38. Adolf Grünbaum, "Can a Theory Answer More Questions Than One of Its Rivals," *British Journal for the Philosophy of Science* 27 (1976): 1–23.

39. Steven Weinberg, *Gravitation and Cosmology* (New York: John Wiley & Sons, 1972), 198. Note that this correction is smaller by an order of magnitude than the correction of 43 seconds of arc per century for Mercury.

40. See Joel Smith, *The Status of Inconsistent Statements in Scientific Inquiry* (doctoral dissertation, University of Pittsburgh, 1987).

41. Unfortunately, recent evidence regarding the mass of Pluto strongly suggests that Pluto is not sufficiently massive to explain the perturbations of Neptune. A different plausible scenario is needed, but I do not know of any serious candidates that have been offered.

42. See "Dynamic Rationality" (note 26 above), 5–12, for a more detailed discussion of various grades of rationality. The term "static" was chosen to indicate the lack of any principled method for changing personal probabilities in the face of inconsistency or incoherence.

43. Ibid., esp. pp. 11–12

44. Thomas S. Kuhn, *The Copernican Revolution* (Cambridge: Harvard University Press, 1957), 3–4.

Bayesian Problems of Old Evidence

1. Introduction

According to "standard" or "classical" Bayesian confirmation theory, a piece of evidence E confirms a theory or hypothesis T, for a given person, if and only if $Pr(T / E) > Pr(T)$, where Pr is the relevant person's subjective (personal) probability function, representing this individual's degrees of belief. $Pr(T / E)$ is the individual's probability for T conditional on E, $Pr(T\&E) / Pr(E)$, which Bayesians argue is equal to the degree of confidence the individual should have in T if, and after, E is learned. In defenses and further "articulations" of this theory, Bayesians have assumed that these "rational degrees of belief" satisfy the standard probability axioms, where the appropriateness of this assumption has been defended in various ways, including "Dutch book" and decision theoretical approaches.[1]

Clark Glymour (1980, 85–93) has recently raised an interesting problem for this theory. It is not uncommon for scientists to find support for a theory in evidence known long before the theory was even introduced, so that, intuitively, there are cases of already known, or "old," evidence confirming "new" theories or hypotheses. Glymour cites the examples of the support for Copernicus' theory derived from previous astronomical observations, the support for Newton's theory of gravitation derived from the already established second and third laws of Kepler, and the support for Einstein's gravitational field equations derived from the already known anomalous advance of the perihelion of Mercury. But if evidence E is already known before theory or hypothesis T is invented, then $Pr(E)$ already equals 1 at that later time, so that, at that later time, $Pr(T / E)$ *must equal*

$Pr(T)$; this follows from the usual axioms of probability and the definition of $Pr(T / E)$. Thus, Bayesian confirmation theory seems to imply that already known evidence *cannot* support newly invented theories, contrary to what seems true in the cases Glymour cites.

After some further clarification of this problem for Bayesian confirmation theory in section 2 below—which will help isolate the version of the problem of old evidence that poses the most potent threat to the theory—I shall in section 3 consider the principal line of defense that has been offered in support of Bayesian confirmation theory in light of the problem. I believe that the general idea embodied in such responses is sound. Roughly, the strategy is: (1) to suppose, contrary to standard or classical Bayesianism, that rational scientists are *not* "logically omniscient" at all times over the set of propositions entertained (as explained below), and then, (2) to argue that, in cases of the kind Glymour cites, there really *is* new evidence, namely, the discovery of some *logical relation* between E and T (which relation might suggest, for example, that the truth of T would *explain* E). In sections 4 and 5, however, I will argue that the ways in which the two components of this strategy have been developed are not adequate, and I will suggest more promising ways of developing them.

2. Clarification of the Problem

Based on a suggestion by Brian Skyrms, Daniel Garber (1983) has suggested that there is an ambiguity in Glymour's problem, and Garber divides the problem into two separate problems, which he calls the "historical" and "ahistorical" problems of old evidence. Garber argues (quickly) that his ahistorical problem can be solved by means of some variant of what he calls a "counterfactual strategy." This strategy has also been discussed, and criticized, by Glymour (1980); the strategy and some of its difficulties will be briefly described later in this section. Since I wish to avoid this strategy, I shall here segment the problem somewhat differently, in a way that I think will allow for more plausible "quick" resolutions of several of its variants. Below I'll also describe the relation between my way of dividing up the problem and Garber's.

In all versions of the problem to be described, with two exceptions to be noted later, I shall assume that the correct assessment of the relation between the evidence E and theory or hypothesis T in question is that E is in fact positively evidentially relevant to T. (I hope that the rationale for the perhaps somewhat awkward terminology I shall employ for labeling the three main problems I shall distinguish will become evident as the discussion unfolds; also, see the outline below.) What I shall call "The Problem Of *Old New* Evidence" arises in cases in which one *first* formulates a theory or hypothesis T and *subsequently* discovers evidence E, which is thus "new" in relation to the time of the formulation of T. In cases of the problem of this first kind, at the point in time at which E is learned,

E in fact *does* increase one's confidence in T. However, later on, when E becomes "old" in relation to the time of its discovery, E can no longer thus increase one's confidence in T. Nevertheless, there seems to be a valid sense in which E is *still* good evidence for T. So, at such later times, E would, in this sense, seem to confirm T—it is in this sense still good evidence for T—even though $Pr(T / E) = Pr(T)$ so that the Bayesian theory says that E does *not* confirm T.

What I shall call (henceforth) simply "The Problem Of *Old* Evidence" arises in cases in which E is learned first, and T is formulated subsequent to the learning of E. This problem can be subdivided in turn. "The Problem Of *Old Old* Evidence" arises at times *after* the formulation of T: if E somehow confirmed T at or around the time of the formulation of T (even if E had probability 1 at or around the time of the formulation of T), then it would seem that, even later on, E is still, in a valid sense, good evidence for T. And "The Problem Of *New Old* Evidence" concerns how, at the time of the formulation of T, E can confirm T even though E already had subjective probability 1. Finally, it will be helpful to divide each of these latter two problems into two cases. In Case 1, T was specifically designed to explain E; and in Case 2, it was not.

