticizing the law and arguing that it is the
conjunctive of elements that should be found to a specific level.
This, too, yields unacceptable consequences, by making the level of
proof of specific elements turn on the fortuity of the number of
elements in a cause of action. Take the example of theft and mur-

43 See, e.g., Kaplan, supra note 3, at 1071–73. For more detailed treatment of
expected utility theory, see David Kaye, The Limits of the Preponderance of the
Evidence Standard: Justifiably Naked Statistical Evidence and Multiple Causation,

44 For a discussion of this and other paradoxes, see Ronald J. Allen, A
Theft has considerably more elements than murder. To convict for theft requires on average that intent to steal be established to a higher probability than intent to kill for a murder conviction. This strikes all legal observers as both unacceptable and absurd.\textsuperscript{45} There is a still deeper formal problem here. Finding the probability of the conjunction of discrete elements may require massive amounts of information, and in any event cannot be done in the simple and direct manner of providing proof of the discrete elements. One of the logical implications of probability theory is, briefly, that virtually any relationship may exist between discrete elements and their combinations. For example, as the probability of two discrete elements each goes up, the probability of their conjunction may go down.\textsuperscript{46} As Professor Ron A. Shapira has summed up the situation, "one of the crucial things which require knowledge in all theories of evidence is a prior partition of the universe of discourse into equivalence classes, as, alternatively, a prior determination of essential properties of objects or relevant experimental variables."\textsuperscript{47} The point is that the path between even such a simple formalization as probability statements about discrete elements and the objectives of trials is quite unclear. The specter of running afoul of "ought implies can" now arises.

There is a second set of problems with expected utility theory: In its simplest manifestation, it ignores base rates, systematic errors in probability assessments, and the fact that it is not the subjective expectation of judges and jurors that the legal system wishes to maximize. More sophisticated versions of expected utility theory, by contrast, have not given a plausible account of how these matters could be taken into account in a way that would increase the probability of furthering the objectives of the legal system. All versions also neglect certain implications of subjective probabilities that will be discussed in the next subsection.

\textsuperscript{45} E.g., id. at 407.

\textsuperscript{46} This is an example of Simpson's paradox, which has only recently been introduced into the evidentiary literature. See Ronald J. Allen, Factual Ambiguity and a Theory of Evidence, 88 Nw. U. L. Rev. 604, 604, 608 (1994). For a thorough development of its implications, see Ron A. Shapira, The Susceptibility of Formal Models of Evidentiary Inference to Cultural Sensitivity, 5 Cardozo J. Int'l & Comp. L. 165, 169–77 (1997).

\textsuperscript{47} Shapira, supra note 46, at 187.
The legal system involves third-party decisionmakers—judges and juries—implementing the wishes and commands of the sovereign people—or less grandly of the policymakers—who are typically legislators and sometimes constitution writers. The utility to be maximized is that of the policymakers, not that of the judges and juries, and the two could be widely disparate. For example, with no knowledge of base rates or the relative accuracy of probability assessments, the lessons of expected utility theory are quite straightforward for factfinders. If the policymaker thinks he is in possession of such knowledge, that knowledge may dramatically affect the expected outcome from the policymaker's point of view. Take a simple example. If no factually liable defendants go to trial, the only kind of error possible is holding a defendant wrongfully liable (a false positive). Increasing the burden of persuasion can only increase the policymaker's expected utility, whatever it does to that of the factfinder. Analogously, policymakers may believe that probability assessments of factfinders are skewed in some fashion, generating the same problem.

No means of accommodating this point has been advanced. The proponents of expected utility theory have simply asserted that beliefs about base rates and the probability assessments can themselves be taken into account when forming subjective probabilities. This is true enough, but it is difficult to see what programmatic implications it may have, for informing the factfinder of this knowledge would have unpredictable effects on the factfinder's appraisal of the evidence. Any particular factfinder may overestimate or underestimate the probabilities of liability, and information about the systemic knowledge may lead to widely disparate adjustments to accommodate that knowledge. In any event, the current state of the law in the United States disconfirms a close connection between expected utility theory and actual trial process.48

48 A full treatment of the relevant issues would have to include the vast area of presumptions, inferences, and explicit modifications of the burdens of proof prevalent in American trial process.
B. Bayes' Theorem

Just as burdens of proof seem to lend themselves to decision theoretic analyses, so does the evidentiary process at trial seem to lend itself to a Bayesian interpretation. Indeed, it is quite natural to think of the evidentiary process as the updating of a prior probability in light of new evidence, as Bayes' Theorem prescribes. There are a number of difficulties with such an analysis from the standpoint of social epistemology, however.

The first worry is computational complexity, which raises the specter of violating "ought implies can." A huge and complicated data set is involved at most trials, even most "simple" trials. No computer, let alone any human, has the computational capacity to do the calculations necessary for the operation of Bayes' Theorem in a reasonable amount of time. Bayesians respond, appropriately, that it is not their fault that the world is complicated. The issue, however, is not fault, but reality: The world is complicated, and that fact constrains normative advice. The Bayesians might still retort that nothing within Bayes' Theorem instructs on what the unit of analysis should be. Thus, the factfinder can lump a bunch of stuff together and update his prior probability using the bunch of stuff as the datum of "new evidence." This move carries only a false promise. The real intellectual work will have been done in the "bunching," and the failure to "bunch" correctly will lead inexorably to false outcomes (except only by chance).

A second worry arises when we reflect upon the description of trials. Factfinders typically have no good sense of what is going on until the end of the trial at closing arguments. Moreover, they are not bound in any way by those arguments and are free to generate their own theories of what happened. This has two implications. First, once the factfinder hears the various theories in closing argument, to operate Bayes' Theorem it must then assign probabilities to the various theories. Those probabilities will be assigned in light of the evidence heard at trial, and thus all that evidence is what is called "old evidence," which simply means it

---

has already been taken into account. Second, even following the assignment of initial probabilities, new theories can emerge during deliberations. The emergence of a new theory requires, for Bayes’ Theorem, a reassignment of initial probabilities of all possible theories, and again the problem of old evidence rears up. The possible scope of Bayesian computations is thus exceedingly limited.\(^{31}\)

Another difficulty with Bayesian approaches to juridical evidence is that the assignments of initial probabilities, which are crucial to the application of the Theorem, are subjective and need respect only the conditions of consistency and summing to 1.0. That means that individuals can begin from radically different perspectives, and each, in Bayesian terms, will be operating equally rationally. Bayes’ Theorem provides no method of adjudicating such differences, which means it does not offer useful guidance for factfinders. In other contexts, such as science, these differences may be marginalized by convergence theorems that demonstrate that over time and with enough new evidence the divergent initial starting points will wash out and the result will converge on the truth. There is nothing even remotely analogous to this in the condition of trials. Jurors are more like scientists reflecting on new theories for the first time than like scientists who have generated substantial evidence over time designed to adjudicate between competing scientific theories. Without something to take the place of convergence theorems, the arguments about Bayes’ Theorem in the law are left with no obvious bridge between the subjective and the objective.\(^{32}\)

To be sure, all theories of juridical evidence will have a subjective component, but the irony of the Bayesian approach is that it implicitly exploits the false hope that by running one’s subjective beliefs through Bayes’ Theorem with the assistance of equally subjective likelihood ratios, something other than a subjective output will result. This is false. The risk is that the allure of the false hope will distract decisionmakers from what tools for reaching objectiv-

\(^{31}\) Allen, supra note 49, at 267–68. The point in the text holds at least as far as discovery is concerned. Justification may be a different matter, but the task at trial is more analogous to discovery than justification. For an analogous discussion of the role of Bayes’ Theorem in the sciences, see Mark Lange, Calibration and the Epistemological Role of Bayesian Conditionalization, 96 J. Phil. 294 (1999).

\(^{32}\) Allen, supra note 49, at 267.
ity they actually have available to them—tools that can be summarized in the notion of painstaking attention to and examination of the evidence and its logical and empirical implications. Moreover, the radical subjectivity of juridical Bayesianism is not a necessary component of theories of juridical proof. For example, the relative plausibility theory discussed in Section III.D below emphasizes the substantive component of factfinding and does not exploit the Bayesian mirage that algorithms may be substituted for substantive engagement with the evidence.

There are further intractable problems with the Bayesian approach. For example, Professor Leonard J. Savage's formalization of subjective probability includes the "sure thing" principle, which is the pivotal axiom that produces the interchangeability of subjective and objective probabilities. This axiom says that if act $A$ is preferred to act $B$ under one set of assumptions about nature, then augmenting the set of assumptions should not cause a reversal of preferences, that is, for $B$ to be preferred over $A$. For example, if you are given a menu and you prefer chicken to turkey and then are told that the kitchen also serves duck, it is a violation of the sure thing principle to say, "Given that additional information, I will switch my order to turkey."

Unfortunately for Bayesian approaches, humans disobey this axiom all the time. Suppose you believe that turkey requires great care in preparation, and you ordered chicken because you are risk averse. Duck is very difficult to prepare, however. Having learned that duck is on the menu, you have greater trust in the chef and so switch from chicken to turkey. This violates the sure thing principle. People regularly disobey this axiom because it requires the articulation of all logical propositions in a probability space—a daunting task even in a quite confined space, and an impossible one when the probability space ranges over all human affairs. (Remember: "Ought implies can"!) This is a particularly acute problem for group decisionmaking. As Savage himself wrote, "It would not be strange, for example, if a banquet committee about to

---

54 The language in the text is largely that of Professor Albert Madansky of the University of Chicago, for whose help with some of the more difficult aspects of subjective probability we are indebted.
agree to buy chicken should, on being informed that goose is also available, finally compromise on duck."

Naturalized epistemology, then, recommends considerable skepticism about Bayesianism for thinking seriously about evidence. Bayesianism is formally elegant, but of little practical use for explicating juridical matters, although it does have some value as an informal heuristic.

