————— Daniel Garber —————

Old Evidence and Logical Omniscience in Bayesian Confirmation Theory

The Bayesian framework is intended, at least in part, as a formalization and systematization of the sorts of reasoning that we all carry on at an intuitive level. One of the most attractive features of the Bayesian approach is the apparent ease and elegance with which it can deal with typical strategies for the confirmation of hypotheses in science. Using the apparatus of the mathematical theory of probability, the Bayesian can show how the acquisition of evidence can result in increased confidence in hypotheses, in accord with our best intuitions. Despite the obvious attractiveness of the Bayesian account of confirmation, though, some philosophers of science have resisted its manifest charms and raised serious objections to the Bayesian framework. Most of the objections have centered on the unrealistic nature of the assumptions required to establish the appropriateness of modeling an individual's beliefs by way of a point-valued, additive function.[1] But one recent attack is of a different sort. In a recent book on confirmation theory, Clark Glymour has presented an argument intended to show that the Bayesian account of confirmation fails at what it was thought to do best.[2] Glymour claims that there is an important class of scientific arguments, cases in which we are dealing with the apparent confirmation of new hypotheses by old evidence, for which the Bayesian account of confirmation seems hopelessly inadequate. In this essay I shall examine this difficulty, what I call the problem of old evidence. I shall argue that the problem of old evidence is generated by the

requirement that the Bayesian agent be logically omniscient, a requirement usually thought to follow from coherence. I shall show how the requirement of logical omniscience can be relaxed in a way consistent with coherence, and show how this can lead us to a solution of the problem of old evidence.

Since, as I. J. Good has conclusively shown, there are more kinds of Bayesianism than there are Bayesians[3], it will be helpful to give a quick sketch of what I take the Bayesian framework to be before entering into the problem of old evidence. By the Bayesian framework I shall understand a certain way of thinking about (rational) belief and the (rational) evolution of belief. The basic concept for the Bayesian is that of a degree of belief. The degree of belief that a person S has in a sentence p is a numerical measure of S's confidence in the truth of p, and is manifested in the choices S makes among bets, actions, etc. Formally S's degrees of belief at some time t_o are represented by a function P_o defined over at least some of the sentences of S's language L.[4] What differentiates the Bayesian account of belief from idealized psychology is the imposition of rationality conditions on S's beliefs. These rationality conditions are of two sorts, synchronic and diachronic. The most widely agreed upon synchronic condition is coherence:

> (D1) A P-function is *coherent* iff there is no series of bets in accordance with P such that anyone taking those bets would lose in every possible state of the world.

Although there are those who would argue that coherence is both necessary and sufficient for S's beliefs to be rational at some given time, I shall assume only that coherence is necessary. One of the central results of Bayesian probability theory is the coherence theorem, which establishes that if P is coherent, then it is a (finitely additive) probability function on the appropriate group of objects (i.e., the sentences of S's language L).[5] In the discussions below, I shall assume that an individual's degrees of belief have at least that much structure. Although there is little agreement about rational belief change, one way of changing one's beliefs is generally accepted as rational by most Bayesians, conditionalization. One changes one's beliefs in accordance with conditionalization when, upon learning that q, one changes one's beliefs from P_o to P_1 as follows:

$$P_1(p) = P_o(p/q)$$

where conditional probability is defined as usual. There are some who take

conditionalization as the sine qua non of the Bayesian account of belief, but I shall regard it as one among a number of possible ways of changing rational belief, a sufficient but not necessary condition of diachronic rationality.[6] Despite this proviso, though, conditionalization will have a major role to play in the discussion of confirmation that follows.

There are two competing ways of thinking about what the Bayesian is supposed to be doing, what I call the thought police model and the learning machine model. On the thought police model, the Bayesian is thought of as looking over our shoulders and clubbing us into line when we violate certain principles of right reasoning. On this view, the axioms of the theory of probability (i.e., coherence) and, perhaps, the dynamical assumption that we should change our beliefs in accordance with conditionalization are the clubs that the Bayesian has available. On the learning machine model, on the other hand, the Bayesian is thought of as constructing an ideal learning machine, or at least describing the features that we might want to build into an ideal learning machine.[7] Unlike others, I do not see a great deal of difference between these two ways of thinking about the enterprise. The Bayesian thought policeman might be thought of as clubbing us into behaving like ideal learning machines, if we like. Or, alternatively, we can think of the ideal learning machine as an imaginary person who behaves in such a way that he never needs correction by the Bayesian thought police. The two models thus seem intertranslatable. Nevertheless, I prefer to think of the Bayesian enterprise on the learning machine model. Although this has no theoretical consequences, I think that it is a better heuristic model when one is thinking about the confirmation of hypotheses from a Bayesian point of view.

