1981

Paradoxes, Gedanken Experiments and the Burden of Proof: A Response to Dr. Cohen's Reply

David H. Kaye

*Penn State Law*

Follow this and additional works at: [http://elibrary.law.psu.edu/fac_works](http://elibrary.law.psu.edu/fac_works)

Part of the [Civil Procedure Commons](http://elibrary.law.psu.edu/fac_works), and the [Evidence Commons](http://elibrary.law.psu.edu/fac_works)

**Recommended Citation**

Paradoxes, Gedanken Experiments and The Burden of Proof: A Response to Dr. Cohen’s Reply

David Kaye*

In a series of papers1 I outlined a solution to what I called the problem of naked statistical evidence. This problem, a hardy perennial in the law of evidence, surfaced in Sargent v. Massachusetts Accident Company,2 bloomed in Smith v. Rapid Transit Company,3 and induced a profusion of secondary growth in the past two decades.4 The problem was restated most recently by L. Jonathan Cohen in the guise of his “paradox of the gate-crasher.”5 According to Cohen, this paradox is one of several that

---

*visiting fellow, Faculty of Law, the University of Southampton; S.B. 1968, Massachusetts Institute of Technology; A.M. 1969, Harvard University; J.D. 1972, Yale University.


5. As I pointed out in each of my previous articles, the gatecrasher paradox is not the only weapon in Dr. Cohen’s armamentarium. In Subjective Probability and the Paradox of the Gatecrasher, 1981 Ariz. St. L.J. 627, he observes that I have not analyzed, point by point, his other arguments against the mathematical theory of probability. Nor shall I. Others have undertaken this task. See the authorities cited in Kaye, The Laws of Probability, supra note 1, at 38 n.18; Schum, A Bayesian Account of Transitivity and Other Order-Related Effects in Chains of Inferential Reasoning, Rice Univ. Dep’t of Psych. Research Rep. No. 79-04, Dec. 30, 1979; Schum, On Factors Which Influence the Redundancy of Evidence and Coroborative Testimonial Evidence, Rice Univ. Dep’t of
reveal that the kind of "probability" at work in litigation does not conform to the axioms of mathematical probability.

In contrast, the analysis I delineated suggests that the familiar theory of probability needs no revision to account for the reluctance of a few courts to permit plaintiffs to prevail on the strength of background statistics alone. I am flattered that Dr. Cohen has found my treatment sufficiently penetrating to warrant a response. Still, I remain satisfied that one need not adopt the esoteric mathematical structure he proposes to explain the burden of proof in civil cases. In Part I, I show that whether or not one accepts the subjective interpretation of probability, nothing in Cohen's most recent paper establishes that forensic probabilities are incommensurable with the usual mathematical axioms. In Part II, I consider, in its own right, Cohen's claim that the subjective interpretation is a "dangerously inappropriate paradigm for the courts."

I. THE SUITABILITY OF THE MATHEMATICAL AXIOMS

In earlier articles I described two distinct lines of argument reconciling probability theory with the disfavored status of naked statistical evidence. The first was that even though resolving a factual issue according to whether the relevant probability exceeds one-half maximizes the expected number of correct decisions, the very policy of accuracy in decisionmaking also justifies finding against the proponent where the probability is calculated from the background statistics alone. As I stated:

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

---

6. See Cohen, supra note 5 at 632.
7. To date, I have dealt only with the easy case in which the plaintiff can provide no satisfactory explanation for his exclusive reliance on background statistics. I hope to analyze the more difficult issue of "justified naked statistical evidence" in the near future.
8. Kaye, Naked Statistical Evidence, supra note 1, at 605 n.19.
9. This point is a restatement, or perhaps an elaboration, of the suggestion made in Tribe, supra note 4, at 1349.
background statistics alone—plaintiffs should be forced to produce it. In the long run, fewer mistaken verdicts should result under this rule of law.10

Cohen does not address this logic in his rejoinder. Yet, the argument suggests that ordinary probabilistic reasoning is compatible with judicial disapproval of naked statistical evidence. Consequently, it is hard to see how the problem of naked statistical evidence—whether posed in the form of the gatecrasher paradox, the blue bus hypothetical, or the few decisions on point—demands any departure from the axioms of probability theory.