C. The Economic Analysis of Evidence

Another form of a priori reasoning that from time to time can run afoul of the admonitions of naturalized epistemology is microeconomic analysis of law, a point implicit in Shapira's argument "that conventional formulae of law and economics are so far removed from the practice of factfinding as to render their application to the law of evidence highly problematic, even as a normative tool." Nonetheless, Posner has recently published a wide-ranging economic analysis of the law of evidence and other

55 Savage, supra note 53, at 207. There are certain complexities here. Empirical work has demonstrated that individuals violate the sure thing principle. See Eldar Shafir, Uncertainty and the Difficulty of Thinking Through Disjunctions, 50 Cognition 403 (1994); Amos Tversky & Eldar Shafir, The Disjunction Effect in Choice Under Uncertainty, 3 Psychol. Sci. 305 (1992). Whether the textual example is a real-life example is more difficult to evaluate, for it might instead be an example of a misspecification of the probability space (by failing to take account of a known probability, for example). The upshot, however, is essentially the same whatever the example exemplifies. If the example is a violation of the sure thing principle, then subjective probability axioms are violated. If it is an example of a misspecified probability space, it demonstrates how at trial the probability space is constantly corrigible based on new information until the point of decision. That in turn means that there is no work for Bayes' Theorem to do until the point of decision, at which point the probability space is formed; but there is no work for Bayes' Theorem to do at that point either, for all evidence will be old evidence.

Another difficulty for subjective probability approaches is that preferences may not be stable over time. Alternatively, individuals may misassess their preferences. See, e.g., Daniel T. Gilbert et al., Immune Neglect: A Source of Durability Bias in Affective Forecasting, 75 J. Personality & Soc. Psychol. 617 (1998). The latter point is at the heart of the debate over euthanasia, for example, with many believing that preferences stated in advance of any particular event may not reflect an individual's views once imminently faced with that event.

56 See, e.g., Richard Lempert, Of Flutes, Oboes and the As If World of Evidence Law, 1 Int'l J. Evidence & Proof 316 (1997).

Naturalized Epistemology

Although Posner’s effort is “eclectic rather than narrowly economic...it slight[s] epistemological and other philosophical perspectives on the trial process, which seem to [him] to have only a very limited utility.” We think the good judge doth protest too much; indeed, he has it exactly backwards. The value of his article is in an inverse relationship to its reliance on a priori microeconomic reasoning. When engaging in his “eclectic analysis” from the stance of the empiricist more interested in the relationship between predictions of formal models and reality than just the predictions themselves, Posner provides some interesting discussions largely confirmatory of pre-existing scholarship in the field of evidence. When, by contrast, Posner shifts to the mode of the law and economics a priori analyst, little of value relevant to the actual operation or understanding of the legal system results. Rather, his analysis highlights the limits of this form of a priori reasoning.

The objective of this Section is to demonstrate that the utility of Posner’s analysis is directly proportional to the extent he works within the empirical tradition and inversely proportional to his reliance on a priori reasoning divorced from the relevant factual setting. We do not intend here to provide a general critique of economics, economic reasoning, or law and economics. Rather, we are critiquing the forms and applications of economic reasoning employed by Posner in this one context.

Even pursuing such limited objectives has certain difficulties, however. Foremost is that Posner’s analysis is not limited to what is

---

58 Posner, supra note 4.
59 Id. at 1479.
60 For example, his discussion of the positive social value of litigation is a helpful antidote to the argument that litigation is simply an argument over spilled milk. See Posner, supra note 4. For previous discussions, see Ronald J. Allen et al., A Positive Theory of the Attorney-Client Privilege and the Work Product Doctrine, 19 J. Legal Stud. 359, 388–96 (1990); Craig R. Callen, Adjudication and the Appearance of Statistical Evidence, 65 Tul. L. Rev. 457, 479 (1991).
conventionally thought of as the law of evidence but includes the rules of discovery and basic structural issues such as the differences between inquisitorial and adversarial proceedings. We ignore most of this for two reasons. First, the portrayal of the "inquisitorial" system bears little relationship to any existing system of which we are aware—apparently deliberately so. Posner remarks that "I wish to make the contrast between the systems as stark as possible, and so I shall treat tendencies as if they were their extremes." Perhaps he has succeeded in doing so, but to what avail is unclear. The question, one would think, is the actual operation of actual systems, not the hypothesized tendencies of hypothetical systems. Perhaps Posner's analysis might lead an investigator to hypotheses for investigation, but all it leads to here is the reiteration of well-known questions about the relative virtues of differing forms of adjudication.\(^6\) Second, Posner's analysis undervalues the extent to

---

\(^6\) Posner, supra note 4, at 1487–88.

\(^6\) It does lead him to some unconventional conclusions. For example, Posner comments that "[i]t is commonly remarked, as though the point were obvious, that the inquisitorial approach is more efficient than the adversarial approach. This article challenges that assumption." Id. at 1488 n.20 (citation omitted). We agree with the assessment of the implications of the comparison of American and some foreign systems, but Posner's argument has no more power to persuade than the American comparativists who assert to the contrary and for the same reason: Neither is exploring the actual operation of any system as it really is; both are merely expressing logical conclusions given their assumptions. The neglect of even the most basic form of empiricism—accurate description of relevant phenomena—in comparative legal scholarship is astonishing; there is typically a yawning chasm separating what comparativists writing in English say about systems and what is actually true of them. E.g., Ronald J. Allen et al., The German Advantage in Civil Procedure: A Plea for More Details and Fewer Generalities in Comparative Scholarship, 82 Nw. U. L. Rev. 705 (1988). Recently a German judge has written of the German criminal process and entitled section I of his article "The Crisis in German Criminal Procedure." He comments:

For years academic writers on German criminal law as well as the country's appellate courts have explored the possibilities of reacting to obstruction of the trial by defence counsel. The criminal justice system is said to be in overload. It is thought that a great part of this is due to dilatory and obstructive tactics of defence counsel or of the defendants themselves. Some even complain that it has become impossible to conclude a criminal trial within an adequate period of time and to reach a verdict. At least one court . . . has reacted with an act of desperation, dismissing an appeal because defence counsel had threatened to boycott [sic] the proceedings with a veritable flood of motions for new evidence. This is a result of the German inquisitorial principle, where it is not for the prosecution or the defence actually to adduce their evidence. Rather, they are required to identify it and then request the court to summon and hear
which discovery mechanisms, the structure of trial, and evidentiary rules can be independent. He thus neglects the similarities in the rules of evidence of these supposedly contrasting systems—the effects of which on adjudication, given differing formal structures, must be explained. Germany has robust privileges, for example; experts have to be qualified as experts everywhere; much of Europe employs a version of the hearsay rule in various contexts; new witnesses, procure documents, and so on. Such motions have no time limits. They may be made up to the very moment when the judge begins reading the sentence.

Michael Bohlander, A Silly Question? Court Sanctions Against Defence Counsel for Trial Misconduct, 10 Crim. L.F. 467, 468–69 (1999) (footnotes omitted). Judge Bohlander’s brief description is good evidence both of the limitations of conventional American comparative scholarship and Posner’s alternative approach. The defender of the standard law and economics approach may think we are being ungenerous in neglecting the support given to empirical conclusions consistent with our own by analyses such as Posner’s, but, risking even more the appearance of lack of generosity, we do not see such analyses as providing any support. Posner is not only talking of “tendencies” rather than the real thing; he also is talking of “tendencies” of fictional entities.


See, e.g., id. at 835–41; Allen, supra note 63, at 735–45.

Article 6 of the Convention for the Protection of Human Rights and Fundamental Freedoms provides in part: “(3) Everyone charged with a criminal offence has the following minimum rights: . . . (d) to examine or have examined witnesses against him and to obtain the attendance and examination of witnesses on his behalf under the same conditions as witnesses against him . . . .” European Convention on Human Rights, Jan. 1971, art. 6.

In a series of cases, the European Court of Human Rights has found that the failure to allow confrontation of witnesses—that is, what we would call the admissibility of hearsay—can violate the Convention. Consider this excerpt from the Kostovski case:

In principle, all the evidence must be produced in the presence of the accused at a public hearing with a view to adversarial argument. This does not mean, however, that in order to be used as evidence statements of witnesses should always be made at a public hearing in court; to use as evidence such statements obtained at the pre-trial stage is not in itself inconsistent with paragraphs 3 (d) and 1 of Article 6, provided the rights of the defence have been respected.

As a rule, these rights require that an accused should be given an adequate and proper opportunity to challenge and question a witness against him, either at the time the witness was making his statement or at some later stage of the proceedings.

Kostovski Case, 166 Eur. Ct. H.R. (ser. A) at 20 (1989) (citations omitted). But see Doorson v. Netherlands, App. No. 20524/92, 22 Eur. H.R. Rep. 330 (1996) (holding that examination of witnesses in the presence of defense counsel was sufficient). The standard assertion that the hearsay rule is not employed in inquisitorial systems is obviously false. Sometimes it is applied quite analogously to its use in the United States. In fact, the analogy is even more complete than it appears. Although in many
the European Court of Justice has found a right to silence and to be free from compelled self-incrimination implicit in the European Convention on Human Rights; and so on.

In addition to emphasizing again that we do not make here any general claims about economics or its utility in any other context, we wish further to clarify the scope of our discussion in one particular. We are assuming that Posner's "economic appraisal" of the law of evidence is directed at explanation. The issue, in other words, is not economics as a theoretical construct that has certain interesting implications if applied in a certain way to specific concepts taken from the field of evidence. Instead, the question is the actual utility of economics as employed by Posner to explain what is observed in the field of evidence.