1. The Problem of Old Evidence

In the course of presenting his own ingenious account of the confirmation of scientific hypotheses by empirical evidence, Clark Glymour offers a number of reasons why he chooses not to follow the Bayesian path. Many of Glymour's arguments are worth serious consideration; but one of the problems Glymour raises seems particularly serious, and seems to go to the very foundations of the Bayesian framework. Glymour writes:

> Scientists commonly argue for their theories from evidence known long before the theories were introduced. Copernicus argued for his theory using observations made over the course of millenia . . . . Newton argued for universal gravitation using Kepler's second and third laws, established before the *Principia* was published. The

argument that Einstein gave in 1915 for his gravitational field equations was that they explained the anomalous advance of the perihelion of Mercury, established more that half a century earlier . . . . Old evidence can in fact confirm new theory, but according to Bayesian kinematics it cannot. For let us suppose that evidence *e* is known before theory *T* is introduced at time *t*. Because *e* is known at *t*, $\text{Prob}_t(e) = 1$. Further, because $\text{Prob}_t(e) = 1$, the likelihood of *e* given *T*, $\text{Prob}_t(e, T)$, is also 1. We then have:

$$\text{Prob}_t(T, e) = \frac{\text{Prob}_t(T) \times \text{Prob}_t(e, T)}{\text{Prob}_t(e)} = \text{Prob}_t(T)$$

The conditional probability of *T* on *e* is therefore the same as the prior probability of *T*: *e* cannot constitute evidence for *T* . . . . None of the Bayesian mechanisms apply, and if we are strictly limited to them, we have the absurdity that old evidence cannot confirm a new theory.[8]

Before trying to understand what is going wrong for the Bayesian and seeing what can be said in response, it will be worth our while to look more closely at the problem itself. There are at least two subtly different problems that Glymour might have in mind here. One of these problems concerns the scientist in the midst of his investigations who appears to be using a piece of old evidence to increase his confidence in a given theory. If we adopt a Bayesian model of scientific inquiry, then how could this happen? How could an appeal to old evidence ever raise the scientist's degree of belief in his theory? This is what I shall call, for the moment, the *historical* problem of old evidence.[9] But there is a second possible problem lurking in Glymour's complaints, what might be called the *ahistorical* problem of old evidence. When we are first learning a scientific theory, we are often in roughly the same epistemic position that the scientist was in when he first put the theory to test; the evidence that served to increase *his* degrees of belief will increase *ours* as well. But having absorbed the theory, our epistemic position changes. The present appeal to Kepler's laws does not any more actually *increase* our confidence in Newton's theory of universal gravitation, nor does the appeal to the perihelion of Mercury actually *increase* our confidence in general relativity any more. Once we have learned the theories, the evidence has done its work on our beliefs, so to speak. But nevertheless, even though the old evidence no longer serves to increase our degrees of belief in the theories in question, there is still a sense in which the evidence in question remains good evidence, and there is still a sense in which it is proper to say that the old evidence confirms the

theories in question. But if we are to adopt a Bayesian account of confirmation in accordance with which e confirms h iff P (h/e) > P (h), then how can we ever say that a piece of evidence, already known, confirms h?[10]