For easier reference, here in outline form are the versions of Glymour's problem to be discussed:[2]

I.   The Problem of *Old New* Evidence: T formulated before the discovery of E; but it is now later and $Pr(E) = 1$, so that $Pr(T / E) = Pr(T)$

II. The Problem of *Old* Evidence: E known before the formulation of T

  A. The Problem of *Old Old* Evidence: It is now some time subsequent to the formulation of T.
    1. T originally designed to explain E
    2. T not originally designed to explain E

  B. The Problem of *New Old* Evidence: It is now the time (or barely after it) of the formulation of T
    1. T originally designed to explain E
    2. T not originally designed to explain E

Garber's ahistorical problem of old evidence, as I understand it, arises in cases I and IIA above, and the historical problem arises in case IIB. Roughly, the counterfactual strategy endorsed by Garber is to argue that, in cases I and IIA, if, now (sometime after both the discovery of E and the formulation of T), we *had not* known E (e.g., if our education in the history of science had been incomplete), then the probability $Pr(T / E)$ *would have been* greater than the probability $Pr(T)$. Thus, on the modification of Bayesian confirmation theory suggested by Garber, appeal is made to *counterfactual degrees of belief*.

Garber admits that there are "some details to be worked out here" (1983, 103), and, as noted above, Glymour (1980) has criticized the strategy. For example, Glymour and Garber both point out that there will not necessarily be any *particu-*

lar degrees of belief that we can say a person would have had in E, or in T given E, if this person's degree of belief in E had been less than 1. Indeed, it seems plausible that in some cases T *would not even have been formulated* had E not been learned. And surely there also will be cases in which the person's knowledge of E *saved his life* at some time in the past, so that had the individual's degree of belief in E been less than 1, the person would be *dead* now. Also, there are of course the well-known difficulties attending the proper interpretation of counterfactual conditionals that would befall any such modification of Bayesian confirmation theory.

Brian Skyrms (1983) has discussed ways of "giving a probability assignment a memory," so that one may, in one way or another, retain information about one's old *actual* degrees of belief in E, and in T given E. This kind of approach is more promising, I think, and it is closer to the approach I shall describe below. To apply this kind of approach, however, it is necessary first to divide Garber's ahistorical problem into problems I and IIA above, for such an approach works only for the first of these.

The problem of old new evidence (problem I) can be handled quite plausibly, I think, as follows. One of the central tenets of Bayesian confirmation theory is that confirmation is a relation between three things: a piece of evidence, a hypothesis or theory, *and a set of background beliefs*. As background beliefs change over time (as well as from person to person), *so does what confirms what*—where, of course, our background beliefs, of various degrees, are given by our subjective probability assignment. At the time (or just before) E was learned, E was not one of our background beliefs of degree 1. At that time—that is, relative to the set of background beliefs that includes E only to some intermediate degree—our degree of belief in T is less than our degree of belief in T conditional on E, so that at that time, E does confirm T according to the Bayesian theory. However, after E is learned, and we face problem I above, *our background beliefs have changed*. At *this* time—that is, relative to the new set of background beliefs—E (or being again told that E, or pondering the discovery of E) can no longer increase our degree of confidence in T; although it did so once, it cannot do so again. I think it is quite natural to say simply that, because of the change in our background beliefs, E simply does not confirm T at the later time, after its evidential impact on T has already been "absorbed." The Bayesian theory of confirmation was *designed* to reflect this possibility: what confirms what, depends on one's background beliefs. What now remains to be explicated in Bayesian terms, however, is that sense in which E may *remain good evidence* for T.

I think it is good idea to distinguish E's *actually confirming* T from E's *being (actual) evidence in favor of* T, as follows. Recall that Bayesian *confirmation* is a relation that obtains between evidence E, theory or hypothesis T, and degrees of belief Pr if and only if $Pr(T / E) > Pr(T)$. Whether or not this relation obtains is independent of whether or not T ever actually gets confirmed: the relation may

obtain, or not, independently of whether or not E will ever get discovered. But we may say that *actual confirmation* is an *event*, involving E, T, and a person with a subjective probability assignment Pr, that takes place when the three-part relation $Pr(T / E) > Pr(T)$ obtains and E is learned. In cases of the problem of old new evidence, such an event happened in the past, once, and it cannot recur (with the same piece of evidence and the same ideal Bayesian agent). As to the second idea, it seems appropriate to say that E *is (actual) evidence for* T, for a given individual, if, at some time in the past, the event of its confirming T, for that individual, took place. Indeed, this seems to be exactly what it means for E to constitute part of our current body of evidence for T: at some time in the past E raised our rational degree of confidence in T. In cases of the problem of old new evidence, therefore, it is clearly consistent and appropriate to say *both* that E confirmed T but no longer does *and* that E is now, but before its discovery was not, part of our body of evidence in favor of T.

To be more precise about the idea of E's being evidence for T, we have to say that it is a relation that may obtain among *four* things: (1) a piece of evidence, (2) a hypothesis or theory, (3) a *time*, and (4) a *history of a set of background beliefs (i.e., a sequence of subjective probability functions indexed by times)*. Roughly: E *is, at time t, part of the body of evidence in favor of theory T relative to history H of background beliefs* if and only if, at some time prior to t in the history H (of a set of background beliefs), the confirmation event took place between E, T, and the state of H (the relevant probability assignment) at that earlier time. Relativity to a history of a set of background beliefs is essential. For it is not difficult to invent cases (or find actual cases) in which, relative to one history, E is now evidence for T, but, relative to an alternative history, E is now evidence *against* T (i.e., *for* $\sim T$). For example, if initially $Pr(T) = 0.5$, $Pr(T / E) = 0.7$, $Pr(T / F) = 0.9$, and $Pr(T / E\&F) = 0.8$, then whether E is evidence for or against T will depend on whether E is learned before or after F. Thus the relativity of *confirmation* to a set of background beliefs shows up in this Bayesian conception of *evidence*. There is no univocal matter of fact about whether or not an E confirms a T or about whether a known E is evidence for or against a T: the first depends on a set of background beliefs, and the second depends also on the history of a set of background beliefs. On this conception of confirmation and evidence, whether E confirms, disconfirms, or is neutral for T, and whether a known E is evidence for, against, or neutral towards T, both involve an element of historical accident pertaining to what our background beliefs, and their history, happen to be.

It might be objected that it is inappropriate to let whether or not E is part of our body of evidence in favor of T depend on *when* it was learned in relation to other evidence (as in the numerical example above). I am not entirely sympathetic (or entirely unsympathetic) with such an objection, but in any case there is a natural framework in which it can be accommodated (suggested to me by Brian

Skyrms in correspondence). Suppose we begin with an initial probability assignment Pr_0, and as evidence $E_1, E_2, \ldots, E_n$ (n = "now") comes in, we update: $Pr_{i+1}(-) = Pr_i(-/E_{i+1})$. Then it is natural to say that E (= E_i, say) is now part of our body of evidence in favor of T if $Pr_n(T) = Pr_0(T / E_1 \& \ldots \& E_n) > Pr_0(T / E_1 \& \ldots \& E_{i-1} \& E_{i+1} \& \ldots \& E_n)$. On this explication, whether or not E is part of our body of evidence in favor of T does not depend on the order in which the evidence comes in: if we permute the subscripts on the E_is, the verdict remains the same. Note also that the "prior" used need not be actual.