To begin, there is much to praise in Posner's effort. It is the first attempt at a comprehensive economic analysis of evidence and will undoubtedly spur considerable work in the field. Moreover, many of his points are persuasive; indeed many of his points, as he notes, are already well-accepted in the field. We thus doubt that Posner is correct that his "conclusion will startle." His conclusion is
that the institutional and doctrinal structure of the American law of evidence has a subtle, though intuitive, implicit, and incomplete economic logic. Most evidence professors, and even a few judges, would, if asked, say that of course the American system of finding facts at trial is inefficient, ludicrously so, and redeemed if at all by the noneconomic values that the system protects.\footnote{71}{Posner, supra note 4, at 1478.}

The support for this empirical proposition is a cite to Judge Marvin Frankel's well-known lament about the adversarial system in a lecture given prior to the effective date of the Federal Rules of Evidence.\footnote{72}{Id. (footnote omitted).} Judge Frankel was not addressing the law of evidence at all, and certainly not the nonexistent Federal Rules of Evidence.\footnote{73}{Id. at 1478 n.3 (citing Marvin E. Frankel, The Search for Truth: An Umpireal View, 123 U. Pa. L. Rev. 1031 (1975)). The paper was based on Judge Frankel's lecture given December 16, 1974. Frankel, supra, at 1031 n.*. The Federal Rules came into effect in 1975.} Perhaps if the set of all comparativists, proceduralists (civil and criminal), and evidence professors was asked about Posner's conclusion, it would startle the bulk of them. By contrast, we predict that most law professors specializing in evidence would if asked say that of course the law of evidence (understood primarily as the Federal Rules of Evidence and its common-law predecessor—the entire system of litigation is another matter) has at least "a subtle, though intuitive, implicit, and incomplete economic logic."\footnote{74}{This is, to be sure, a factual issue that, to our knowledge, has not been the subject of robust empirical inquiry, and thus we remain open to the possibility of its being in error. But based on our experience of teaching and writing in the area over a twenty-five year period, the economics of trial generally and of presenting evidence specifically are standard fare in standard evidence courses, even if commonly spoken of in conventional rather than microeconomic terminology. We are unaware of any support for Posner's assertions about what evidence professors would assert.}
Many (actually, close to all, we predict) would say that the "economic logic" of some of the rules is pretty explicit—at least in the sense that they are concerned with cost-benefit relationships. Federal Rule 102 refers to construing the rules "to secure . . . elimination of unjustifiable expense and delay,"75 and Federal Rule 403 permits the exclusion of evidence on "considerations of undue delay, waste of time, or needless presentation of cumulative evidence."76 American evidence professors would also point to various notice provisions designed to forbid costly surprises at trial, such as in Federal Rules 412–15, 609, and 807.77 They would point to a series of rules excluding evidence of specific instances of conduct largely because such presentations, and the responses they would engender, are rarely worth the cost. Finally, they would mention how Federal Rules 408 and 410 are motivated in part by the desire to facilitate compromises, in large measure because compromises are more efficient than trials.78

What is noteworthy about Posner's argument is not the trivial point that one can discern some economic value from some of the rules of evidence. Rather, notwithstanding Posner's qualification that his work is eclectic, it is his relentlessness in pursuing the implications of certain formalisms that is striking. We think aspects of these portions of his argument are problematic, and we concentrate on them. In our judgment, they tend to confirm the veritistic value of the approach of naturalized epistemology and the concomitant skepticism with which implications of a priori reasoning should be approached.

1. The Implications of Rational Choice Theory

Posner's analysis relies heavily on a simplistic expected utility model of decisionmaking, which is what he means by "rational choice" in that article. According to his model, as the expected cost of an act goes up, the incidence of that act goes down in a direct relationship.79 There are more sophisticated versions of rational

---

75 Fed. R. Evid. 102.
76 Fed. R. Evid. 403.
79 See Posner, supra note 4. This is the premise of virtually his entire article.
choice, and Posner is an expert in them, but they are not employed in his article. The model of behavior he does employ is common in much a priori theorizing about human behavior in the law and economics literature, but the critical question is its relationship to reality. Although one can hardly tell from the legal literature, Posner certainly knows that economists who actually do empirical work view simple expected utility models as relatively poor predictors of behavior. For example, a review of the tax literature in the *Journal of Economic Literature* concludes that simple expected utility models are poor predictors of reality and rather imperfect guides to policy. The authors of the review suggest that the economic models, when used to advance the goal of deterrence, may not even be close approximations of reality: "This is an important area, because the econometric results to date suggest that the use of a 'stick' to enforce compliance with tax laws may not have any long-run impact." The authors conclude that there must be something other than simple expected utility that explains why people pay their taxes. They speculate that factors not considered by the microeconomists, such as morals and social dynamics, may have great impact on the economic models but point out that research exploring these factors is currently lacking.

Recently, Uri Gneezy and Aldo Rustichini empirically tested the predictions of the utility maximizer model of human behavior in a field study involving a day care, and the results were disconfirming. Parents arrived late to collect their children, which imposed costs on the school. When a money fine was introduced, the number of

---

*See, e.g., Posner, supra note 61. "Rational choice" is often employed in the legal literature as though it referred to a well-defined, unproblematic entity. In fact it is neither well-defined nor unproblematic. See, e.g., Isaac Levi, Review Essays, 97 J. Phil. 387 (2000) (reviewing James M. Joyce, The Foundations of Causal Decision Theory (1999)) (explaining some of the controversies within the field of rational choice).


*Id. at 844.

*Id. at 852.
late arriving parents actually increased significantly. Similarly, Professor Erling Eide cites multiple factors beyond the threat of economic sanction that explain behavior:

The reasons why people are more or less law-abiding are manifold [...] In criminometric studies it might be useful to distinguish between norm variables (representing desires for various courses of action), want (or taste) variables (representing preferences for various outcomes), ability variables (representing intellectual, psychic and physical characteristics), punishment variables (representing the probability and severity of punishment), individual economic variables (representing legal and illegal income opportunities), and environmental variables (other than punishment and economic variables).

From this complex array of variables, Posner has employed one—the effect of the imposition of costs upon behavior—and has addressed neither how that variable may interact with any others, nor the implications of the interactions for the law of evidence.

We see little point to this as an explanatory exercise. Rather plainly, what matters is how people and the system behave in fact, not how they are predicted to behave by the application of formal tools, no matter how elegant, for which there is substantial discomfiting, if not disconfirming, data. We respectfully suggest that, to be useful, an explication of the law of evidence must deal with the disturbances to the expected utility theory caused by the data rather than simply ignoring them without explanation. Posner’s article is

---

87 Erling Eide, Economics of Criminal Behavior: Survey and Bibliography, at 21–22 (U. of Oslo, Inst. for Private Law, Law and Economics Working Paper No. C 5). Other empiricists concur. In his 1990 book, Why People Obey the Law, Professor Tom R. Tyler reports the results of a study testing the contribution of various variables to the decision of individuals to obey the law, including the deterrent efficacy of the law. As he summarizes the data:

The findings of the Chicago study also support the suggestion that the influence of deterrence on compliance may be overrated. The Chicago study used an approach to measurement patterned after that of Paternoster and a similarly designed panel study, and found little evidence of deterrence effects. Although the study does not question the assumption that deterrence works, other studies may well.

not very helpful to the extent it rests without explanation upon an overly simplistic version of rational choice theory.  

2. Bayes' Theorem and the Meaning of Relevance

Posner cautiously suggests Bayes' Theorem may be a useful heuristic in analyzing juridical proof, a point that others have made before him and with which we agree. As much more than a casual heuristic, however, Bayes' Theorem has little to recommend it in the juridical context, as we have previously discussed. Posner also endorses the use of Bayes' Theorem to explicate the meaning of relevancy, developed in the legal literature by Professor Richard Lempert.  

As elegant as the Bayesian theory of relevance is, naturalized epistemology reminds us to ask how it comports with the facts about human reasoning—both what people actually do and what they can do. An obvious truth (obvious by both analysis and inspection) is that virtually all evidence is highly contingent. Accordingly, a useful likelihood ratio cannot be formed to test the relevancy of a piece of evidence unless all the other pieces of evidence, as well as how they all interact, are already known. That is

---

88 Curiously, one of the strongest proponents in the legal literature of the theory of behavior that Posner employs in his evidence article, Professor Steven Shavell, apparently has conceded when writing in an economic journal that the theory he propounds so vigorously in the legal literature does not accurately portray the reality of deterrence and that more study on other variables is needed. Shavell, with Professor A. Mitchell Polinsky, writes:

We also have not discussed social norms as a general alternative to law enforcement in channeling individuals' behavior. By a social norm, we mean a rule of behavior (for example, that people should not litter or should not discriminate on the basis of race) whose violation may have the following consequences: the violator may experience an internal sanction (guilt, remorse); others may impose on the violator external, extra-legal social sanctions (gossip, ostracism); and others may experience utility or disutility from punishment of the violator. There is an emerging literature on social norms that seems worth amplifying because of the influence that social norms have on behavior, because of their role as a substitute for and supplement to formal laws, and also because of the possibility that laws themselves might influence social norms.


89 Kaplan, supra note 3, at 1083–86; Lempert, supra note 3, at 1022–27.

90 See supra Section III.B.

91 Posner, supra note 4, at 1522 n.95 (citing Lempert, supra note 3, at 1025).

92 See Allen, supra note 49, at 270.
why likelihood ratios are not discussed after evidence is produced. Instead substantive arguments that describe how the bit of evidence in question will be integrated into the fabric of the party's story are discussed. Posner recognizes the point—"Evidence that is cumulative must be distinguished, however, from evidence necessary to complete a mosaic of proof. A costly bit of 'additional' evidence might be cost-justified because it fits in with other evidence to establish the truth convincingly."93—but he confuses the rule with the exception. The issue of relevancy generally is whether evidence fits into a "mosaic of proof"; this is not limited to the issue of cumulative evidence. Obversely, only in exceptional cases can a plausible case be made for testing admissibility by a Bayesian likelihood ratio.94

By focusing on the formalism of Bayes' Theorem, Posner's analysis also misses the deeper significance of the relevancy rules. The relevancy rules, unlike the formal Bayesian analysis, recognize that relevancy determinations require extensive substantive knowledge of the entire case to be made intelligently and thus cannot be reduced to algorithms like Bayes' Theorem. That means that relevancy determinations are very difficult to make intelligently prior to possessing that knowledge. Even with all the evidence, people can reasonably disagree about whether any particular piece of evidence rationally fits into a "mosaic" or not. Thus, relevancy rulings must be made cautiously and contingently, and they must be made in a manner respectful of the bifurcated nature of factfinding when a jury is involved. They must be made, in other words, along the lines that Federal Rule 104(b) lays out—a rule that makes considerable sense viewed as instantiating the "mosaic" view of relevancy but that bears a Bayesian interpretation only awkwardly.95

---

93 Posner, supra note 4, at 1524.
94 DNA evidence is the contemporary standard example, but it is a complete mystery whether DNA evidence can be incorporated algorithmically into trials in a manner that increases the accuracy of decision. "Algorithmically" here is an important qualifier. Obviously DNA evidence can easily be a primary determinant of the truth of competing stories, but for such a purpose no formal analysis of the type discussed here need be employed.
95 The definition of relevancy in Federal Rule 401 can be read as consistent with a Bayesian test for relevancy, but it is also consistent with any imaginable rational test for relevancy as well.
3. Rootless Theorizing

Another concern about Posner's economic approach to the law of evidence can be summarized in the phrase "rootless theorizing." The results of formal systems are dependent upon their axioms and rules of deduction. The relationship between the results of deductions and reality, however, depends on the truth of the axioms and the nature of the rules of deduction. Posner's focus on the logical implications of formalisms may deflect consideration from the truth content of his larger enterprise. We give two examples of this.