My second way of justifying the legal rule without inventing new axioms involved distinguishing between “subjective” and various “objective” interpretations of probability. I employed this distinction to dispute the implicit assumption that because a measured proportion $p$ of a class of objects has some property (501/1000 persons gatecrashed the rodeo, 4/5 blue busses belong to defendant, etc.), the object in question in a particular court case has the probability $p$ of possessing the same property. I challenged this assumption because it ignores a crucial feature of the situation in these cases. Presumably, the plaintiff could have produced more probative evidence with little effort. The failure to adduce such evidence suggests that the evidence would not have supported plaintiff's claim. Hence, the probability calculated as the relative frequency $p$ seems overstated. One way to incorporate this inference from non-production into a formal analysis that uses the standard axioms is to apply an elementary formula known as Bayes' Theorem.11

It is this second line of argument, and especially my reference to the subjective interpretation of probability, that has attracted Cohen's attention. He insists that the subjective view is not suitable to modeling forensic decision making. Ironically, I have come to think that the debate over the subjective conception of probability is something of a red herring here. Contrary to what I may have suggested in my earliest writing,12 the subjective interpretation is not essential to my defense of the usual probability axioms. After all, Bayes' formula, and every other result in

10. Kaye, The Laws of Probability, supra note 1, at 40 (footnotes omitted); cf. The Paradox of the Gatecrasher, supra note 1, at 106 ("it may be appropriate to treat the subjective probability as less than one-half, and therefore insufficient to support a verdict for plaintiff, simply to create an 'incentive for plaintiffs to do more than establish the background statistics.' ").
12. Compare Kaye, The Paradox of the Gatecrasher, supra note 1, at 106 ("This distinction between objective and subjective interpretations . . . is crucial") with Naked Statistical Evidence, supra note 1, at 609 ("My argument is most easily developed with the aid of the subjective interpretation").
probability theory, works as well with "objectively" ascertained probabilities as with subjectively estimated ones, and the probabilities that jurors must estimate, however incoherently, can be construed as relative frequencies, degrees of confirmation, objective chances, or any of the other interpretations proffered by philosophers of mathematics. If I am right in this regard, it follows that the truth of Cohen's thesis—that the gatecrasher example establishes that the probability axioms break down as applied to forensic proof—in no way turns on whether one adopts a subjective interpretation of probability.

Of course, Cohen is not fond of the prevalent relative frequency view of probability either. In this latest essay he unveils a new hypothetical case intended to show that what counts in forensic proof is "the build-up of a suitable weight of evidence" rather than "a mere frequency (which could be just accidental)." The example involves a man killed in an automobile accident. The factual issue is said to be whether, but for the negligence of the defendant, the man would have survived to age 70. Cohen points out that it would not do to deduce the probability of this event by looking to the number of males who were once the age of the deceased but who subsequently lived to be 70 and then comparing this number to the total number of males who lived at least as long as the deceased. This relative frequency calculation would be too crude. Additional characteristics of the deceased permit a more refined calculation. It seems that the deceased was a coal miner, suffered from incipient silicosis, had decided to work above ground, and so on. At this stage, Cohen is simply partitioning the sample space so as to compute the relevant probability on the basis of all the available information. Equivalently, one could use Bayes' formula, with probabilities ascertained from relative frequencies, to incorporate the effect of each additional datum on the initial, crudely ascertained probability of survival. There is no need to introduce an ordinal grading according to "the weight of the evidence." The various items of evidence

\[ \frac{P_1}{1-P_1} = \frac{P_o}{1-P_o} = \text{LE}_{E_1} \]

where \( P_1 \) is the probability of \( H \) considering \( E_1 \) as well as the prior data, and \( \text{LE}_{E_1} \) is the likelihood ratio for \( E_1 \). This ratio is defined in terms of the conditional probabilities of \( E_1 \).
are merely used to arrive at the appropriate relative frequency.

Nonetheless, Cohen goes on to say that calculating relative frequencies fails because "accidental" rather than "causal" factors may be involved. In the context of the coal miner case he writes: "It could well be the case that the deceased's surname contained six letters and that the frequency of those with six-letter-names surviving to age 70 is greater than the frequency in the population at large. But that fact would be quite irrelevant to the matter at issue because it is only an accident and not the manifestation of some causal link between name-length and longevity."