As to the problem of *old* evidence (problem II above), I think we can quickly dispense with versions A1 and B1. If our knowledge of E inspired and guided our formulation of theory T, where the intention was to give an explanation of E, then it would seem that E does not confirm T in the first place, and E does not constitute evidence in favor of T. In this case, the Bayesian theory gives the right answer: for both IIA1 and IIB1, we have $Pr(T / E) = Pr(T)$. (Compare Garber 1983, 104.) This is not to say that T must be completely *without support*, however. T may derive support from the fact (if it is a fact) that it has some of the virtues usually thought to attach to good theories, such as simplicity, analogy with other well-confirmed theories, independent evidence, and the very fact (if it is a fact) that its truth *would* successfully explain E, and so on. Indeed, it would seem that if T had none of these virtues, then it *would* be entirely without support, despite the truth of E; therefore, if T does have support, it is not from E itself.

This leaves problems IIA2 and IIB2. But if problem IIB2 can be solved, so can IIA2, by employing the very same considerations employed in resolving problem I earlier. If it can be shown, in resolving IIB2, that *at the time of the formulation of T, T gets confirmation*, in the Bayesian sense, in virtue of some aspect of its relation to E, then, for problem IIA2, we should say that whatever exactly it was that confirmed T at the earlier time *no longer does so* at the later time, even though it will still be, in the sense clarified earlier, a part of our body of *evidence* in favor of T. Let us thus turn to problem IIB2.

3. The Basic Bayesian Defense

If, in cases of problem IIB2, T *does* receive confirmation in the Bayesian sense explained above, then, *when* it does, there must be some proposition F such that (1) one's subjective probability of F increases from some value short of 1 to 1, and (2) the prior probability of T is less than the posterior probability of T (i.e., the prior probability of T conditional on F). For reasons already explained, F cannot be the same as E. Daniel Garber (1983) and Richard Jeffrey (1983) have recently sought to show how some proposition other than E might plausibly play the role of F.

The *total* strategy in these defenses involves basically three steps; and the two authors concentrate on different steps of the total strategy. First, it is argued that

classical Bayesianism's assumption of "logical omniscience" is clearly unrealistic, and this first step includes some "nonclassical" formulation of Bayesianism that is intended to be more realistic yet at the same time appropriately logically restrictive on rational agents' degrees of belief. Classical Bayesianism's assumption of logical omniscience may be formulated as follows: All logical truths have subjective probability 1; and if propositions A and B are logically incompatible, then $Pr(A \lor B) = Pr(A) + Pr(B)$. These two conditions, together with the condition that all subjective probabilities are greater than or equal to 0, are just the usual axioms of subjective probability theory. Of course, given the "richness" of the languages we use, it is very unrealistic to suppose that any rational scientist's degrees of belief will be sensitive to *all* the logical facts encompassed by these axioms. So it is desirable, even independently of the problem Glymour has raised for Bayesianism, to formulate weaker versions of the usual axioms of subjective probability.

The second step of the total strategy is to describe a *logical relation* that holds between E and T—or between E, T, and others of one's beliefs—whose obtaining perhaps suggests that the truth of T (or the truth of T in the presence of one's background beliefs) would explain E. It is then argued, in the third step, that it is not E itself that confirms T in cases of problem IIB2, but rather the discovery of this logical relation between E and T. That the obtaining of this relation *can be discovered*, and need not have been known all along, should now be a possibility in light of the successful completion of step one of the strategy. Thus, where "$T \vdash E$" expresses the proposition that E and T are so logically related to each other (and perhaps also to one's background beliefs), step one of the strategy makes it theoretically possible that $Pr(T / T \vdash E) > Pr(T)$. It is the task of step three to argue that this inequality *should* indeed obtain.

Roughly speaking, Garber focuses on step one, without providing much in the way of argument for step three, while Jeffrey concentrates on step three; for various reasons (good ones I think, as explained below), neither provides much detail in the way of carrying out step two.

As to step one, Garber advances a theory of what might be called "limited logical omniscience." He begins with a truth-functional language L, with atomic sentences a_i, and builds language L^* from L by adding new atomic sentences $A \vdash B$, where A and B are any sentences of L (i.e., truth-functional compounds of the a_is). Treating sentences $A \vdash B$ as atomic has the effect of making each of them *formally*, in L^*, logically independent of all the other atomic sentences: L^*-atomic sentences neither L^*-imply nor are L^*-implied by other L^*-atomic sentences. Thus, there are no axiomatic constraints on what subjective probabilities (between 0 and 1) may be assigned to atomic sentences: any sprinkling of numbers between 0 and 1 (inclusive) on the atomic sentences is allowed.

"Extrasystematically," however, we will want to understand "$\vdash$" as meaning, say, *implies*, or *explains*, and this makes it desirable to put some sort of formal constraint on the relation. Recall that step two of the basic Bayesian response to

the problem is to identify the appropriate logical relation that is discovered between T and E; "$\vdash$" is intended to be interpreted as denoting that appropriate relation. Let us thus briefly digress from step one and consider Garber's treatment of step two. Garber doesn't actually insist on any particular interpretation of "$\vdash$". He states:

> Depending on the context of investigation, "$\vdash$" may be understood as truth-functional implication, or implication in . . . the global language of science. We can even read "$h_i \vdash e_i$" as "e_i is a positive instance of h_i," or as "e_i bootstrap confirms h_i with respect to some appropriate theory," as Glymour demands. (1983, 112)

I agree with Garber that it need not be part of an adequate solution to the problem Glymour has raised that a particular interpretation of "$\vdash$" be specified, i.e., that a particular logical relation between T and E be described. For, as Garber suggests, different interpretations of "$\vdash$" may be appropriate in different particular, actual cases of problem IIB2. What relation will be appropriate depends, as Garber suggests, on the context of investigation. Indeed, in different actual investigations, different relations (of the appropriate kind) between an E and a T *will be* discovered, assuming that Garber's general approach to problem IIB2 parallels what transpires in actual cases of IIB2. Of course, the Bayesian defense would be strengthened if analyses of actual, historical cases of problem IIB2 could be given, supplying particular appropriate interpretations of "$\vdash$". For the most part, in any case, both Garber and Jeffrey use the interpretation of "$\vdash$" as "logically implies" to guide some of their intuitions and support various moves in their analyses; hence, one would suppose, the symbolism.