First, in summing up his comparison of different systems of adjudication, Posner notes that in the United States:

[T]he conviction rate is lower in bench trials than in jury trials. This is significant because in most states the decision in a criminal case as to whether to be tried by a judge or by a jury is entirely the defendant's. If juries are less accurate guilt determiners than judges, innocent defendants will choose to be tried by judges rather than run the risk of jury mistake, while guilty defendants will choose to be tried by juries, hoping for a mistake. The acquittal rate should therefore be higher in bench trials—and it is.96

This appears to be empirical vindication of the economic model, but it all rests upon the assertion that "in most states the decision in a criminal case as to whether to be tried by a judge or by a jury is entirely the defendant's." This is false. Posner has provided an economic rationale for an incorrect proposition.97

At present, some states provide for a defendant's unilateral right to a bench trial. Other states require prosecutorial and court consent. Still others allow for a defendant to waive a jury trial in all but capital cases or cases where the death penalty is sought. In some jurisdictions, the court must consent to the defendant's waiver. In other states, the court accepts the defendant's waiver only upon consent of the Government. In

96 Posner, supra note 4, at 1501 (footnote omitted).
97 Posner has said in his helpful and generous comments on a draft of this article that this is not an "economic" argument. We so classify it because it is a standard application of the simple expected utility model that is at the core of Posner's "economic" arguments about evidence. If we misclassify it, it remains nonetheless an interesting example of rootless theorizing, which in any event is our main subject in this Section.
one state, Ohio, if the defendant's jury waiver is proposed either shortly before or during the trial, the trial judge and prosecutor must consent. One other state, North Carolina, does not appear to permit the accused to waive a jury trial in a felony case under any circumstances.\n
---

\*Adam H. Kurland, Providing a Federal Criminal Defendant with a Unilateral Right to a Bench Trial: A Renewed Call to Amend Federal Rule of Criminal Procedure 23(a), 26 U.C. Davis L. Rev. 309, 321–23 (1993) (footnotes omitted). We have updated Kurland's research on a defendant's right to waive a jury trial in all fifty states and D.C.


Ten states require only court consent before allowing defendants to waive a jury trial. See McCorquodale v. State, 211 S.E.2d 577, 581–82 (Ga. 1974); Palmer v. State, 25 S.E.2d 295, 300–01 (Ga. 1943); Haw. Rev. Stat. Ann. § 806-61 (Michie 1999); Haw. R. Penal Proc. 23(a); Me. Rev. Stat. Ann. tit. 15, § 2114 (West 1980); Me. R. Crim. P. 23(a); Mass. Gen. Laws Ann. ch. 263, § 6 (West 2000) (incorporating 1992 amendment to the effect that "consent to ... waiver shall not be denied in the district court or the Boston municipal court if the waiver is filed before the case is transferred for jury trial to the appropriate jury session," as long as when there is more than one defendant, all
Posner’s argument, in addition to resting on an apparently false premise, demonstrates the manipulability of formal arguments. An equally plausible variation of his argument can be made that the evidence of higher acquittal rates by judges demonstrates that juries are, and are believed to be, more, not (as in Posner’s original argument) less, accurate decisionmakers. If juries are believed to be more accurate decisionmakers, innocent parties will choose juries, but prosecutors will read the signal and dispose of many cases of innocent defendants in one manner or another (such as dismissal or really good plea bargains). Pre-trial proceedings will take a large proportion of the innocent defendants who would have had jury trials out of the system, leaving a much higher proportion of guilty defendants going to jury trials who are subsequently found guilty by highly accurate juries. A high proportion of guilty individuals hoping for a mistake will also be tried by judges, and their lower conviction rate proves that judges are less accurate decisionmakers than juries. Is any of this true? Who knows? Knowledge is not advanced by this kind of reasoning.

The second example is Posner’s discussion of what is known as the Blue Bus hypothetical and whether naked statistical evidence suffices for a verdict. Posner constructs an economic explanation premised on the assertion that “[t]he law’s answer is ‘no.’” Posner relies on some of the evidentiary literature for this conclusion and on the case of Smith v. Rapid Transit. Unfortunately,

Smith is difficult to view as a “statistical evidence” case....
The plaintiff did not rely on any such evidence. She merely as-
serted that she was forced off the road by a bus and in addition proved that Rapid Transit, Inc. was the only bus company operating regularly on the road where the accident occurred. In appraising the strength of the evidence, the court concluded that it was a matter of "conjecture" who owned the bus and that "[t]he most that can be said of the evidence in the instant case is that perhaps the mathematical chances somewhat favor the proposition that a bus of the defendant caused the accident. This was not enough."

The language of the case is hardly the language of rejection of the adequacy of statistical proffers. The "perhaps" alone should be sufficient to make that point clear. In any event, cases raising clean issues of naked statistical evidence are rare, perhaps even nonexistent, but the impression from the cases is inconsistent with Posner's assumption. An example is the United States Court of Appeals for the Ninth Circuit's opinion in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, on remand from the Supreme Court:

Plaintiffs do not attempt to show causation directly; instead, they rely on experts who present circumstantial proof of causation. Plaintiffs' experts testify that Bendectin is a teratogen because it causes birth defects when it is tested on animals, because it is similar in chemical structure to other suspected teratogens, and because statistical studies show that Bendectin use increases the risk of birth defects. Modern tort law permits such proof, but plaintiffs must nevertheless carry their traditional burden; they must prove that their injuries were the result of the accused cause and not some independent factor. In the case of birth defects, carrying this burden is made more difficult because we know that some defects—including limb reduction defects—occur even when expectant mothers do not take Bendectin, and that most birth defects occur for no known reason.

California tort law requires plaintiffs to show not merely that Bendectin increased the likelihood of injury, but that it more likely than not caused their injuries. In terms of statistical proof, this means that plaintiffs must establish not just that their mothers' ingestion of Bendectin increased somewhat the likeli-

102 43 F.3d 1311 (9th Cir. 1995).
hood of birth defects, but that it more than doubled it—only then can it be said that Bendectin is more likely than not the source of their injury. Because the background rate of limb reduction defects is one per thousand births, plaintiffs must show that among children of mothers who took Bendectin the incidence of such defects was more than two per thousand.\footnote{Id. at 1320 (footnote and citation omitted). Another example is \textit{Kramer v. Weedhopper of Utah}, 490 N.E.2d 104 (Ill. App. Ct. 1986), which involved a bolt that allegedly sheared off, causing harm. Ninety percent of the bolts were supplied by defendants; ten percent were supplied by another supplier, not part of the litigation. Id. at 106. No other evidence could identify which supplier supplied the bolt in question. The trial court entered summary judgment for the defendant; the appellate court reversed, holding that this presented a sufficient case to go to the jury. Id. at 105, 107-08. Whether a relative risk approach to the meaning of a preponderance of the evidence is sensible is a different question. See, e.g., Sander Greenland \& James M. Robins, Epidemiology, Justice, and the Probability of Causation, 40 Jurimetrics J. 321 (2000).}

Posner's economic argument again appears to be founded on a misconception of the law—the most that can be said of the cases favorable to his argument is that the adequacy of a naked statistical case is an open question—and thus he has constructed an economic rationale for an arguably nonexistent entity, demonstrating the noteworthy flexibility but not the veritistic prodigiousness of his technique.

Posner concludes this section of his article with another economic argument. This is based on the hypothetical Blue Bus case, where the plaintiff brings evidence that he was hit by a bus, and fifty-one percent of the buses on this route were owned by company \textit{A} while forty-nine percent of the buses were owned by company \textit{B}.\footnote{Posner, supra note 4, at 1508.}

There is still another objection to allowing the bus case to go to the jury. If \textit{B}, though responsible in fact for almost half the accidents, is never held liable and \textit{A} is always held liable, \textit{A} will have a big incentive to be careful and \textit{B} little or no incentive to be careful. As a result, over time, more than half the accidents will be caused by \textit{B}, increasing the error rate resulting from allowing juries to base decisions on the ratio of the companies' buses on the route in question. Eventually, \textit{A}, having higher li-
ability costs, will probably withdraw from the route; the rule on burden of proof will have created a monopoly!\textsuperscript{105}

Our doubts about the value of Posner's form of a priori reasoning for explicating the law of evidence are well captured by this paragraph. Taking the argument on its own terms—terms which we reject, of course—the consequence Posner fears will occur only if $A$ is remarkably stupid. In the world Posner is hypothesizing, all $A$ need do is take at most three buses out of service. If $A$ takes three buses out of service, the ratio of buses now favors $A$ (it changes from 51/49 to 48/49), and $B$ will henceforth be held liable for all accidents.\textsuperscript{106} Of course, another economic fear might now arise: $B$, also not being massively stupid, might respond by taking two of its buses out of service. Perhaps the prediction would now be that "bus reduction" would become recursive, eventually resulting in only one bus company with only one bus driven by a very careful driver, which would surely not serve the needs of the community! Therefore, the logic would go, allowing verdicts based on naked statistics is not economically sensible.

Neither a monopoly nor the essential elimination of the industry would result from allowing probabilistic verdicts. If such verdicts began to accumulate inaccurately against bus companies, they would invest in precautions, many of which are cheaply available.

Posner might respond that we are making economic arguments: Both error reduction and subjective expected utility may be advanced by permitting decision on naked statistics, and individuals will intelligently assess the value of investments in precaution. Our claim, however, is not that there is no role for policies concerned with errors and costs, nor is it that incentives are irrelevancies. Rather, the point is to critique a form of a priori reasoning that is curiously out of touch with the phenomenon supposedly under investigation.\textsuperscript{107}

\textsuperscript{105}Id. at 1510.

\textsuperscript{106}Or, if company $A$ is really smart, it will take exactly two buses out of service. If both companies have forty-nine buses in service, the probability of liability would be exactly 0.5, meaning plaintiffs injured by buses could never recover.

\textsuperscript{107}The economics analysis is not redeemed by the argument that some economists, including Posner, make to the effect that empirical adequacy of a model or its assumptions does not matter; only the predictions of the model do. We have attempted to show that the tools employed by Posner permit virtually any prediction
More importantly, as employed by Posner in this context, the technique seems curiously out of touch with the essential justification for the technique in the first place. Apart from aesthetics, the primary value of any formalization is its capacity to generate true (or, within pure mathematics, valid) answers. A "formalization" that can be used to justify inconsistent states of affairs is, in our view, not a "formalization" at all. In any event, it serves no obvious purpose. Typically an argument or approach is less, not more, valuable if more aspects of different problems can be defended or explained with it. At the limit, an argument that explains everything explains nothing. This is the root of the common complaint that law and economics work tends toward being ad hoc—a complaint that, regardless of the general utility of microeconomics for explicating the law, is borne out by much of Posner's argument here. We do not mean by this to consign economics, so far as the law of evidence is concerned, to the trash bin of history. This would be an eminently ridiculous proposition. Costs, benefits, and incentives are obviously material concerns to the structuring of dispute resolution. As we said at the beginning of this Section, our point is considerably narrower—to analyze the contributions of this one expression of economic analysis.