On this basis, Cohen concludes that "the type of Pascalian probability that is characteristically in the courts is [not] a mere frequency," for a mere frequency "could just be accidental."

I must confess that I find this portion of Dr. Cohen's paper hard to follow, but there seem to be at least two serious flaws in the discussion. For one, the example itself is deceptive. Surely the real reason that name-length is legally irrelevant to the issue of expected lifespan is that we know very well (or think we do) that there is no statistically significant association between these two variables. Were we actually confronted with a finite population in which substantially more men with six-letter names became septuagenarians, then the evidence would be relevant. Such evidence would, in the language of the Federal Rules of Evidence, tend "to make the existence of [a] fact that is of consequence to the determination of the action more probable or less probable than it would be.

\[
LE_i = \frac{P(E_i|H)}{P(E_i|Hc)}
\]

For example, if \(E_i\) denotes the fact that the deceased was a coal miner with incipient silicosis, the likelihood ratio could be obtained from survival table showing (a) the proportion of males (of the age of the deceased) who survive to age 70 who are coal miners with incipient silicosis, and (b) the proportion of males who reach age 70 who are not coal miners with incipient silicosis. If it were found that these proportions stood in the ratio of two to seven, \(LE_i\) would be taken to be \(\frac{2}{7}\) under a relative frequency approach. For (1) the revised probability \(P_o\) would then be .4, as Dr. Cohen's postulates. It is easy to show that for \(n\) items of evidence \(E_1, E_2, \ldots, E_n\) about the deceased, the probability of survival, considering these successive refinements, is related to the crude figure for all males of the age of the deceased by the equation

\[
P_n = \frac{P(E_1|H)E_2 \ldots E_n|Hc)}{P_o}.
\]

where \(LE_n = \frac{P(E_n|HE_1E_2 \ldots E_{n-1})}{P(E_n|HcE_1E_2 \ldots E_{n-1})}\)

15. See Cohen, supra note 5 at 633.

16. Id.
without the evidence.\textsuperscript{17} This is to say, contrary to what Cohen implies, that there is no legal requirement of causal as opposed to purely statistical association. For example, evidence as to gender is relevant in wrongful death cases even though, to my knowledge, the courts do not entertain the belief that possessing a pair of X chromosomes \textit{causes} longer life. It suffices that there is an unexplained but well documented statistical dependence between gender and lifespan in contemporary society.\textsuperscript{18}

Second, even if the claim that a "causal" rather than a statistical relation is a prerequisite to legal relevance were well taken, this would not show that to decide a disputed factual question one must look to an ordinal grading according to the weight of the evidence instead of the weight of the evidence expressed as a cardinal probability number. No doubt, Cohen is correct when he insists that a \textit{mere} frequency does not always supply a suitable figure for the pertinent probability. We have already seen that the class of objects must be defined with some care before a final relative frequency is measured. If, as part of this effort, we also need a "causal-propensity criterion" to avoid computing the wrong relative frequency, then so be it.\textsuperscript{19} It is still probabilities conforming to the mathematical axioms that determine, subject to certain constraints of policy, how contested factual issues should be resolved.

Upon inspection, then, Cohen’s latest arguments do not add much to his case against mathematical probability. The problem of naked statistical evidence can be handled without substituting inductive probabilities for the familiar mathematical ones. Moreover, this is so whether one in-