In any case, some formal constraint or other on $\vdash$ would seem to be desirable, if plausible, given the more or less vague intended interpretation. Garber assumes:

$$(K^*) \; Pr(A \; \& \; B \; \& \; A \vdash B) = Pr(A \; \& \; A \vdash B).$$

This principle guarantees that if a person assigns subjective probability 1 to A and to $A \vdash B$, then this individual will also assign probability 1 to B. Thus, $\vdash$ will behave somewhat like implication in classical Bayesianism, although, as Garber points out, K^* by itself doesn't rule out its being interpreted as conjunction, for example, or biconditionalization.

Garber distinguishes what he calls "global Bayesianism" from what he calls "local Bayesianism," and the distinction shows up when we look, *extrasystematically*, into the structures of the L^*-atomic sentences a_i and $A \vdash B$. As Garber explains, while these are atomic sentences of L^*, they may, from a broader perspective, have complex logical structure: truth-functional structure, quantificational structure, modal logical structure, and so on—including of course a special interpretation of "$\vdash$" in the case of the $A \vdash B$s. In virtue of the internal structures

of the L^*-atomic sentences, some of them may be logically true from such a broader perspective. Global Bayesianism includes the thesis that those that are logically true, whether from the broad perspective or just from the point of view of L^*, must be assigned subjective probability 1. Local Bayesianism requires only that those that are logically true from the perspective of L^*—i.e., *those that are tautological truth-functional compounds of the a_is and the $A \vdash B$s*—be assigned subjective probability 1. (A similar argument applies to extrasystematically logically incompatible L^*-atomic sentences.)

Garber proposes that local Bayesianism can handle problem IIB2, since although it may be *true*—even logically true—that $T \vdash E$ (when "$\vdash$" is interpreted), it is nevertheless allowable that $Pr(T \vdash E) < 1$, so that it may also be the case that $Pr(T/T \vdash E) > Pr(T)$. In cases of problem IIB2, according to Garber, it is $T \vdash E$, and not E, that actually confirms T; and although E is old, $T \vdash E$ may not be old, for "local Bayesians."

I agree that Garber's defense does all that needs to be done in the way of carrying out step two of the strategy, as explained above. However, there are serious problems with his approach to step one, as I shall argue in the next section. In addition to taking a more critical look at Garber's theory of local Bayesianism, I shall also in the next section examine the approach to a more realistic Bayesianism offered by Hacking (1967). I then will try to characterize a more adequate kind of approach to step one (without, however, attempting actually to carry out such an approach). Finally, in section 5, I shall turn to step three of the strategy.

4. Bayesianism and Logical Fallibility

Although I sympathize with the idea that axioms of subjective probability theory should be weakened to allow for the failure of logical omniscience in rational individuals, I think Garber's approach does not go nearly far enough in the way of allowing logical fallibility. To extend Garber's analogy, his method is to draw a "line" demarcating the "local" from the "nonlocal," and then to insist on logical omniscience only on the local side of the line. The line Garber draws (as an example of the approach at least) is, so to speak, between truth-functional logic and logic that attends to more features of the logical form of statements than truth-functional logic does. That is, Garber's approach requires rational locally Bayesian agents to assign probability 1 to all *tautologies* (of L^*), and to recognize all cases of pairs of sentences that are logically incompatible in virtue of their (L^*) truth-functional logical structure (in the sense that the subjective probability of their disjunction will in each such case equal the sum of their subjective probabilities). The same need not hold for cases of sentences (namely, some "atomic" sentences of L^*) that are logically true in virtue of non-truth-functional (or even perhaps, as far as L^* is concerned, "extrasystematic" *truth-functional*) features of

their form, or for pairs of sentences that are logically incompatible in virtue of ($L*$) non-truth-functional features of their forms.

But this seems to be an inappropriate place to draw the relevant "line." For there are extremely complex tautologies (of $L*$), so complex that it would be more difficult to recognize them as logically true than to recognize as logically true certain simple sentences that are logically true in virtue of their (say) quantificational logical form. Simple sentences of the form "For all x, if Fx then Fx," would be atomic in $L*$. And it seems completely inappropriate not to require an agent to assign probability 1 to such sentences, while at the same time insisting that the agent assign probability 1 to arbitrarily complex tautologies of $L*$.

Of course the choice of making L and $L*$ *truth-functional* languages is just an example. They could instead be first-order languages, where sentences containing modal logical structure, second-order quantifiers, and so on, would be considered atomic. However, the same objection would apply to any such proposal: There will always be extremely complex logically true sentences of the local language, and extremely simple logically true sentences "outside" the local language, where it will be inappropriate to insist on probability 1 for the former while not so insisting in the case of the latter. Thus, in one way, Garber's theory of local Bayesianism requires too much, and in another, too little.[3]

Of course there may be some atomic sentences a_i of $L*$ that, from an extrasystematic point of view, are extremely *truth-functionally* complex. Some of these may be tautologies from the extrasystematic point of view. Local Bayesianism does *not* require subjective probability 1 for them. But this seems arbitrary in view of the fact that local Bayesianism *does* require subjective probability 1 for tautologies of $L*$ that have *exactly the same truth-functional form* from the point of view of $L*$ as extrasystematically extremely complex tautological $L*$-atomic sentences have from the extrasystematic point of view. That is, an $L*$-atomic sentence a_i may be an extremely complex tautology from an extrasystematic point of view, and have the same complex form from that point of view as a sentence A has from the point of view of $L*$. Yet local Bayesianism requires subjective probability 1 for A and allows any subjective probability for a_i.