D. The Relative Plausibility Theory and Naturalized Epistemology

Not all theorizing about evidence is a priori. One example is the relative plausibility theory that was constructed in response to the empirical and analytical inadequacies of the expected utility and Bayesian approaches. The critical insight of the relative plausibility theory is that legal fact finding involves a determination of the comparative plausibility of the parties' explanations offered at trial rather than a determination of whether discrete elements are found

---


to a specific probability. In civil cases the factfinder is to identify the most plausible account of the relevant events, whereas in criminal cases the prosecution must provide a plausible account of guilt and show that there is no plausible account of innocence. The structure of liability is provided by the formal elements, but that is different from the proof process, which proceeds in a largely comparative fashion over the stories advanced by the parties. Once the most plausible account of the relevant events is determined, liability flows deductively from the formal structure of the law. The relative plausibility theory as developed in the legal literature bears a close relationship to the empirical work on jury decisionmaking done by Professors Nancy Pennington and Reid Hastie. It also bears a close relationship to the work done on hypothesis comparison through the use of connectionist approaches, such as in the work of Professor Paul Thagard. From the perspective of naturalized epistemology, there are several advantages of the relative plausibility theory including:

- It appears to explain what factfinders actually do.
- It unmistakably explains what advocates actually do at trial.
- It avoids the formal difficulties of Bayesianism, as it has so far been developed in the literature.
- The paradoxes of proof are marginalized because they are distributed evenly over both sides of a dispute.
- Computational complexity is largely eliminated as a problem because litigation focuses on the plausibility of coherent stories advanced by the parties rather than on discrete items of evidence.

In addition to better explaining the basic structure of trials than its competitors, the relative plausibility theory also produces falsifiable predictions, a number of which are suggested above. One prediction is that actual litigation proceeds in a comparative fashion. Support for this prediction of the relative plausibility theory is ubiquitous in the case law. Examples could be listed endlessly be-

---

12 For an extended defense of the relative plausibility theory, see Allen, supra note 46; Allen, supra note 109.
cause virtually all trials reduce to the comparison of competing claims. For example, in *MCI Communications Corp. v. American Telephone & Telegraph Co.*, the court wrote: “Considering the trial evidence and argument as well as the instructions tendered, the jury’s obvious choice was between a nationwide market (espoused by MCI) or a more limited market (advocated by AT&T).” The jury’s “obvious choice” was not between whether the plaintiff had proved its allegations by a preponderance of the evidence or not but between the two markets espoused by the parties. *Swajian v. General Motors Corp.* dealt with whether an accident was caused by a faulty axle, as alleged by the plaintiff. The defense was driver error as a result of intoxication. The trial court excluded the defendant’s evidence of intoxication. On appeal, the issue was not presented as whether the excluded evidence was admissible to demonstrate that causation had not been proven by a preponderance of the evidence but, instead, whether the evidence was admissible as proof of an alternative story:

Armed with this evidence, the jury could have concluded that driver error contributed significantly to, if not caused, decedent’s accident. As it was, the jury was presented with the following factual scenario: the two month old vehicle was travelling down a straight, flat road in good weather when it swayed and went out of control for no apparent reason. The only explanation proffered was that there was a defect in one of the axles. Without the evidence of intoxication the jury was left with no reason for the loss of control other than Swajian’s allegations.

---

113 708 F.2d 1081 (7th Cir. 1983).
114 Id. at 1174.
115 916 F.2d 31 (1st Cir. 1990).
116 Id. at 34. Still another example is *Wyletal v. United States*, 907 F.2d 49 (7th Cir. 1990):

Josephine Wyletal, a lively eighty-five year old widow with a cataract in her left eye, was walking eastbound on the north side of Oakton Street in Skokie, Illinois at 11:00 a.m. on a cloudy day in November. She was not wearing her glasses. At the same time, a letter carrier for the United States Postal Service was hurriedly delivering mail by foot heading west on the north side of Oakton Street. As Mrs. Wyletal proceeded along the sidewalk, the postal carrier emerged from a recessed doorway. Neither saw the other and they collided [resulting in damages]....

...
Inconsistent pleading cases also provide support for the relative plausibility theory by typically instructing the jury to decide which story is most plausible. In *McCormick v. Kopmann*, the court upheld the trial judge who had sent a case to the jury containing inconsistent claims. In one claim, the plaintiff alleged that one defendant was liable for having killed her husband and that the husband was not contributorily negligent as a result of being drunk (she alleged he was sober). In another claim against different defendants, she alleged liability under the Dramshop Act for having sold sufficient alcohol to her husband to render him intoxicated. The jury was essentially instructed to return a verdict against the party—plaintiff or either defendant—most likely liable for the event, just as the relative plausibility theory would predict.

In his economic critique of evidence law, Posner adopts aspects of the relative plausibility theory, although he implicitly rejects it in one particular that bears upon cases like *McCormick*. He asserts that if

\[ \text{[T]he plaintiff's story had a probability of .42 of being true, the defendant's story a probability of .30 of being true, and the probability that another story or stories is true was .28, then the plaintiff should lose because he has failed to prove that his story is more likely than not true.}\]

Applied to *McCormick*, this means that the plaintiff would have to show one of the defendants to be more likely than not liable. This is wrong, as the case at least implicitly recognizes. It is instructive to explicate the error, however.

One either knows or does not know the implications of the story or set of stories comprising the missing 0.28 probability. If these implications are known, each party should get the benefit of the

---

During the bench trial, Mrs. Wyletal and the postal carrier presented conflicting testimony as to how the collision occurred. Mrs. Wyletal testified she was hit from the back, while the postal carrier testified they collided head-on.

Id. at 49-50.
118 Id. at 725, 732.
119 Id. at 724.
120 Id.
121 Id. at 725.
122 See Posner, supra note 4, at 1512 & n.74.
123 Id. at 1513.
124 See supra notes 117-21 and accompanying text.
probability associated with the story or stories that favor them. If the implications are not known, there is no good reason to systematically disfavor plaintiffs by attributing all the ambiguity to them. In civil cases, given mutual discovery, the parties can be expected to search for and produce evidence of whatever stories they think can plausibly support their legal claims. Indeed, the only argument for systematically disfavoring plaintiffs is an unpersuasive one that posits plaintiffs will bring actions where the probability of their story is extremely low but where defendants cannot respond. This conceptual problem has no obvious empirical counterpart. If a plaintiff can make a plausible claim, even if in the abstract low-probability case, and a defendant simply has no response, under the present rules the plaintiff is likely to win. In any event, ignoring the ambiguity that no party, remember, wishes to litigate will advance all theories of trials (for example, risk reduction, optimizing expected returns, fairness). The inconsistent claims cases implicitly recognize this point.

Posner's argument that all ambiguity should be allocated against plaintiffs and the state is another interesting example of how apparently straightforward, logical, perhaps "economic" arguments are often unresponsive to the actual conditions about which they purport to be theorizing. Plaintiffs could not possibly establish that of all the ways that the universe might have been on the day and at the place in question, half plus one favor liability. The analogous requirement for criminal cases is also impossible to establish. Take as an example any criminal case in any populated area, such as the O.J. Simpson murder case. There were roughly seven million people in the greater Los Angeles area on the night in question. Were Posner's argument an accurate explication of the law, all Simpson's

---

126 This is also one of two reasons why ambiguity should not be assessed against the state in criminal cases: First, the state is not likely to go around looking for false criminal charges that it can prosecute solely because the wrongly accused defendant will not have evidence supporting his innocence. The second reason is that in criminal cases a low probability of guilt will rationally give rise to the belief that some other explanation than the guilt of the defendant is true, and therefore the defendant should be acquitted, a point that can be brought out by counsel. The relative plausibility theory is not falsified by analogous factfinding in civil cases, either, as the factfinder's own knowledge and experience is relevant to fashioning the possible stories explaining the litigated events.

127 Posner, supra note 4, at 1513.
counsel would have had to do is present to the jury with the phone book of the relevant area and put it to the prosecution to eliminate all these alternative hypotheses. Unless one knows—which here means has sufficient evidence to establish—that the probabilities of these alternative hypotheses are zero, each must count in the defendant’s favor. It would be quite astonishing if the cumulative probability of approximately seven million low-probability events does not equate with reasonable doubt.  

That defense attorneys do not pursue the logical implications of Posner’s argument strongly suggests that his argument is false as a proposition about the phenomenon under investigation. By contrast, it is good evidence in favor of the factual accuracy of the relative plausibility theory and in favor of the idea that the juridical world is deeply comparative in the sense advanced by the relative plausibility model.  

Various decisions of the Seventh Circuit, including three opinions authored by Judge Posner, are consistent with the relative plausibility theory. First, *Spitz v. Commissioner* involved a prosecution for tax fraud. The trial court was unimpressed with the taxpayers’ expla-

---

127 Under the relative plausibility theory, by contrast, the question is whether any of the individuals in the area plausibly committed the crime. In the absence of evidence that any one did, it would not be plausible. This highlights another difference between the relative plausibility theory and the Bayesian approach—a difference that corrects a false conventional belief about criminal trials. The conventional belief is that defendants do not have to raise alternative theories or provide evidence. While formally this is true, functionally it is false. The reason it is false is explained by the relative plausibility theory. If theories supported by evidence are not presented to the factfinder, they exist only to the extent they preexist in the factfinder’s mind.

128 For discussions of these and many related issues, see sources cited supra note 109; Ronald J. Allen et al., Probability and Proof in *State v. Skipper*: An Internet Exchange, 35 Jurimetrics J. 277 (1995). It also bears noting that Posner’s qualification of the relative plausibility theory in the case of multiple possible explanations eliminates the legal significance of the theory. In any case in which the probability of the plaintiff’s case does not exceed 0.5, the total probability of the alternative explanations, whether advanced by the defendant or not, must meet or exceed 0.5, resulting in a decision for the defendant. It is just this manner of viewing the preponderance of the evidence standard that the relative plausibility theory critiques and for which it provides a substitute. In any event, nothing turns on whether there is any legal significance to the theory under Posner’s qualification; the proper issue is its empirical validity.