Now that we have a grasp on the problems, we can begin to look for some possible ways of responding. One obvious response might begin with the observation that if one had *not* known the evidence in question, then its discovery *would have* increased one's degrees of belief in the hypothesis in question. That is, in the circumstances in which e really does confirm h, if it *had been* the case that P (e) > 1, then it *would also have been* the case that P (h/e) > P (h). There are, to be sure, some details to be worked out here.[11] If P (e) were less than one, what precisely would it have been? What, for that matter, would all of the *rest* of the P-values have been? If such details could be worked out in a satisfactory way, this counterfactual gambit would offer us a reasonably natural solution to the *ahistorical* problem of old evidence. This solution amounts to replacing the identification of confirmation with positive statistical relevance with a more subtle notion of confirmation, in accordance with which e (ahistorically) confirms h iff, if e had been previously unknown, its discovery would have increased our degree of belief in h. That is, e (ahistorically) confirms h iff, if P(e) (and, of course, P(h)) were less than one, then P(h/e) would be greater than P(h). In what follows I shall assume that the ahistorical problem of old evidence can be settled by some variant or other of this counterfactual strategy.[12]

It should be evident, though, that however well the counterfactual strategy might work for the *ahistorical* problem of old evidence, it leaves the *historical* problem untouched. When dealing with Einstein and the perihelion of Mercury, we are not dealing with a *counterfactual* increase in Einstein's confidence in his theory: we are dealing with an *actual* increase in his degree of belief. Somehow or other, Einstein's consideration of a piece of old evidence served to increase his confidence in his field equations, not counterfactually, but actually. This is something that the counterfactual solution cannot deal with.

How, then, are we to deal with the historical problem of old evidence, the cases in which considerations involving old evidence seem actually to *raise* an investigator's confidence in one of his hypotheses? We can put our finger on exactly what is going wrong in the Bayesian account if we go back and examine exactly when a piece of old evidence does seem to confirm a new hypothesis. It is appropriate to begin with the observation that Glymour's conclusion is not *always* implausible. There are, indeed, some

circumstances in which an old e *cannot raise* the investigator's degree of belief in a new h. For example, suppose that S constructed h *specifically* to account for e, and knew, from the start, that it would. It should not *add* anything to the credibility of h that it accounts for the evidence that S knew all along it would account for. In this situation, there is not confirmation, at least not in the relevance sense of that term.[13] The evidential significance of the old evidence is, as it were, *built into* the initial probability that S assigns to the new hypothesis. Where the result is paradoxical is in the case in which h was concocted *without* having e in mind, and only later was it *discovered* that h bears the appropriate relations to e, i.e., that h (and perhaps some suitable auxiliaries) entails e, that e is a positive instance of h, or the like. Just *what* the relationship in question is a matter of some debate. But it seems clear that in the cases at hand, what *increases* S's confidence in h is not e itself, but the *discovery* of some generally logical or mathematical relationship between h and e. In what follows I shall often assume for simplicity that the relation in question is some kind of logical entailment. But although the details may be shaped by this assumption, the general lines of the discussion should remain unaffected.

With this in mind, it is now possible to identify just which part of the Bayesian framework is generating the problem. In the Bayesian framework, coherence is almost always taken to imply that the rational subject S, the constraints on whose degrees of belief the Bayesian is trying to describe, is *logically omniscient*. Since logical (and mathematical) truths are true in all possible states of the world, if P is to be coherent, then coherence must, it seems, preclude the possibility of S's accepting a bet against a logical truth. Consequently, coherence seems to require that S be certain of (in the sense of having degree of belief one in) all logical truths and logical entailments. Now for logically omniscient S it is absolutely correct to say that old evidence e does not increase his confidence in a new hypothesis h. Because of S's logical omniscience, S will see immediately, for every new hypothesis, whether or not it entails his previously known evidence (or, perhaps, bears the appropriate logical relations to it). No hypothesis ever enters S's serious consideration without his knowing explicitly just which of his past observations it entails. So every new hypothesis S takes into consideration is, in a clear sense, *based* on the previously known observations it entails: the initial probability assigned to every new hypothesis already takes into account the old evidence it entails. For no hypothesis h and evidence e can the logically omniscient S ever

discover, after the fact, that h entails e. And, as I have suggested above, in such a circumstance, it is perfectly intuitive to suppose that the previously known evidence does not confirm the new hypothesis in the sense of raising its probability. The historical problem of old evidence, then, seems to be a consequence of the fact that the Bayesian framework is a theory of reasoning for a logically omniscient being.