Suppose some evidence statement E and some theory T are, extrasystematically, quite complex, but complex *only with respect to the kind of logical form to which the local language is sensitive*. As an example, let's say E and T are quite complex *truth-functionally*, from the extrasystematic point of view. Suppose also that T truth-functionally implies E from the extrasystematic point of view. From the point of view of the local language, however, E, T, and $T{\vdash}E$ are all atomic. Suppose that we are in a case of problem IIB2, so that $Pr(E) = 1$. Now it must be true that $Pr(T{\vdash}E) < 1$ if Garber's approach is to work. But it seems quite arbitrary to think that $Pr(T{\vdash}E)$ may be less than 1, while there are sentences A and B of the local language that have the same forms from the point of view of the local language as T and E have from the extrasystematic point of view,

respectively, so that $Pr(A \supset B)$ has to equal 1. In this case, $T \supset E$ *has the same truth table* from the extrasystematic point of view as $A \supset B$ has from the point of view of the local language! And if one assigns subjective probability 1 to the conditional $A \supset B$ out of local logical omniscience—in virtue of having perceived the truth-functional logical connection between A and B—then it is hard to see why this individual would miss the (identical kind of) logical connection in the case of T and E.[4]

The point here, of course, is that Garber's local Bayesianism makes the prospects for successfully carrying out step three of the general strategy look very bleak. In the case described in the previous paragraph, for example, it seems very implausible that $Pr(T / T{\vdash}E) > Pr(T)$, if it is required that the agent be logically omniscient with respect to the parallel logical relations between A and B.

It is worthwhile considering an alternative way of "drawing the line," advanced by Ian Hacking in his well-known article "Slightly More Realistic Personal Probability" (1967). Hacking first notes that the axioms of subjective probability theory can be stated in terms of the idea of logical possibility, rather than the ideas of logical truth and incompatibility. Thus, the usual axioms can be restated as follows (this is not the particular axiomatization considered by Hacking in his article):

For all propositions A and B:

    (1) $Pr(A) \geq 0$;

    (2) $Pr(A) = 1$, if not$-A$ is not *logically possible*; and

    (3) $Pr(A \lor B) = Pr(A) + Pr(B)$, if $A \& B$ is not *logically possible*.

Hacking then suggests that it is unrealistic to assume these axioms for subjective probability, since it is unrealistic to assume that a rational agent will always be able to recognize cases of logical possibility and logical impossibility as such. Instead, we should assume axioms stated in terms of *personal possibility* and *sentences*, where a sentence (see below) is personally possible if the relevant individual *does not know* (in a special sense of "know," see below) that the sentence is false. Thus, a sentence is only required to have subjective probability 1 if it is *known* to be true (rather than: if it is logically true), and the probability of a disjunction is only required to have subjective probability equal to the sum of the probabilities of its disjuncts if it is *known* that the conjunction of the two disjuncts is false.

The reason for stating the axioms in terms of sentences rather than propositions is that "proposition" is usually understood in such a way that "two logically equivalent propositions" are really the same proposition (Hacking 1967, 318, cites Carnap 1947, 27). Thus, it would be absurd to entertain the possibility that an agent knows one proposition but fails to know a *different but logically equivalent* proposition. However, we want to allow the possibility that an agent can know one thing, but—in part because the agent does not know the relevant logical

equivalence – does not know a second thing that is logically equivalent to the first. So it is natural to say that what the agent knows to be true is one *formulation* of a proposition and what the agent fails to know to be true is a different but logically equivalent *formulation* of the same proposition. Thus, (unambiguous) sentences are a natural choice for the objects of personal probability, sentences that *express* propositions, in a personal language closed under the linguistic connectives of negation, conjunction, disjunction, and conditionalization, for example. (For presumably the same reason, Garber's theory of local Bayesianism is also formalized in such a way that the objects of subjective probability are sentences – of formal languages, in fact.)

Hacking's special sense of knowledge – called "the examiner's view of knowledge" – is one in which certain traditional closure conditions for knowledge are explicitly rejected, e.g., "a man can know how to use *modus ponens*, can know that the rule is valid, can know *p*, and can know $p \supset q$, and yet not know *q*, simply because he has not thought of putting them together" (1967, 319). It is clear that such closure conditions must fail, if we relax the classical assumption of logical omniscience. Otherwise (for example), the agent must assign probability 1 to all first-order logical truths if (*roughly*) probability 1 is assigned to all the axioms of a (complete) deductive calculus in which the only rule is *modus ponens*.

I think Hacking's way of making subjective probability theory more realistic – of abandoning the assumption of logical omniscience – goes too far. The axioms are much too weak. However, one of the ways in which it has been argued that classical personalism – the classical axioms assuming logical omniscience – are reasonable is by way of "Dutch book arguments." And Hacking suggests a revised Dutch book argument to show the reasonableness of his weaker axioms.

One of the ways in which it has been attempted to justify classical personalism is to prove that if a person's subjective probabilities do *not* satisfy the classical axioms, and if the person is willing to accept any bet whatsoever whose odds are determined in the natural way from the person's subjective probabilities, then it is possible for a clever betting opponent to offer bets, all acceptable to the agent, such that *no matter which propositions bet on turn out to be true and which false*, the agent is assured of a loss.[5] The clever betting opponent need know no more than the agent *about matters of fact* in order to identify a series of bets (called a "Dutch book"), each acceptable to such an agent, but that will assure a net loss to the agent. All that is necessary is that the opponent be able to detect the "incoherence" (violation of the classical axioms); then, using simple mathematical techniques, a Dutch book can be found. The ability to detect an incoherence requires only logical and mathematical sophistication – and not knowledge of matters of fact.

Hacking, however, wants to put knowledge of matters of fact and knowledge of logical facts on a par, for the purpose of assessing a person's rationality.[6] Thus,

Hacking suggests that an appropriate betting opponent for carrying out a Dutch book argument must be *one who knows no more than the agent in the examiner's sense of knowledge*. But this has the consequence that if the agent is *unaware of an incoherence* in his subjective probabilities, then so must be an appropriate betting opponent. But this means that a person will turn out to be rational in the Bayesian sense as long as the person is *not aware of an incoherence*.

Put in other terms, if it is personally possible to *you* that no Dutch book can be made against you, then this must also be personally possible to any appropriate betting opponent, so that the opponent couldn't know of a Dutch book against you either. So it seems that Hacking's slightly more realistic personal probability requires only that you not know that you're not coherent. But this seems to be too severe a weakening of classical personalism.

Thus, while Garber's approach seems to require too much of a rational agent in one sense, and too little in another, Hacking's approach simply requires too little. Recall that the central idea behind all the difficulties raised for Garber's local Bayesianism is that of *complexity* of sentences, logical relations, and inferences. The counterexamples all suggested (roughly) either that "local Bayesianism requires that the agent perceive such-and-such enormously complex logical fact," or that "local Bayesianism does not require the agent to perceive such-and-such extremely simple logical fact." From this point of view, the problem with Hacking's slightly more realistic personalism is that it *allows the agent to set the standard* governing how complex a logical fact has to be in order for him not to be required to perceive it, and its implications, and yet still be considered to be rational on the theory. All this suggests, to borrow Garber's analogy again, that the appropriate place to "draw the line" between the logical facts an agent has to perceive and those that one need not perceive, in order to be considered rational, should correspond to the *complexity* of the facts on the two sides of the line. Or perhaps we should conceive of rationality as coming in (objective) degrees, corresponding to where the line in fact falls for particular agents. Although the development of such a measure is, of course, beyond the scope of this paper, I think we have seen plenty of considerations indicating that, if such a measure *could* be defined, then it would be the appropriate tool for use in developing a version of Bayesianism that is truly more realistic than classical Bayesianism and yet at the same time still reasonably restrictive in the right way.