\footnotesize
\textsuperscript{17} FED. R. EVID. 40. For a careful analysis of legal relevance using mathematical probability theory and confirming this conclusion, see Lempert, \textit{supra} note 14.
\textsuperscript{18} Since Cohen does not explicitly define what he means by "accidental" and "causal," I may be using these terms somewhat differently than he. It might be said, for instance, that gender \textit{is} a causal factor in lifespan inasmuch as the "appropriate counterfactual inference," Cohen, \textit{supra} note 5, at 634, involves an alteration in probable lifespan. Under this view, however, the only "accidental" links would seem to be those that have been observed more often than not but which we somehow know would not be detected in such number in a much larger sample of observations. But this merely says that the apparent association between the two variables is not statistically significant. Accordingly, the evidence is of little probative value and should be excluded—though not because of any defect in probability theory or the relative frequency interpretation.
\textsuperscript{19} As developed by Pierce, Popper, and others, propensity theories of probability are proposed as distinct alternatives to relative frequency interpretations. \textit{E.g.,} J. MACKIE, \textit{TRUTH, PROBABILITY AND PARADOX: STUDIES IN PHILOSOPHICAL LOGIC} 155-57, 179-86 (1973). As described by Cohen in the coalminer example, however, the propensity "criterion" seems to supplement a simple relative frequency theory. For present purposes, nothing much turns on this distinction. Construing probability as the manifestation of innate "dispositions" or "propensities" does not entail denying that probabilities behave in accordance with the standard mathematical axioms. For some views on the merits of the propensity interpretations of mathematical probabilities, \textit{compare} D. MELLOR, \textit{THE MATTER OF CHANCE} (1971) (defending propensity theory) with J. Mackie, \textit{supra}, at 187 ("there are strong reasons for doubting whether . . . propensity theory has ontologically valid applications.").
interprets probabilities as subjective estimates, relative frequencies, causal
propensities, or the like. Persistent riddles as to how a sample space
should be partitioned and how the probabilities of elementary events
should be obtained ought not be confused with the very different issue of
whether these numbers — however they may be arrived at — obey the
probability axioms. With respect to the latter question, it seems clear
that the mathematical theory of probability has shown itself to be a useful
tool in describing and criticizing many features of the law of evidence.

At the same time, I believe that the applicability of the subjective inter-
pretation of probability to juridical proof — the principal subject of Dr.
Cohen's paper — is an important and interesting question in its own
right. It is to this topic that I now turn.

II. THE SUITABILITY OF THE SUBJECTIVE INTERPRETATION

Most interpretations of probability are strained when applied to the
unique, non-repeatable events that are ordinarily the subjects of litigation.
The subjective view, however, seems congenial to modeling legal factfind-
ing. For heuristic purposes at least, jurors can be visualized as making
personal estimates of the probability that the events in question took place
in the manner described by the plaintiff or defendant and deciding the
case in accordance with these estimates. The results of such a model can
then be employed in normative or descriptive ways.

In *The Provable and the Probable*, Dr. Cohen denied the aptness of
this subjective interpretation on two grounds. I took issue with him on
both. In rebuttal, he now defends one of his original arguments, omits
any mention of the other, and introduces yet a third. His remaining
opening argument is that the statements a juror might make about bet-

---

20. The need to draw such a distinction always arises when it comes to making use of axiomatized
systems. In applying Euclidean geometry, for instance, the suitability of the postulates (particularly
the parallel postulate) is one issue; the procedure for measuring the entities presumed to be subject to
these axioms (lines, angles, etc.) is another matter. But see H. REICHENBACK, THE PHILOSOPHY OF
SPACE AND TIME (1953).


22. Roughly, these were that the betting odds conceptualization of subjective probability cannot
apply because there is no independent procedure for resolving hypothetical bets about facts contro-
verted by litigants, and because the odds a juror would accept depend in part on the magnitude of the
bet.


24. Although I did not realize it at the time, the principal argument I advanced in the Laws Of
Probability to refute Cohen's claim that subjective probabilities are poorly defined because coherent
betting quotients depend on the size of the wager was already set out quite clearly and forcefully in
the philosophical literature. See D. Mellor, *supra* note 19, at 34-37. Cohen's continued refusal to
recognize this argument, let alone to come to grips with it, is disappointing.
ting odds or lottery tickets would carry no conviction and have no mean-
ing without some realistic way of settling the bet or the lottery. He goes
so far as to say that bets or lotteries about litigated facts could not be
settled without deterministic metaphysics, perfect social sciences, and in-
delible traces of the past.25