It has generally been recognized that the Bayesian framework does not seem to allow the Bayesian agent to be ignorant of logical truths, and thus does not allow a Bayesian account of logical or mathematical reasoning. Although this has been considered a weakness of the framework, it has usually been accepted as an idealization that we *must* make in order to build an adequate account of the acquisition of *empirical* knowledge. What the problem of old evidence shows is that this idealization will not do: without an account of how the Bayesian can come to learn *logical* truths, we cannot have a fully adequate theory of *empirical* learning either. So if we are to account for how old evidence can raise the investigator's degree of belief in new hypotheses, we must be able to account for how he can come to know certain logical relations between hypothesis and evidence that he did not know when he first formulated the new hypothesis.

The problem of old evidence is not of course the only reason for seeking an account of logical learning consistent with Bayesian principles. There is an even deeper concern here. With the assumption of logical omniscience, there is a philosophically disturbing asymmetry between logical and empirical knowledge in the Bayesian framework. Although it may be unfortunate that we lack omniscience with respect to empirical truths, the Bayesian account makes it *irrational* to be anything but logically omniscient. The Bayesian agent who is *not* logically omniscient is incoherent, and seems to violate the only necessary condition for synchronic rationality that Bayesians can agree on. This is an asymmetry that smacks of the dreaded analytic-synthetic distinction. But scruples about the metaphysical or epistemic status of that distinction aside, the asymmetry in the treatment of logical and empirical knowledge is, on the face of it, absurd. It should be no more *irrational* to fail to know the least prime number greater than one million than it is to fail to know the number of volumes in the Library of Congress.[14]

The project, then, is clear: if the Bayesian learning model is to be saved, then we must find a way to deal with the learning of logical truths within the Bayesian framework. If we do this correctly, it should give us both a way of

eliminating the asymmetry between logical and empirical knowledge, and a way of dealing with the problem of old evidence. This is the problem taken up in the following sections.

2. Two Models of Logical Learning

A solution to the problem of old evidence requires that the Bayesian be able to give an account of how the agent S can come to know logical truths that he did not previously know. In this section I shall present and discuss two possible Bayesian models of logical learning. Because of the immediate problem at hand, the models will be formulated in terms of a particular kind of logical truth, those of the form "p logically entails q," symbolized by "p ⊢ q," although much of what I say can be extended naturally to the more general case. In this section I shall not discuss the precise nature of the logical implications dealt with here (i.e., truth-functional entailment vs. first order quantificational entailment vs. higher order quantificational entailment, etc.), nor shall I discuss the nature of the underlying language. These clarifications and refinements will be introduced as needed in the succeeding sections. But even without these refinements, we can say some interesting things about the broad paths we might follow in providing a Bayesian account of logical learning.

The two models of logical learning that I would like to discuss are the *conditionalization model* and the *evolving probability model*. On the conditionalization model, when S learns that p ⊢ q, he should change his beliefs from P_o to P_1 as follows:

$$P_1(-) = P_o (-/p \vdash q)$$

On the evolving probability model, on the other hand, when S learns that p ⊢ q, he is required to change his beliefs in such a way that P(q/p) = 1, and to alter the rest of his beliefs in such a way that coherence is maintained, or at least in such a way that his beliefs are as coherent as they can be, given his imperfect knowledge of logical truth.[15]

Which, if either, of these models should the Bayesian adopt? The conditionalization model has obvious attractions, since it fits neatly into the most popular Bayesian account of belief change in general. But however attractive it might be on its face, the conditionalization model has one obvious difficulty. I pointed out earlier that coherence seems to require that all logical truths get probability one. Consequently we are left with an unattractive choice of alternatives. It *seems* as if we must either say that the conditionalization model fails to allow for any logical learning, since in the

case at hand P_1 must always equal P_o; or we must radically alter the notion of coherence, so that logical truths can get probability of less than one. Let us then set the conditionalization model aside for the moment and see if we can make do with evolving probability.