129 954 F.2d 1382 (7th Cir. 1992).

130 Id. at 1383.
In reversing, the Seventh Circuit, per Judge Posner, commented that "in general and in this instance the plausibility of an explanation depends on the plausibility of the alternative explanations. However implausible the [defendants'] explanation might seem in isolation, it does not stand alone, but must be compared with the government's alternative explanation . . . ." \[132\]

Second, in Brackett v. Peters,\[133\] the defendant was charged with felony murder following the rape and beating of the victim.\[134\] While in the hospital as a consequence of the assault, the eighty-five year-old victim's physical and mental condition deteriorated, and she eventually died, having been asphyxiated by some food lodged in her trachea.\[135\] The question on appeal was whether a rational finder of fact could have found that the defendant caused the death of the victim.\[136\] In affirming the conviction, Judge Posner, writing for the court, commented that:

\[T]here are dangers in inferring consequence from sequence. But they are slight when as in this case the event not only follows the act closely in time but is the kind of event frequently produced by the kind of act, and no persuasive evidence of an alternative causal sequence is presented . . . .\[137\]

Third, in United States v. Morales,\[138\] the court, per Judge Posner, reversed a conviction for firearm possession and remanded for a new trial,\[139\] asserting that:

The issue becomes: not was it highly likely beforehand that a sequence such as that described by Officer Maher would actually occur, but, given that the gun and clip were found in the sinks, was the prosecution's hypothesis as to how they got there

\[132\] Id. at 1384.
\[133\] Id. at 1384–85 (citations omitted).
\[134\] 11 F.3d 78 (7th Cir. 1993).
\[135\] Id. at 79.
\[136\] Id.
\[137\] Id.
\[138\] Id. at 80.
\[139\] 902 F.2d 604 (7th Cir. 1990).
\[159\] Id. at 609.
substantially more probable than the hypothesis that someone other than Morales put them there?140

Two instructions typically given to juries better support the relative plausibility theory than the elements approach. The first is that jurors are to rely on their common sense,141 but doing so would mean the jury would disregard the judge's instructions to find elements by a preponderance of the evidence and focus instead on the competing claims of the parties. The other instruction is that jurors are not to draw inferences until all the evidence is in.142 This instruction is a striking embarrassment to a Bayesian understanding of the structure of litigation. The factfinders are explicitly instructed to do the opposite of what the Bayesian argument requires. By contrast, this instruction is not at all in tension with relative plausibility theory. Under the relative plausibility theory, the objective is to test the explanatory power of the stories of the parties, which might be put at risk by reaching conclusions too hastily. The contrast between Bayesian and relative plausibility explanations of juridical proof could not be more stark. This instruction is devastating to the Bayesian arguments but is easily explained by relative plausibility.

The data provided here are admittedly anecdotal. Nonetheless, it is obvious both that there are considerable data supporting the relative plausibility theory and that the theory could be falsified by well-formulated studies.

The relative plausibility theory also explains many discrete aspects of the rules of evidence, a point deserving some elaboration. Many aspects of trial implicitly embrace the relative plausibility theory in order to advance the veritistic consequences at the heart of naturalized epistemology. For example, various rules of completeness override technical regulatory or exclusionary rules of evidence. These rules provide data to factfinders in conventional story form by admitting surrounding material relevant to specific

140 Id. at 608. For examples in opinions authored by other judges, see, e.g., United States v. Crosby, 75 F.3d 1343, 1347 (9th Cir. 1996); Wyletal v. United States, 907 F.2d 49, 50 (7th Cir. 1990).
141 E.g., Comm'n on Pattern Jury Instructions, Dist. Judges Ass'n, 5th Cir., Pattern Jury Instructions (Civil Cases) § 3.1, at 31 (1999).
142 E.g, Comm'n on Pattern Jury Instructions, Ass'n of Supreme Court Justices, 1 New York Pattern Jury Instructions—Civil ¶ 1.11, at 16 (1974).
testimony. One example is Federal Rule 106: "When a writing or recorded statement or part thereof is introduced by a party, an adverse party may require the introduction at that time of any other part or any other writing or recorded statement which ought in fairness to be considered contemporaneously with it." Another example is the res gestae rule, found in many jurisdictions, which permits virtually any background matter to the litigated question to be adduced in order to flesh out the relevant events, regardless of the technical admissibility of the background material. Still another example is the standard practice of trying conspirators jointly "to give the jury a fuller picture of the scheme." The list continues. Motive is never an element but always admissible in criminal cases, and for that matter in civil cases, and convictions are difficult to obtain without proof of motive. Motive fills in the gaps and converts formal structures into human events. Federal Rule 612 provides that, "if a witness uses a writing to refresh memory for the purpose of testifying," the writing is admissible regardless of exclusionary rules. Again, the result is to put an entire human episode before the factfinder. Opening statements and closing arguments are also more consistent with the relative plausibility theory than the elements model. Opening statements inform the factfinder of the story to be told, and closing arguments attempt to demonstrate that one story is more plausible than its competitors.

Some rules of evidence may appear inconsistent with the relative plausibility model but in fact are not. For example, exclusionary rules keep evidence from the jury, retarding the emergence of the full picture, but there are few general exclusionary rules. Apart from the constitutional exclusionary rules whose purpose is to vin-

---

143 Fed. R. Evid. 106.
146 Fed. R. Evid. 612.
147 Another example of the completeness principle is the commonly-held belief among criminal practitioners that defendants generally must testify to stand a reasonable chance of acquittal. If the person most knowledgeable about the truth of the state's case chooses not to testify, the inference is virtually irresistible that the state's case is accurate. It is just this point that makes the admission of prior records so controversial. Defendants must testify, but admitting prior records may engender the sense that this particular defendant is someone who should be kept off the streets, regardless of whether he committed this particular act. An error of a wrongful conviction, in other words, is not much to be regretted.
dictate rights, there are only two general exclusionary rules: relevancy and hearsay. Relevancy exclusions do keep information from juries, but only that information that no person could reasonably rely upon or whose "danger of unfair prejudice, confusion of the issues, or misleading the jury" substantially outweights its probative value. This is a quite liberal standard for admission, and thus quite supportive of the relative plausibility theory. Moreover, to the extent the relevancy determinations exclude evidence, they do so based on a veritistic rationale that the relative plausibility theory also honors.

The treatment of hearsay under the Federal Rules is also consistent with the relative plausibility theory. The most important forms of hearsay, admissions and prior statements, are largely excluded from the definition of hearsay or are admitted under exceptions. In addition, there are numerous explicit exceptions—twenty-eight in the Federal Rules of Evidence—and a residual exception. Moreover, the previously mentioned res gestae principle acts as a catch-all rule of admission for many hearsay statements. The hearsay rule keeps only the runkest and least reliable form of evidence from the factfinder, which is quite consistent with the relative plausibility theory and its veritistic implications.

Privilege rules do keep information from the jury, but even here the inconsistency with our basic thesis is not striking. Most privilege rules are designed to sacrifice truth-seeking to other values, and thus they are indifferent to the comparison of the conventional view of the proof rules and the relative plausibility theory. Moreover, an important privilege—the attorney-client privilege—is best understood as promoting rather than retarding the production of information.

148 See, e.g., Fed. R. Evid. 401; Mueller & Kirkpatrick, supra note 78, § 4.2.
149 Fed. R. Evid. 403.
150 See, e.g., Fed. R. Evid. 801(d).
151 See Fed. R. Evid. 803, 804, 807.
152 In fact, there are even more hearsay exceptions. See, e.g., Fed. R. Evid. 703 (allowing certain hearsay that is the basis of expert testimony). See generally Ronald J. Allen, The Evolution of the Hearsay Rule to a Rule of Admission, 76 Minn. L. Rev. 797, 800 (1992) (concluding that the hearsay rule has been largely eroded by the various exceptions).
153 See Allen et al., supra note 60, at 361-62. The relative plausibility theory also explains cross-examination. Cross-examination brings out more information and thus
One set of rules—the burden of proof rules—is no more consistent with the relative plausibility theory than its competitors. Like expected utility and Bayesian approaches, the relative plausibility theory predicts that elements will not be the focus of decision at trial. Both expected utility and Bayesian approaches, however, view the question at trial as liability or not. The relative plausibility theory predicts the focus at trial will be competing accounts of what occurred. For the reasons already advanced, the relative plausibility theory does not suffer from the logical problems afflicting its competitors, and considerable evidence of its accuracy can be found in the cases. By contrast, virtually no such evidence can be found for the other theories.

IV. NATURALIZED EPISTEMOLOGY AND THE SPECIFIC RULES OF EVIDENCE

For any rule of evidence that has a veritistic rationale, we can properly analyze that rule from the standpoint of social epistemology. That means, of course, asking an essentially empirical question: Does this rule of inclusion or exclusion in fact increase the likelihood that factfinders, given what they are actually like, will achieve knowledge about disputed matters of fact? For ease of reference, let us paraphrase this as asking: Does this rule of inclusion or exclusion maximize veritistic value? This, in fact, is precisely the question we take many evidence scholars to be asking already—albeit not framed in these precise terms, and albeit not by those scholars attracted to the a priori formalisms discussed in Part III. It is part of the argument of this paper, however, that this is the question all evidence scholarship should be asking.

Many rules that on their face invite one kind of veritistic analysis require a very different kind in practice. For example, Federal Rule 404, on its face, excludes character evidence in most contexts, though, in fact, the exception in Federal Rule 404(b) largely swallows the rule. Thus, while it might seem that we should ask whether excluding character evidence maximizes veritistic value, the real question is whether admitting it does. The same may be said for the hearsay rule. Although on its face, the hearsay doctrine facilitates the construction of stories, even if not quite the one to which a particular party is attached.

\[\text{facilitates the construction of stories, even if not quite the one to which a particular party is attached.}\]

\[\text{154 See infra Section IV.C. (discussing Fed. R. Evid. 404).}\]
is a rule of exclusion, in reality it is a rule of admission: What the advocate must really know is how to get the proffered hearsay admitted under one of the multitude of exceptions to the nominal rule of exclusion. Thus, the pertinent veritistic question concerns the veritistic credentials of the grounds on which hearsay is admitted, rather than the veritistic reasons for excluding it in most cases.