Although one senses a bit of hyperbole here, I have no doubt that were
jurors really to gamble on the truth of their findings, a great many of
their bets could not be settled even with the expenditure of inordinate
sums of money and the most invasive forms of interrogation. As Cohen
notes, only H.G. Wells’ time machine would permit every controversy to
be resolved with certainty. But this observation is not a telling rejoinder to
the suggestion that the probabilities implicitly at work in legal factfinding
can be given a conceptually meaningful subject interpretation. Betting
odds, lottery tickets (and bids for reference contracts) are merely heuris-
tic devices for describing, in ways that are easily visualized, preferences
and choices under conditions of risk. Decision theory reveals that as long
as a person acts in accordance with his preferences among alternative
risky outcomes and that as long as these preferences have certain plausi-
ble properties, that person can be said to be predicking his choices on
subjective, but mathematically well-behaved probabilities. Cohen’s posi-
tion to the effect that “nothing whatever [is] at risk” is difficult to recon-
cile with the obvious fact that in the uncertain world of litigation, jurors
always risk making erroneous decisions. By definition, conscientious jurors
care about this. They would prefer to avoid such outcomes. Their deci-
sions are thereby “constrained by considerations of risk” — the omnipres-
tent risk or error, even if in most cases such errors are never discovered.26

If need be, the same point can be made in terms of wagers and the like.
The purpose of depicting hypothetical wagers, lotteries, or reference con-
tracts is not to persuade jurors to make book on disputed issues of fact.
Rather, it is to exhibit a reasonably concrete but entirely theoretical pro-
cedure for generating, in principle, a set of numbers conforming to the
probability axioms. Of course, this procedure cannot be implemented in
the legal context, or for that matter, in any other. A living specimen of
the “rational” decisionmaker has yet to be exhibited. But this limitation
does not indicate any defect in the proof that subjectively produced esti-

25. See Cohen, supra note 5 at 632.
26. It may be worth noting that factually erroneous verdicts do not invariably escape detection.
Subsequent information tending to confirm or contradict a verdict does sometimes come to light, a
consideration that may further prompt conscientious jurors to make as accurate an assessment of the
probabilities in question as can be had on the basis of the evidence at bar. A recent, scandalous
illustration of such belatedly discovered information is provided in The Times (London), June 20,
1981, at 1, c. 4.
mates can have the same mathematical properties as objectively measured ones. In fact, the impracticability of the procedure is typical of all gedanken experiments. One can deduce the famous equation $E = mc^2$ by thinking about what would happen in a rigid box with a burst of photons bouncing off an inside surface. Yet, no physicist would seriously consider obtaining an experimental verification of the mass-energy relationship in this way.

So, to undermine the subjective interpretation, it must be demonstrated that even the completely hypothetical, inexorably rational creature of decision theory would be at a loss to state the odds, lottery tickets, or bids that he would find acceptable if a wager, lottery or sale actually could be held. The structure of this individual’s preferences for the relevant (admittedly imaginary) outcomes (that he will be proved correct in having agreed that event $X$ is as plaintiff contends, that he will be given a certain sum of money with some probability $p$, and so on) must be shown to contradict the decision theoretic axioms of rational choice. Since, by hypothesis, the decisionmaker is rational and willing to think about such outcomes, it is not surprising that no such showing has been made.

This leaves only Dr. Cohen’s newly stated objection to the subjectivist thesis. In essence, he complains that subjective probabilities are, well, subjective. They can vary from person to person. Initially, one might have thought that this characteristic would have commended the subjective interpretation to us. Jurors hearing the same evidence do sometimes reach different conclusions. Yet, Cohen finds the personal quality of the subjective interpretation unpalatable because it puts “the probability with which a point has been proved... wholly outside the framework of rational controversy.”

Again, the complaint seems misguided. Without question, there is no logical inconsistency in A stating that “for me, the probability is .7,” and

28. Seen in this light, Cohen’s quarrel with the subjectivist thesis rests on a claim about human psychology—namely, that the human mind cannot or will not produce by subjective means a set of numbers that conform to the probability axioms unless it is placed in an actual, rather than a hypothetical condition of risk. This claim misses the point. No one expects any human being to achieve perfect coherence even in real-world situations. The rigorous subjective formulation of probability works for hypothetical decision makers only. Why should placing these unreal entities in hypothetical situations preclude them from achieving consistency in their efforts to ascertain probabilities?