The evolving probability model does not have the obvious difficulties that the conditionalization model has. It does, however, require a major change in the way we think about coherence. If we adopt the evolving probability model, then we are implicitly removing coherence as a synchronic constraint on rational belief. The best that we can say is that an individual ought to regard coherence as an ultimate goal. That is, the evolving probabilist seems forced to the position according to which the synchronic condition for rationality is not coherence itself, but only that the rational individual try to become as coherent as he can. Although this is intuitively not unattractive, it does have at least one unattractive consequence. If it is not required that an individual be coherent at any given time, then it would seem that nothing very strong could be said about the general characteristics that a rational individual's beliefs would have to satisfy at any given time. All of the wonderful theorems of the mathematical theory of probability would not apply to the rational investigator, but would only apply at the limit, at the end of inquiry, when his beliefs became fully coherent. But although this is somewhat unattractive, we could probably learn to live with this consequence if the evolving probability model turned out to be otherwise adequate to the task.

Unfortunately, though, it does not. Even if we could accept the required weakening of the constraint of coherence, there are three other problems that should give us serious pause. For one, the evolving probability model as stated gives us very little guidance as to how we ought to change our beliefs upon discovering that h ⊢ e. If the *only required* changes in our beliefs upon learning that h ⊢ e are to alter P(e/h) to one and restore coherence, we can always find a way of changing our beliefs consistent with the evolving probability model that will raise, lower, or leave P(h) unchanged. Suppose at t_o, $P_o(E/h) < 1$. That is, $P_o(e \& h) < P_o(h)$. Suppose, then, that we learn that h ⊢ e, so that we alter P_o in such a way that $P_1(e/h) = 1$. That is, we now have $P_1(e \& h) = P_1(h)$. But it is clear that this result can be arrived at in any of three ways: we can lower $P_o(h)$, leave $P_o(h)$ unchanged and raise $P_o(e \& h)$, or we can raise both $P_o(h)$ and $P_o(h \& e)$ to the same level. Each of these ways of altering one's beliefs is consistent with the evolving probability model. Consequently the evolving probabil-

ity model can tell us nothing general about the effect that learning that $h \vdash e$ may have on the rest of one's beliefs. The effect it has is determined by the *way* in which one changes from $P(e/h) < 1$ to $P(e/h) = 1$, and the evolving probability model says nothing about this.[16]

There is a second, more philosophical difficulty connected with the evolving probability model. Although the evolving probability model gives the Bayesian a way of dealing with logical learning, something of the original asymmetry between logical and empirical learning still remains. Upon learning an empirical truth, one (presumably) changes one's beliefs through conditionalization, whereas upon learning a logical truth, one changes one's beliefs through evolving probabilities. This continuing asymmetry should make us feel somewhat uncomfortable. The asymmetry could be eliminated, of course. We could declare that the evolving probability scheme is the way to change one's beliefs whether we learn empirical truths or logical ones, and give up conditionalization altogether, even for empirical learning. One might say, for instance, when S learns that e, he should simply change his beliefs in such a way that $P_1(e) = 1$, along with whatever other changes are necessary to restore coherence. But this is not very satisfactory. It would subject empirical learning to the same kind of indeterminacy that logical learning has, on the evolving probability model, and prevent our saying anything interesting of a general nature about empirical learning as well.

These two problems are serious. But there is a third problem even more serious than the previous two. Although the evolving probability model may give us a way of thinking about logical learning within the Bayesian framework, it is utterly incapable of dealing with the problem of old evidence. I argued that in the circumstances that give rise to the problem it is learning that our new hypothesis entails some piece of old evidence (or is related to it in some appropriate logical or mathematical way) that raises our degree of belief in h. But if we adopt the evolving probability model, learning that $h \vdash e$ in those circumstances will not change our beliefs at all! The evolving probability model tells us that when we learn that $h \vdash e$, we should alter our beliefs in such a way that $P(e) = 1$. But in the cases at hand, where e is old evidence, and thus $P(e) = 1$, $P(e/h)$ *already* equals 1 (as does "$P(h \supset e)$"). So, in the cases at hand, the evolving probability model will counsel *no change at all* in our degrees of belief. Thus learning that $h \vdash e$ can have no effect at all on our degree of belief in h, if e is previously known.