Continuing this theme, all of the following questions would fall within the purview of the social epistemology approach to evidence:

- What standards of relevance (under Federal Rule 402) and their implementation (under Federal Rule 104) would in fact maximize veritistic value?
- Are out-of-court statements based on present sense impressions (which are admissible under Federal Rule 803(1)) in fact more reliable, such that they do not need to be tested with cross-examination under oath and with the benefit for the trier of fact of evidence of demeanor? What about out-of-court statements made for the purpose of medical diagnosis or treatment (admissible under Federal Rule 803(4))? What about the other hearsay exceptions?
- What kinds of statistical and probabilistic evidence are in fact probative without being confusing and misleading (and so are admissible under Federal Rule 403)?
- How can statistical evidence be integrated with unquantified evidence?
- What standard for the admissibility of scientific evidence (under Federal Rule 702) will in fact maximize veritistic value (taking account of the cognitive limitations of both judges who must apply the standard and jurors who must weigh the evidence)?

---

155 See Fed. R. Evid. 802; Allen, supra note 152.
156 Fed. R. Evid. 402.
157 Fed. R. Evid. 104.
158 Fed. R. Evid. 803(1).
159 Fed. R. Evid. 803(4).
160 Fed. R. Evid. 403.
161 Fed. R. Evid. 702.
What is the proper role for instructions on inferences, presumptions, and comments on or summary of the evidence?

Should jurors be instructed to take (or refrain from taking) notes, to discuss the evidence among themselves prior to the close of evidence at trial, to reach tentative conclusions, or to ask questions at trial of witnesses, parties or lawyers?

In fact, however, all these questions central to the social epistemology enterprise present a threshold issue: To what extent is the available empirical evidence probative as to expected veritistic values in the context of real trials? In other words, to what extent do cognitive shortcomings that are manifest in the laboratory reflect cognitive shortcomings that will not, in fact, be compensated for by other aspects of the trial process—aspects such as group deliberation, cross-examination, and the like? These are serious questions that threaten to cut short the social epistemology inquiry before it gets started.

Fortunately, there are some answers. The best way to appreciate them is by looking at a paradigmatic case of a social epistemology inquiry into one particular area of the law of evidence.

A. The Case of Demeanor

In a well-known 1991 article, Professor Olin Guy Wellborn III examined the empirical evidence concerning (as we would put it) the instrumental value of demeanor evidence for maximizing veritistic value. As Wellborn comments:

The importance of demeanor as an indicator of credibility is commonly cited as a premise of the general requirement of live testimony, the hearsay rule, and the right of confrontation. The importance placed upon demeanor information is highlighted by the strict limits traditionally placed upon trial use of depositions and transcribed testimony taken in other proceedings. The opportunity of the trier to observe the demeanor of witnesses is a principal basis for the deference accorded by reviewing courts to factual determinations of trial courts and hearing officers.

The assumption that demeanor provides highly useful information plays an important role in other procedural doctrines.\footnote{Id. at 1077 (footnotes omitted).}

In fact, however, it appears that demeanor evidence has little instrumental value as a maximizer of veritistic value. "[T]he experimental evidence indicates that ordinary observers do not benefit from the opportunity to observe nonverbal behavior in judging whether someone is lying."\footnote{Id. at 1088; see id. at 1078 n.10, 1079 n.12 (citing the evidence discussed). See generally Paul Ekman, Why Don't We Catch Liars? 63 Soc. Res. 801 (1996) (discussing possible explanations for the phenomenon). For other recent work on this topic, see, e.g., Thomas H. Feeley & Mark A. deTurck, Global Cue Usage in Behavioral Lie Detection, 43 Comm. Q. 420 (1995). But see Mark A. deTurck & Gerald R. Miller, Training Observers to Detect Deception: Effects of Self-Monitoring and Rehearsal, 16 Hum. Comm. Res. 603 (1990) (finding that training can help to improve detection of deception, at least deception by unskilful liars); Mark G. Frank & Paul Ekman, The Ability to Detect Deceit Generalizes Across Different Types of High-Stake Lies, 72 J. Personality & Soc. Psychol. 1429 (1997) (finding that lie detection improves when the stakes are high, probably because the liar gives off more nonverbal cues).}

In addition, "a good body of experimental evidence consistently shows that jurors simply cannot tell whether a witness's perception and memory are accurate... [N]either verbal nor nonverbal cues are effectively employed..."\footnote{Wellborn, supra note 162, at 1091-92.} Wellborn, however, is appropriately cautious in his normative recommendations in light of the empirical facts about demeanor evidence. The conclusion he draws is neither that we should eliminate live testimony—such testimony may still maximize veritistic value in factfinding, even if it does not enhance credibility determinations, and in any case, live testimony serves non-veritistic purposes as well\footnote{Wellborn, supra note 162, at 1090-91; see, e.g., id. at 1089 nn.70, 73 & 74 (citing material that discusses that body of evidence). See generally Elizabeth F. Loftus, Psychologists in the Eyewitness World, 48 Am. Psychologist 550 (1993) (discussing psychological studies of eyewitness identification and eyewitness testimony); Siegfried Ludwig Sporer et al., Choosing, Confidence, and Accuracy: A Meta-Analysis of the Confidence-Accuracy Relation in Eyewitness Identification Studies, 118 Psychol. Bull. 315 (1995) (analyzing thirty studies to test the link between confidence and accuracy in eyewitness identification). For recent work in a similar vein, see C.A. Elizabeth Luus & Gary L. Wells, The Malleability of Eyewitness Confidence: Co-Witness and Perseverance Effects, 79 J. Applied Psychol. 714 (1994); Gary L. Wells & Amy L. Bradfield, "Good, You Identified the Suspect": Feedback to Eyewitnesses Distorts Their Reports of the Witnessing Experience, 83 J. Applied Psychol. 360 (1998).}—nor that we should eliminate the hearsay rule...
(since the more important "hearsay dangers" are lack of cross-examination and absence of the oath). The facts about demeanor evidence do suggest several modest reforms, however. For example, appellate courts should give less deference to the factual findings of trial courts, especially as they bear on witness credibility, since "a transcript is actually as good a basis for a credibility determination as live testimony."

Wellborn is cautious in another way that makes his article a good model for the social epistemology of evidence. He is sensitive to the dangers mentioned above about generalizing from laboratory results to the rules of evidence as they apply at trial. In particular, he identifies four pertinent differences that might undermine the utility of experimental results—what he calls "context, cross-examination, deliberation, and preparation."

- **Context:** "In a trial, each witness's testimony has a much more substantial context—the other evidence in the case—than the respondents' stories in the experiments."

- **Cross-examination:** Experiments typically do not elicit information from respondents via the trial method of direct and cross-examination.

- **Deliberation:** "Jurors deliberate and make decisions by consensus rather than individually; by contrast, the experimental subjects decide alone whether to believe a respondent's statement."

- **Preparation:** Real witnesses are often rehearsed or coached prior to testifying, in part to enhance their credibility. "Experimental respondents normally make their true or false statements without rehearsal or coaching."

The question then is whether these differences vitiate the value of existing empirical evidence. In the case of experimental evidence on the probative value of demeanor, Wellborn concludes

---

167 Id. at 1094.
168 Id. at 1095 (making this point but eventually rejecting, on policy grounds, the suggestion that appellate courts should give less deference to the factual findings of trial courts).
169 Id. at 1079.
170 Id.
171 Id.
172 Id.
173 Id.
that they do not. "The four named trial conditions probably decrease, rather than increase, the utility of nonverbal deception cues." So, for example, Wellborn points out that other empirical evidence shows that suspicious questioning (like that in a cross-examination) and stress (a by-product of an aggressive cross-examination) tend to make even mock jurors "view their responses as deceptive even when they are honest, which significantly increases [lie] detection errors." Additionally, some studies have employed real attorneys, as well as direct and cross-examination, and yet still found that "[e]ven experienced lawyers, free to question the witness as they chose, were unable to lead mock jurors to believe accurate eyewitnesses more than inaccurate eyewitnesses."

In the end, the details of the demeanor case matter less than the guidelines Wellborn's approach suggests. With respect to proffered empirical evidence bearing on the instrumental value of some evidentiary rule for maximizing veritistic value, we must always ask whether differences in context, cross-examination, deliberation, and preparation between the experimental and trial settings affect the utility of the empirical data for evaluating the evidentiary rules within a social epistemology framework.

We now turn to sketch two areas of evidence law that, when analyzed from the standpoint of naturalized epistemology, cry out for reform and/or additional research. These examples are merely illustrative and in many respects are familiar to scholars of the rules in question. They suggest, however, the structure of the questions and problems with which evidence, as a branch of social epistemology, should concern itself.

**B. Probabilistic Evidence**

Probabilistic evidence is increasingly important in trials, yet there remains a serious question about what exactly jurors make of it. Does such evidence "mislead" and "confuse" jurors within the meaning of Federal Rule 403, and if so, when and how?

---

174 Id.
175 Id. at 1080; see id. at 1080 n.19 (citing that evidence).
176 Id. at 1090 (quoting R.C.L. Lindsay et al., Mock-Juror Belief of Accurate and Inaccurate Eyewitnesses, 13 Law & Hum. Behav. 333, 338 (1989)).
The best-known instances of probabilistic evidence concern DNA matching of a defendant's blood sample to blood at the scene of a crime, as in the O.J. Simpson case. Probabilistic evidence sometimes yields astounding probabilities such as one in fifty-seven billion. What does such evidence really mean? It gives us the "random match probability" ("RMP"), that is, "the theoretical likelihood that a randomly selected person from the general population (or from the population of certain large ethnic or racial groups) would genetically match the trace evidence as well as the defendant." The worry about such evidence, however, is that lay persons may understand it to mean things it does not mean. Such evidence, for example, does not give us a "source probability," that is, "the likelihood that the defendant is the source of the trace." That probability is affected by all the other evidence for and against the claim that the defendant is the source.

Of course, if the odds are, for example, fifty-seven billion to one, and Simpson matches, then it is unclear who else on the planet could also be a possible source. Professor Jonathan J. Koehler has urged in a number of papers that such enormous probabilities are still misleading because of the fact that error rates in labs (on the order of one in a hundred results) undermine the validity of the astronomical probability. As Koehler comments:

Do the tiny DNA frequencies—frequencies on the order of 1 in millions, billions, and trillions—have any probative value beyond that which is given by the laboratory error rate when the error rate is many orders of magnitude greater than DNA frequency? My answer is that they do not....
Even granting Koehler this much, it does not follow that such evidence should be excluded under Federal Rule 403: The obvious solution seems to be for triers of fact to also consider error rates. Some empirical evidence, however, suggests that this will not suffice. Koehler and colleagues, drawing on recent empirical psychology, observe that:

[J]urors may overweight extremely small RMPs for two reasons. First, vividness theory suggests that people give inferential weight to evidence in proportion to its vividness or memorability. Very small statistics, such as one in millions or billions, may be quite vivid and memorable, and therefore exert a large impact on verdicts. Second, people often combine probabilistic items of evidence by averaging them. When an averaging strategy is used to estimate the disjunctive probability of error from either of two sources, one of which is several orders of magnitude smaller than the other, it substantially overweights the contribution of the smaller error source. In this way, jurors provided with RMPs and laboratory error rates may attach great significance to very small—but minimally diagnostic—RMPs.