5. The Evidential Significance of $T \vdash E$

I turn now to part three of the basic Bayesian defense – an argument to the effect that $Pr(T / T \vdash E)$ should be greater than $Pr(T)$ before the discovery of $T \vdash E$. The main thing Garber does in this connection is show that it is *possible* that $Pr(T / T \vdash E) > Pr(T)$. That is, he proves that, under some fairly general conditions, there *are* probability functions such that $Pr(T / T \vdash E) > Pr(T)$ (and

(K^*) is satisfied). No argument is given that one's subjective probability function *should* satisfy this inequality in the relevant kind of situation, although Garber points out in a note (1983, 131) that the discovery of $T \vdash E$ will increase the probability of T if and only if the probability of $T \vdash E$ is higher given T than given not-T. Richard Jeffrey (1983), in the part of his article about the problem of old evidence, basically assumes the adequacy of Garber's steps one and two, and, taking a hint from Garber's note, provides an argument for step three.

Here is a simplified version of Jeffrey's main result. Jeffrey proves that $Pr(T \mid T \vdash E) > Pr(T)$ if the following four conditions obtain:

(1) $Pr(E) = 1$ and $Pr(T) > 0$;

(2) $Pr(T \vdash E)$ and $Pr(T \vdash \sim E)$ are both strictly between 0 and 1, and $Pr(T \vdash E \ \& \ T \vdash \sim E) = 0$;

(3) $Pr(T \mid T \vdash E \ \mathrm{v} \ T \vdash \sim E) \geq Pr(T)$; and

(4)   Garber's condition (K^*) (see above), in particular, $Pr(T \ \& \ T \vdash \sim E) = Pr(T \ \& \sim E \ \& \ T \vdash \sim E)$.[7]

The proof of this version is simple. In view of (3), it suffices to establish

$$Pr(T \mid T \vdash E) > Pr(T \mid T \vdash E \ \mathrm{v} \ T \vdash \sim E),$$

which, given (2), is true if and only if

$$Pr(T \mid T \vdash E) > Pr(T \vdash E \mid T \vdash E \ \mathrm{v} \ T \vdash \sim E)Pr(T \mid T \vdash E) \\ + Pr(T \vdash \sim E \mid T \vdash E \ \mathrm{v} \ T \vdash \sim E)Pr(T \mid T \vdash \sim E).$$

(1), and an application of (4), implies that the second term on the right-hand side equals 0. (1) and (3) imply that the right-hand side is greater than 0, so that $Pr(T \mid T \vdash E) > 0$. And (2) implies that $Pr(T \vdash E \mid T \vdash E \ \mathrm{v} \ T \vdash \sim E) < 1$, giving us the desired inequality.

Let us consider the four conditions. Condition (1) is part of the specification of the problem (we assume, of course, that T initially enjoys *some* credence). Condition (2) is plausible in light of the agent's *logical nonomniscience* and the intended interpretation of "$\vdash$", as long as we assume that the agent fully believes that T is not inconsistent. Condition (3) will be discussed below; and condition (4) just specifies part of the intended interpretation of "$\vdash$".

Condition (3) expresses the idea that "your confidence in [T] would not be weakened by discovering that it implies something about [the relevant phenomenon]" (Jeffrey, 1983, 150). Conversely, in order for (3) to be true, it must also be the case that your confidence in T *would* be weakened (or left unchanged) by the discovery that it does *not* imply either E or $\sim E$. There is, however, the intuition that the more a hypothesis or theory implies (the "stronger" it is logically), the less chance it has of being true—an intuition that says more, I think, than just that the *probability* of a hypothesis can be no greater than propositions it implies.

And it seems odd that a theory should be *disconfirmed* just by the fact that it is *silent* on a certain issue. Here, given that T implies *something* about the relevant phenomenon, there is, of course, always the chance that it implies *something false* about it. Thus, in order for (3) to be true, we must be antecedently relatively more confident that, *if T implies *something* (E or $\sim E$) about the relevant phenomenon, then it implies a *truth* about it (i.e., E), than we are that T would imply *a falsehood* about it (i.e., $\sim E$). But this seems to run against the spirit of the idea of allowing the agent to be *logically nonomniscient*, as I shall presently explain.

Condition (3) (in the presence of the other conditions) implies that:

$$\frac{Pr(T \vdash E)}{Pr(T \vdash E) + Pr(T \vdash \sim E)} \times Pr(T \mid T \vdash E) \geq Pr(T).$$

Suppose now that the agent is "so logically nonomniscient" that $Pr(T \vdash E) = Pr(T \vdash \sim E)$. This is not implausible if T and E are sufficiently complex in the right way. Now, given that T was not designed just with the intention of explaining E, T may already have been confirmed somewhat; say $Pr(T)$ is equal to 0.6. In that case, it is clear from the last displayed inequality that it is impossible for condition (3) to be satisfied. Note also that the agent's having the same degree of confidence in $T \vdash E$ as in $T \vdash \sim E$, while at the same time assigning probability 1 to E, is not incompatible with (4), together with a high degree of confidence in T, for the agent's degree of confidence in $T \vdash E$ and in $T \vdash \sim E$ may be quite low.

Thus, in order for (3) to be satisfied, the agent must *either* assign a high probability to $T \vdash E$, compared with the probability assigned to $T \vdash \sim E$, *or* have a relatively low degree of confidence in T—or both. When would we expect this to be true? To me, this disjunctive condition strongly suggests (though strictly speaking it doesn't imply) that T *was designed to explain E*. This hypothesis would certainly *explain* why $Pr(T \vdash E)$ is much greater than $Pr(T \vdash \sim E)$, if it is: the agent thinks he is pretty good at coming up with theories that would, if true, explain things. And it would also explain a low initial degree of confidence in T: the theory hasn't been around very long and thus has not received *independent* confirmation. But if T was designed to explain E, then we have only a solution to problem IIB1, above, and not to IIB2. (Earlier it was noted that, for IIA1 and IIB1 situations, T may derive support from its explaining E, but not from E itself.)