I have offered three reasons for being somewhat cautious about adopting the evolving probability model of logical learning. These arguments suggest that we turn to the conditionalization model. We must of course subject the conditionalization model to the same tests to which we subjected the evolving probability model. We must examine how well it determines the new probability function, how well it deals with the problem of asymmetry, and most important of all, how well it deals with the problem of old evidence. But first we must deal with the most basic and evident difficulty confronting the conditionalization model: can any sense be made of a probability function in which $P(h \vdash e)$ is anything but 0 or 1? Will allowing probability functions in which $0 < P(h \vdash e) < 1$ force us into incoherence in both the technical and nontechnical senses of that word?

3. Coherence and Logical Truth: An Informal Account

As I noted earlier, the standard definition of coherence, (D1), seems to require that all logical truths get probability 1. For surely, if h entails e, it entails e in every possible state of the world, it would seem. And if we were to assign probability less than one to a sentence like "$h \vdash e$," then we would be allowed to bet that "$h \vdash e$" is false, a bet that we would lose, no matter what state of the world we were in. Thus if we require P to be coherent, logical omniscience seems inescapable, and the conditionalization model of logical learning seems untenable.

One way out of this problem might be to eliminate coherence as a necessary condition of rational belief. But this is not very satisfying. If we were to eliminate coherence, we would have no synchronic conditions on rational belief at all; the Bayesian framework would reduce to an idealized psychology. It might help to reintroduce coherence as an ultimate goal of inquiry, as the evolving probabilist implicitly does. But, as I suggested in the course of our examination of the evolving probability model, this is not very attractive. This ploy has the unfortunate consequence of allowing us to say nothing of interest about the characteristics that a rational person's beliefs would have to exhibit at any given time. Explicitly relativizing coherence to an individual's state of knowledge with respect to logical truth might seem attractive, and has actually been proposed.[17] But this will give us little of the mathematical structure that we want. Moreover, it has the extra problem of introducing the philosophically problematic notion of knowledge explicitly into the Bayesian framework.

But all is not lost. Although it does not seem advisable to eliminate or

weaken coherence, perhaps a more careful examination of the coherence condition itself may give us a way of weakening the requirement of logical omniscience. The definition of coherence is obviously relativized to another notion, that of a *possible state of the world*. How we understand that notion should have important consequences for the constraints that the coherence condition imposes on an individual's beliefs. And how we understand the notion of a possible state of the world, it turns out, depends on what we think the Bayesian learning model is supposed to do.

One popular conception of the Bayesian enterprise is what I shall call *global Bayesianism*.[18] On this conception, what the Bayesian is trying to do is build a global learning machine, a scientific robot that will digest all of the information we feed it and churn out appropriate degrees of belief. On this model, the choice of a language over which to define one's probability function is as important as the constraints that one imposes on that function and its evolution. On this model, the appropriate language to building into the scientific robot is the *ideal language of science*, a maximally fine-grained language L, capable of expressing all possible hypotheses, all possible evidence, capable of doing logic, mathematics, etc. In short, L must be capable, in principle, of saying anything we might ever find a need to say in science.

Now, given this global framework, there is a natural candidate for what the possible states of the world are: they are the maximal consistent sets of sentences in L. But if *these* are what we take to be the possible states of the world, then logical omniscience of the very strongest sort seems to be demanded, and the conditionalization model of logical learning goes out the window. For if the possible states of the world are the maximal consistent sets of sentences in the most fine-grained, ideal language of science, then they are, in essence, the *logically* possible states of the world. And if I am coherent with respect to these states, i.e., if I am not allowed to enter into bets that I would lose in every such logically possible state of the world, then I must have degree of belief one in all logical truths.

But there are reasons for thinking twice before accepting this conclusion. Although global Bayesianism is a position often advanced, it is a very implausible one to take. For one thing, it does not seem reasonable to suppose that there is any one language that we can be sure can do everything, an immutable language of science of the sort that the Vienna Positivists sought to construct. Without such a language, the scientific robot model of Bayesianism is untenable, as is the idea that there is some

one unique set of logically possible states of the world to which we are obligated to appeal in establishing coherence. But even if it were possible to find a cannonical and complete language for science, it would not be of much use. One of the goals of the Bayesian enterprise is to reconstruct scientific practice, even if in an idealized or rationalized form. Typically when scientists or decision makers apply Bayesian methods to the clarification of inferential problems, they do so in a much more restricted scope than global Bayesianism suggests, dealing only with the sentences and degrees of belief that they are actually concerned with, those that pertain to the problem at hand.