Naturally, whether actual jurors come close to attaining this coherence in their unarticulated probability estimates is an empirical question. See Schum & Martin, Empirical Studies of Cascaded Inference in Jurisprudence: Methodological Considerations, Rice Univ. Dep’t of Psych. Research Report No. 80-01, May 30, 1980. I merely contend that for theoretical purposes, subjectively ascertained probabilities can be assigned to events of interest in litigation.

29. See Cohen, supra note 5 at 630.
in B asserting that "for me, the probability is only .3." If A and B are not lying, the two propositions have the same truth value. But this truth-functional equivalence does not preclude an advocate from insisting that the better subjective estimate is .7. Acceptance of the subjective thesis does not make it irrational to tell B "I understand that your indifference point now falls at odds of 3 to 7. But for you to accept the bet at those odds would be sheer foolishness. My client's claim is by no means as unlikely as you seem to think. Here is why . . . ." And, at this point, counsel gives his or her concluding statement.

In other words, the subjective view does not commit one to the position that there is no right answer. It simply recognizes that it is not always possible for persons with different attitudes, experiences, and information to arrive at identical answers. Rational discussion of subjective probability figures is therefore not barred. Thus, returning to my treatment of the paradox of the gate-crasher, I argued in part that to conclude that the subjective probability in plaintiff's favor is .501 would be to give a wrong answer. This answer would be wrong, I suggested, because it apparently overlooks the fact of non-production. In making this argument, I was asserting not only that .501 is the wrong figure for me, but also that it is wrong for any person to be satisfied with this number for his subjective probability. I think it fair to say that my discussion was located somewhere within the universe of rational discourse.

The notion that subjective probabilities are expressions of degrees of belief and that such "partial beliefs" can be rationally justified is nothing new to philosophers who study the foundations of probability and induction. A propensity theorist may be persuaded that what justifies some particular degree of belief is an objective property of an event, which he calls its "propensity." Other theorists may find the justification in some relation that the proposition partly believed bears to other propositions.

30. Rational individuals with identical utilities, data, and background information would arrive at the same subjective probabilities.
31. In suggesting that there are interpersonal standards by which subjective probabilities can be judged, I might be thought to be injecting an element of "objectivity" into subjective probability. Such objectivity would be present if the standards were not merely conventional, like those at work in the appreciation of haute cuisine, but somehow "objectively" correct. All I mean to say here is that the subjective theory treats probabilities as statements of personal beliefs about the outcomes of events. Whether these beliefs can be said to have some objective basis is an issue that is not resolved by the use of the word "subjective." Indeed, in this context, the phrase "personal probability" might be more apposite.
32. See, e.g., J. Mackie, supra note 19, at 158.
33. See, e.g., D. Mellor, supra note 19, at 2, § 18-19.
34. Carnap, for instance, wrote that a rational person should assess the probability of a proposition by its relation to his "total evidence" for and against it. R. Carnap, Rephers and Systematic
Analytical philosophers might insist that the interpersonal standards of justification to which I have alluded are implicit in some set of canonical examples of "correct" probabilistic reasoning. Needless to say, still other views are possible. The point that bears emphasizing here is that one can surely hold that subjective probabilities are well suited to the study of forensic proof while firmly maintaining a discreet agnosticism as to which camp of philosophers has found the correct answer to the question of what it really is that makes one degree of belief more justified than all other degrees of belief.

I should be surprised if this defense of subjective probability against the charge of standardless subjectivism will prove entirely satisfying to Dr. Cohen. It does not, after all, go to the root of the philosophical problem. Nevertheless, it is all that is needed to cope with the issue of naked statistical evidence. Dr. Cohen's arguments not only fail to elucidate any irreparable incongruity between mathematical probability theory and the standard of proof in civil litigation; they also do not reveal any insurmountable conceptual barrier to using the subjective interpretation to understand the probabilities implicitly at work in legal factfinding. Dr. Cohen's latest paper is thus much like his erudite volume, *The Provable and the Probable*. Both works are replete with intriguing insights, provocative passages, and ingenious arguments. Yet, for all the creativity and resiliency displayed in defense of the claim that the paradox of the gatecrasher points up some fundamental dissonance between mathematical probability theory and forensic proof, the case for Dr. Cohen's "Baconian" probabilities is considerably less powerful than Dr. Cohen will as yet admit.