Regardless of whatever connection there may be between the disjunctive condition just discussed and cases in which T was invented with the intention of explaining E, we have seen that there is what would seem to be an important class of cases outside the scope of Jeffrey's approach, as further suggested below.

Failure of condition (3) is, of course, compatible with $Pr(T \mid T \vdash E) > Pr(T)$: it could be true that finding out merely that T implies *something* about the relevant

phenomenon (i.e., that it either implies E or implies $\sim E$) would *decrease* one's confidence in T, while finding out that T implies *a truth* (i.e., E) about the relevant phenomenon would *increase* one's confidence in T. For example, one can imagine an investigator's having somehow hit upon the idea that Einstein's equations *might* have precise implications pertaining to the apparent orbit of Mercury, *without* having actually *gone through any calculations* to determine what the precise implications might be; and the investigator might conceivably, pessimistically, think it unlikely that any such precise consequences of the equations would match the previous (in this hypothetical example) precise observations of the orbit. In this case, nevertheless, a match would, for this investigator, provide striking confirmation.

More formally, note that conditions (1) and (4) above imply that $Pr(T / T \vdash \sim E) = 0$. Thus, $Pr(T / T \vdash E) > Pr(T)$ if and only if

(*) $Pr(T / T \vdash E) > Pr[\sim (T \vdash E) \ \& \ \sim [T \vdash \sim E) / \sim (T \vdash E)]$
$\times Pr[T / \sim (T \vdash E) \ \& \ \sim (T \vdash \sim E)]$.

And it is clear that this relation may be satisfied even if Jeffrey's condition (3), which, given his others, is equivalent to

(3') $Pr(T / T \vdash E \ v \ T \vdash \sim E) > Pr[T / \sim (T \vdash E) \ \& \ \sim (T \vdash \sim E)]$,

is not. Clearly (when Jeffrey's other conditions are met), the left-hand side of (*) is greater than the left-hand side of (3') (see the derivation of this above), and the right-hand side of (*) is less than the right-hand side of (3'), thus making (*) "easier to satisfy" than (3'). And, as suggested in the last paragraph with an example, it seems that there are genuine cases of confirmation involving old evidence in which (*) holds but (3'), i.e., (3), does not.

But how to "justify" (*)? Of course, we should not hope for a universal justification of (*), applicable in all cases of theories T and evidence statements E. For example, as Jeffrey points out,

> a purported theory of acupuncture that implies the true value of the gravitational red shift would be undermined thereby: [for example,] its implying *that* is likely testimony to its implying everything, i.e., to its inconsistency. (147)

Also, note that we are virtually "back to square one." Condition (*) is *equivalent* to $Pr(T / T \vdash E) > Pr(T)$, given Jeffrey's conditions (1) and (2) and Garber's (K^*), all of which are quite plausible given the specification of our problem, given logical nonomniscience, and given the intended interpretation of "$\vdash$"!

There are good reasons to think that there can be no *single* justification of (*); there *are* no "more primitive" assumptions that will justify (*) and be satisfied in all and only those situations in which $T \vdash E$ should be taken as confirming T. As Jeffrey's acupuncture/red shift example shows, whether or not there is a rational increase in confidence in T as a result of discovering that $T \vdash E$ will depend on what T and E *are about*, on the relationships between what T is about and what

E is about, and on our background beliefs. And of course these relevant items differ quite a lot from case to case; we should not expect them to be amenable to one single, systematic, formal treatment in the form of "more primitive" assumptions.

The case is parallel, I think, to the Bayesian explication of "E confirms T." The explication is "$Pr(T / E) > Pr(T)$", thus taking into account one's background beliefs, codified in Pr. Any justification of a statement of the form "$Pr(T / E) > Pr(T)$" will have to work from more or less particular information about the relation between T, E, and one's background beliefs, such information as "$T \vdash E$," or "E is a positive instance of T where T is subjectively probabilistically independent of the relevant object's satisfying the antecedent of T," or "E bootstrap confirms T relative to an appropriate theory," and so on. Such information as this pertains to particular cases, and no particular such piece of information will apply generally.

Similarly, it seems to me, the Bayesian should simply *explicate* "$T \vdash E$ confirms T (relative to Pr)" as "$Pr(T / T \vdash E) > Pr(T)$," without expecting there to be any single formal kind of *justification* of the latter for exactly the cases in which $T \vdash E$ *should* be taken as confirming T. Jeffrey's analysis will shed light in many cases; but in other cases, there may be other formal conditions, incompatible with his condition (3), that will imply $Pr(T / T \vdash E) > Pr(T)$. And it is conceivable that in other cases, it will be a formally "intractable" feature of Pr that the inequality holds; i.e., there is no simple, more primitive, relation holding between various items in virtue of which the inequality holds.

This latter kind of possibility, is, incidentally, closely connected with part of the Bayesian rationale for appeal to subjective probability distributions in the first place. As Charles Chihara has put it,

> To take account of heterogeneous information and evidence obtained from a variety of sources, all of different degrees of reliability and relevance, as well as of intuitive hunches and even vague memories, the Bayesian theory provides us with a subjective "prior probability distribution," which functions as a sort of systematic summary of such items. (1981, 433)

Of course, the prior probability distribution does not literally *summarize* the relevant *items* (hence the phrase "sort of" in the quotation), but rather summarizes the *effects* that the agent's absorption of such items has on the evidential situation at hand—the point being, in part, that there may be no simple formal relation between the relevant items themselves, the expression of which would accomplish the same task. And it seems that in some cases, there will not even be any simpler *subjective probabilistic* relations between items that justify or explain why $Pr(T / E) > Pr(T)$, or why $Pr(T / T \vdash E) > Pr(T)$.

As in the case of step two of the basic Bayesian strategy, I think the best hope for step three is (at least for starters) in the analysis of concrete, particular, histori-

cal cases of confirmation, which come complete with particular sets of background beliefs, particular theories, and evidence statements, and what they are about.

Notes

1. For more on Bayesian confirmation theory, and the Dutch book and decision theoretical approaches alluded to, and for further references, see, for example, Savage (1972), Jeffrey (1983), Hesse (1974), Chihara (1981), and Eells (1982). For criticisms, see, for example, Kennedy and Chihara (1979), Kyburg (1983), and Glymour (1980).