This suggests a different way of thinking about the Bayesian learning model, what one might call *local Bayesianism*.[19] On this model, the Bayesian does not see himself as trying to build a global learning machine, or a scientific robot. Rather, the goal is to build a hand-held calculator, as it were, a tool to help the scientist or decision maker with particular inferential problems. On this view, the Bayesian framework provides a general formal structure in which one can set up a wide variety of different inferential problems. In order to apply it in some particular situation, we enter in only what we need to deal with in the context of the problem at hand, i.e., the particular sentences with which we are concerned, and the beliefs (prior probabilities) we have with respect to those sentences.

So, for example, if we are interested in a particular group of hypotheses h_i, and what we could learn about them if we were to acquire some evidence e_i, then our *problem relative language* L' would naturally enough be just the truth-functional closure of the h_i and the e_i. Our probability functions would then, for the duration of our interest in this problem, be defined *not* over the maximally specific language of science L, but over the considerably more modest problem-relative language L'.

In working only with the problem relative L', we are in effect treating each of the h_i and e_i as *atomic sentences*. This is not to say that h_i and e_i *don't* have any structure. Of course they do. It is by virtue of that structure, which determines their *meanings*, that we can tell in a given observational circumstance whether or not a given e_i is true, and it is by virtue of that structure that we know what it is that our degrees of belief are degrees of belief about! But none of this extra content is entered into our Bayesian hand-held calculator. Whatever structure h_i and e_i might have in some language *richer* than L' is submerged, so to speak, and the h_i and e_i treated as unanalyzed wholes from the point of view of the problem at hand. This

extra structure is not *lost,* of course. But it only enters in *extrasystematically,* so to speak, when, for example, we are assigning priors, or when we are deciding whether or not a particular observational sentence is true in a particular circumstance.

This seems to open the door to a Bayesian treatment of logical truth. In some investigations we are interested only in sentences like "h_i" and "e_i." But in others, like those in which the problem of old evidence comes up, we are interested in other sentences, like "$h_i \vdash e_i$." Sentences like "$h_i \vdash e_i$" certainly have structure. Depending on the context of investigation, "$\vdash$" may be understood as truth-functional implication, or implication in L, 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.[20] But whatever *extrasystematic* content we give sentences like "$h_i \vdash e_i$," in the context of our problem-relative investigation we can throw such sentences into our problem-relative language as *atomic sentences,* unanalyzed and unanalyzable wholes, and submerge whatever content and structure they might have, exactly as we did for the h_i and e_i.

Suppose now that we are in a circumstance in which logical relations between sentences are of concern to us. Say we are interested in some implicative relations between hypotheses and evidence, sentences of the form "$h_i \vdash e_i$." The problem-relative language will be the truth-functional closure of all the h_i, e_i, and sentences of the form "$h_i \vdash e_i$," where *each* of these sentences, *including* those of the form "$h_i \vdash e_i$" is treated as an *atomic sentence* of the problem relative language. Now the crucial question is this: what constraints does coherence impose on probability functions defined over this language? In particular, does coherence require that all sentences of the form "$h_i \vdash e_i$" get 0 or 1? If not, then we are out of the woods and on our way to an account of logical learning through conditionalization.

As I argued, in order to decide what follows from coherence, we must determine what is to count as a possible state of the world. Now in giving up global Bayesianism and any attempt to formulate a maximally fine-grained language of science, we give up in effect the idea that there is some one set of logically possible states of the world that stands behind every inferential problem. But how then *are* we to understand states of the world? The obvious suggestion is this. In the context of a particular investigation, we are interested in some list of atomic sentences and their truth-functional

compounds: hypotheses, possible evidence, and statements of the logical relations between the two. Insofar as we are uncertain of the truth or falsity of any of these atomic sentences, we should regard each of them as true in some states of the world, and false in others, at least in the context of our investigation. And since, in the context of investigation, we are interested in no other sentences, our problem relative states of the world are easily specified: *they are determined by every possible distribution of truth values to the atomic sentences of the local language L'.* This amounts to replacing the *logically* possible worlds of the global language with more modest *epistemically* possible worlds, specified in accordance with our immediate interests.