Koehler and his colleagues tested these hypotheses with jury-eligible subjects at the University of Texas and in the local community. Subjects reviewed written summaries of a murder case, in which the strongest evidence was the RMP connected to the DNA evidence. Some subjects received no RMP, some received the RMP without information on laboratory error rates, and some received both pieces of information. The researchers found “that introduction of the RMP had a strong effect on mock jurors’ verdicts, both when laboratory error rates were absent and present. . . . Introduction of laboratory error rates . . . did not significantly affect conviction rates . . . .” Their conclusion, accordingly, was that

---

183 See Koehler et al., supra note 181, at 212 nn.42–43 (citing sources).
184 Id. at 212 (footnotes omitted).
185 For the details of the methodology in the two studies, see id. at 212–15.
186 Id. at 213. They add:

The probability that a juror would convict in the hypothetical case doubled or tripled when the RMP was introduced. Separate introduction of the highly diagnostic laboratory error rate had little impact. But when the aggregated error rate [which “reflected the combined possibility of error from coincidental matches and laboratory mistakes,” id.] was introduced, and jurors were not
RMP evidence was "potentially prejudicial" and that even when jurors are also given information about error rates, they are still "likely to be hopelessly confused about the weight to accord" the evidence.

Federal Rule 403 calls for a balancing test, and nothing said so far establishes how the balance should be struck: RMPs may be prejudicial, but they are also rather clearly probative as well. But there is a more serious worry arising from Wellborn's cautionary notes about applying laboratory results to real-world trials. As Koehler and his colleagues concede:

An important issue for future research is whether the findings described here will persist even when corrective steps are taken. For example, expert testimony or judicial instructions explaining that laboratory error rates should be considered even in cases involving very small RMPs may be sufficient to sensitize jurors to the normative issue. We also caution that the studies presented here do not consider the effects of group deliberation.

In addition, the Koehler experiments did not include live testimony and, in particular, did not include live cross-examination. Surely a skilled attorney (such as one who has carefully read Professor Koehler's articles) might effectively demolish RMP evidence in the eyes of a jury.

In the case of probabilistic evidence, then, social epistemology can make no concrete recommendations yet. We know that in the laboratory, lay people make a mess of RMPs and error rates. To date, we have no real idea what jurors in real trials will make of such evidence.

afforded separate access to a small RMP, convictions rates [sic] declined by nearly half.

Id. at 213-14.

106 Id. at 215.


109 Koehler et al., supra note 181, at 216–17.
C. Character Evidence

Character evidence figures in the Federal Rules in two contexts: Federal Rule 404(a) excludes most character evidence in criminal trials (except in special circumstances) and all character evidence in civil trials, but Federal Rule 404(b) has the practical consequence of making it possible to admit most character evidence by making admissible evidence "of other crimes, wrongs, or acts" to prove "motive, opportunity, intent, preparation, plan, knowledge, identity, or absence of mistake or accident."190 Federal Rule 608, in turn, permits evidence of character in order to impeach a witness.191

There is now a large social psychology literature examining the explanatory power of the concept of character: Do individuals act "in conformity with character"—with stable long-term behavioral dispositions—or do they, rather, act in situation-specific ways such that the notion of "character" is an unreliable predictor of subsequent conduct? A now-dominant view—"situationalism"—holds that people's actions are situation-specific, rather than reflecting stable dispositions constitutive of character.192 Situationalism runs strongly counter to common sense about explanations of behavior. As two situationalists explain:

"People tend to (a) offer dispositional explanations for behavior instead of situational ones, even when it should be transparent that the behavior is produced by situational factors . . .; (b) make overly confident predictions about behavior on the basis of a small amount of information concerning dispo-

190 Fed. R. Evid. 404.
191 Fed. R. Evid. 608.
sitions; and (3) [sic] describe the self as well as others in terms of internal dispositions rather than context-specific factors.  

One natural question for social epistemology to ask is: if situationalism is correct, what becomes of character evidence? Professor Richard Friedman has addressed the issue in the context of impeachment, arguing that even non-situationalist but plausible psychological premises require a radical revision of the impeachment rules. Situationalism strengthens the case. If "manipulations of the immediate social situation can overwhelm in importance the type of individual differences in personal traits or dispositions that people normally think of as being determinative of social behavior," then why think bad behavior in some out-of-court context has bearing on truth-telling in court, under oath, in front of a jury, with a threat of perjury? The kind of situation in which character impeachment evidence is generated and the kind of situation in which the witness testifies could not be more different.

In any case, Friedman has already addressed impeachment at some length. We want to consider here Federal Rule 404 in light of situationalism. The whole premise of character evidence—namely, "to prove the character of a person in order to show action in conformity therewith"—is exactly the premise of lay psychology that situationalism repudiates. As two leading situationalists, Professors Lee Ross and Richard E. Nisbett, write: "[P]eople are inveterate dispositionists. They account for past actions and outcomes, and make predictions about future actions and outcomes, in terms of the person—or more specifically, in terms of presumed personality traits or other distinctive and enduring personal dispositions." In fact, however, "standard correlation coefficients determined in

---

195 Ross & Nisbett, supra note 192, at xiv.
196 Fed. R. Evid. 404(b).
197 Ross & Nisbett, supra note 192, at 90.
well-controlled research settings” show that “personality traits” lack substantial “explanatory and predictive power.” 198

Federal Rule 404(b) does not permit the use of character evidence for “inveterate dispositionist” purposes. Other “crimes, wrongs, or acts” may be highly probative of “motive, opportunity, intent [etc.]” for reasons perfectly congenial to the situationalist: Insofar as the other acts are similar along one of the relevant dimensions, their probative value is clear. The real worry, rather, is that Federal Rule 404(b) in effect undermines the bar on more generalized character inferences of the sort situationalism tells us are not warranted. A jury may be warned that the evidence of prior wrongdoing is only to be considered with respect to defendant’s “opportunity” to commit the crime with which he is currently charged, yet the suspicion is strong that jurors will also draw the inference that the defendant has “bad character” and draw damaging inferences accordingly.

This means, of course, that there is always an argument for exclusion of Federal Rule 404(b) evidence on Federal Rule 403 grounds. If situationalism is correct, the Federal Rule 403 dangers are quite substantial: If the jury draws (forbidden) inferences from putative traits of character, the jury will be seriously misled and prejudiced, since situationalism teaches us that character traits have relatively little predictive power. The more radical conclusion that situationalism suggests, however, is that the Federal Rule 404(b) exception that swallows the rule is a bad idea: Situationalism would counsel that the Federal Rule 404(a) bar on character evidence be preserved.

We do not, however, want to overstate the conclusion. For one thing, this argument against Federal Rule 404(b) and the admission of character evidence is premised on the truth of situationalism. Even a casual reading of the social psychology literature suggests certain weaknesses in the evidence for situationalism. Most obviously, the predictive value of situational variants is limited. Thus, the famous 1973 study of Good Samaritan behavior 199 found that “[i]f the subjects were in a hurry . . . , only about 10 percent helped [the person needing assistance]. By contrast, if they were not in a

198 Id. at 91.
199 Darley & Batson, supra note 192.
hurry..., about 63 percent of them helped. What about that ten percent? Would it not be reasonable to invoke their good character relative to the majority, in explaining their behavior?

Moreover, there are difficulties with the notion of a "situation." Ross and Nisbett make the following claim:

While knowledge about John is of surprisingly little value in predicting whether he will help the person slumped in the doorway, details concerning the specifics of the situation would be invaluable. For example, what was the appearance of the person in the doorway? Was he clearly ill, or might he have been drunk or, even worse, a nodding dope addict? Did his clothing make him look respectably middle class or decently working class, or did he look like a homeless derelict?

Supposing that these factors are relevant, how exactly do they show that knowledge of character is irrelevant? Doesn't it make perfectly good sense to say that someone of a genuinely charitable (or altruistic or sensitive) character (perhaps that ten percent we met above) thinks only of human need, and not of the class status of the person in need? Would it not be quite natural to say that the people who let class status determine their decision to help those in need betray something about their character?

Even if situationalism is correct, there still remains the question of what real jurors—who hear character evidence in context, subject to cross-examination, under instruction from the judge to consider it only with respect to Federal Rule 404(b) issues, and who then deliberate about its significance with others—actually do with such evidence. It is possible that the faulty inference that situationalism would reject is not one jurors will make, even if experimental subjects and ordinary people do draw those inferences.

CONCLUSION

In this paper, we have tried to sketch and defend a theoretical framework for evidence scholarship and naturalized epistemology. We have also shown both the kinds of theoretical approaches it rules out (Part III) and the kinds of questions and inquiries it de-

---

200 Ross & Nisbett, supra note 192, at 4.
201 Id. at 3.
mands (Part IV). As noted earlier, there is already a body of evidence scholarship operating within what we would call the naturalized epistemology framework, examining topics ranging from hearsay\footnote{E.g., Margaret Bull Kovera et al., Jurors' Perceptions of Eyewitness and Hearsay Evidence, 76 Minn. L. Rev. 703 (1992); Peter Miene et al., Juror Decision Making and the Evaluation of Hearsay Evidence, 76 Minn. L. Rev. 683 (1992).} to expert evidence.\footnote{E.g., Ronald J. Allen, Expertise and the \textit{Daubert} Decision, 84 J. Crim. L. & Criminology 1157 (1994); Ronald J. Allen & Joseph S. Miller, The Common Law Theory of Experts: Deference or Education?, 87 Nw. U. L. Rev. 1131 (1993).} We hope we have shown why there are good philosophical reasons supporting the practical reasons for this kind of evidence scholarship to predominate in the exploration of the law of evidence. A priori formalisms that too often have commanded the allegiance of law professors may have their place, but that place is limited to the suggestion of avenues for research. It does not extend to generating useful conclusions about the real world.