2. Note that to complete a classification of *situations* in which an *E* can confirm a *T*, we would have to add the case of "*New New* Evidence": the case in which *E* is new relative to the theory *T* (i.e., learned after *T* is formulated), and in which we are presently at, or just subsequent to, the moment of the discovery of *E*. Of course, this situation does not present a problem for Bayesian confirmation theory of the kind under discussion. For, as required by the theory, if *E* actually confirms *T*, then *E* increases our confidence in *T* at the time of *E*'s discovery, the time specified in the description of this kind of situation. It is, of course, central to Bayesian confirmation theory that confirmation of a hypothesis or theory implies "rational increase in confidence" in the hypothesis or theory, given one's background beliefs (full and partial), which are supposed to be systematically codified in a subjective probability (degree of confidence) function. Current controversy concerning the idea that confirmation coincides with confidence increase involves, I think, mainly (1) Glymour's problem of old evidence and (2) the plausibility of Glymour's (1980) "bootstrap" conception of evidence, which seems to be inconsistent with the idea of confirmation's implying increase in confidence. As to (1), that problem, and the possibility of a Bayesian resolution to it, is the focus of this paper. As to (2), see Horwich (1983) for discussion.

3. Martin Barrett made essentially these points in my seminar on confirmation theory.

4. Garber doesn't require what he calls condition (*), that $Pr(A \vdash B) = 1$ if $A \supset B$ is a tautology of L^* and must thus itself be assigned subjective probability 1. This leaves open the possibility that *even though locally Bayesian agents assign probability 1 to all tautologies*, they may do so, even in the infinity of complex cases, on grounds other than a perception (intuitive or otherwise) of truth-functional logical structure. But how might this possibility "realistically" be realized (even setting aside the implausibility of assigning probability 1 to all tautologies of L^*? Surely we cannot suppose that, for example, for very many cases of tautological conditionals, their probability 1 status is secured by the agent's being told by a source believed to be totally reliable that they are true. Garber acknowledges "a kind of *informal* contradiction in requiring that S be certain of $A \supset B$ when A truth-functionally entails B in [L^*], while at the same time allowing him to be uncertain of $A \vdash B$" (p. 118), while at the same time insisting (correctly) that this is no formal contradiction. Despite the intuitive implausibility of failure of (*) *given that* all tautologies must have subjective probability 1, Garber seems, in the end, neutral (or noncommittal), with respect to the condition (118).

5. For details on the Dutch book argument, originally due to de Finetti (1964), see, for example, Shimony (1955), Kemeny (1955), or Skyrms (1984, chapter 2).

6. Indeed, Hacking has emphasized (1967, 312–13), plausibly I think, the appropriateness of investigating certain logical and mathematical issues by empirical methods. Though not cited by Hacking, one plausible example of this is the question of which of a number of blackjack ("21") strategies yields a player the highest expectation against the house, which plays a fixed strategy. Though mathematically intractable, Monte Carlo techniques using computer simulation have been successfully employed. The point is that questions pertaining to how one can "in principle" come to know such-and-such are irrelevant to rationality and confirmation, where what is relevant is what one in fact knows, and the reliability and efficiency of one's actual methods of coming to know such-and-such, given one's background belief and knowledge.

7. In Jeffrey's version, *E* is contrasted not just with ~*E* but is, more generally, a member of a

set of mutually exclusive propositions about some phenomenon, such as the tides. The version just given, however, would seem to be an instance of this, since if *E* is about some phenomenon, then ~*E* would seem to be about the same phenomenon. Also, Jeffrey uses "*H*" rather than "*T*."

References

Carnap, Rudolf. 1947. *Meaning and Necessity*. Chicago: University of Chicago Press.

Chihara, Charles. 1981. "Quine and the Confirmational Paradoxes." In *Midwest Studies in Philosophy*. Vol. 6, *The Foundations of Analytic Philosophy*, eds. Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein. Minneapolis: University of Minnesota Press.

de Finetti, Bruno. 1964. "Foresight: Its Logical Laws, Its Subjective Sources." Trans. Henry E. Kyburg, Jr., from the 1937 French. In *Studies in Subjective Probability*, eds. Henry E. Kyburg, Jr., and Howard E. Smokler. New York: John Wiley & Sons, Inc.

Eells, Ellery. 1982. *Rational Decision and Causality*. New York: Cambridge University Press.

Garber, Daniel. 1983. "Old Evidence and Logical Omniscience in Bayesian Confirmation Theory." In *Minnesota Studies in the Philosophy of Science*, Vol. 10, *Testing Scientific Theories*, ed. John Earman. Minneapolis: University of Minnesota Press.

Glymour, Clark. 1980. *Theory and Evidence*, Princeton: Princeton University Press.

Hacking, Ian. 1967. Slightly More Realistic Personal Probability, *Philosophy of Science* 34:311–25.

Hesse, Mary. 1974. *The Structure of Scientific Inference*. Berkeley and Los Angeles: University of California Press.

Horwich, Paul. 1982. *Probability and Evidence*. New York: Cambridge University Press.

——. 1983. "Explanations of Irrelevance." In *Minnesota Studies in the Philosophy of Science*, Vol. 10, *Testing Scientific Theories*, ed. John Earman. Minneapolis: University of Minnesota Press.

Jeffrey, Richard C. 1983a. *The Logic of Decision*, 2d ed. Chicago: University of Chicago Press.

——. 1983b. "Bayesianism with a Human Face." In *Minnesota Studies in the Philosophy of Science*, Vol. 10, *Testing Scientific Theories*, ed. John Earman. Minneapolis: University of Minnesota Press.

Kemeny, John. 1955. Fair Bets and Inductive Probabilities. *Journal of Symbolic Logic* 20:263–73.

Kennedy, Ralph, and Chihara, Charles. 1979. The Dutch Book Argument: Its Logical Flaws, Its Subjective Sources. *Philosophical Studies* 36:19–33.

Kyburg, Henry E., Jr.. 1978. Subjective Probability: Criticisms, Reflections, and Problems. *Journal of Philosophical Logic* 7:157–180. Reprinted, with modifications, in his *Epistemology and Inference*, Minneapolis: University of Minnesota Press, 1983.

Savage, Leonard J. 1972. *The Foundations of Statistics*, 2d ed. New York: Dover Publications.

Shimony, Abner. 1955. Coherence and the Axioms of Probability. *Journal of Symbolic Logic* 20:1–28.

Skyrms, Brian. 1983. "Three Ways to Give a Probability Assignment a Memory." In *Minnesota Studies in the Philosophy of Science* Vol. 10, *Testing Scientific Theories*, ed. John Earman. Minneapolis: University of Minnesota Press.

——. 1984. *Pragmatics and Empiricism*. New Haven: Yale University Press.