Now if the possible states of the world are those determined by all possible assignments of truth values to the atomic sentences of the local language L', then coherence imposes one obvious constraint on the scientist's degrees of belief: if sentence T in L' is true on all possible assignments of truth values to the atomic sentences of L', then $P(T) = 1$. That is, if T is a *tautology* of L' then $P(T) = 1$. Coherence understood in this way, however, relativized to the problem-relative states of the world, does *not* impose *any* constraints on the *atomic sentences* of L'. Since for any atomic sentence of L' there are states of the world in which it is false, we can clearly assign *whatever* degree of belief we like to *any* of the atomic sentences without violating coherence, i.e., without being caught in the position of accepting bets that we would lose in every (problem-relative) state of the world. *And this holds even if one of those atomic sentences is extrasystematically interpreted as "h logically entails e."*

This seems to get us exactly what we want. It seems to allow us to talk about uncertainty with respect to at least *some* logical truths, and in fact, it allows us to do this without even violating coherence! This is an interesting and slightly paradoxical result. In order to see better what is going on, and make sure that there is no contradiction lurking beneath the surface of the exposition, I shall try to set the result out more formally.

4. Coherence and Logical Truth: a Formal Account

In the previous section we dealt informally with relatively modest local languages, a few hypotheses, a few evidential sentences, a few logical relations. But the coherence result I argued for can be shown formally to hold for much larger languages as well. Let us consider first the language L, the truth-functional closure of a countably infinite collection of atomic

sentences, {a_i}. Let us build the larger language L* by adding to L some new atomic sentences, those of the form "A ⊦ B," where A and B are in L, and again, closing under truth-functional operations. L* is a truth-functional language that allows us to talk about truth-functional combinations of an infinite set of atomic sentences {a_i}, and relations of implication between any truth-functional combination of these sentences.[21] So it is clearly adequate to handle any of the problem situations that we had been discussing earlier.

Now, L* is just a truth-functional language generated by a countably infinite number of atomic sentences, i.e., those of the form "a_i" or "A ⊦ B." So, if the possible states of the world are identified with possible assignments of truth values to the atomic sentences of L*, on analogy with what I argued above with respect to the more modest local languages, then imposing coherence will fix no degrees of belief with respect to the atomic sentences of L*. There will be coherent P-functions that will allow us to assign *whatever* values we like in [0, 1] to the atomic sentences of the form "A ⊦ B," however these may be interpreted extrasystematically. The only specific values fixed by the requirement of coherence will be those of the tautologies and truth-functional contradictions in L*, i.e., the tautological and contradictory *combinations* of atomic sentences of L*.

This almost trivial result follows directly from the fact that, from the point of view of the probability function, sentences like "A ⊦ B" are uninterpreted and treated on a par with the a_i, treated like structureless wholes. But, interestingly enough, a similar result can be obtained without such a strong assumption. That is, we can introduce a certain amount of structure on the atomic sentences of the form "A ⊦ B" without restricting our freedom to assign them probabilities strictly between 0 and 1.

In introducing the atomic sentences of the form "A ⊦ B" into our local problem-relative languages, I emphasized that "A ⊦ B" could be interpreted extrasystematically in a variety of different ways, as "A truth-functionally entails B," that is, as "'A ⊃ B' is valid in L," as "A entails B in some richer language" (e.g., in the maximally fine-grained ideal language of science), or as some logical or mathematical relation other than implication, e.g., as "B is a positive instance of A," or as "B bootstrap confirms A with respect to some appropriate theory," in the sense in which Glymour understands this relation. For the purpose of adding some additional structure, though, let us assume that we are dealing with some variety of implication or other. Now if "A ⊦ B" is to be read as "A implies

B," we may want to require that our Bayesian investigator S recognize that atomic sentences of the form "A ⊦ B" have *some* special properties, however implication is understood. Although we do not want to demand that S recognize all true and false ones, it does seem reasonable to demand that S recognize that *modus ponens* is applicable to these particular atomic sentences of L*. That is, we might require that if "A ⊦ B" is to be properly read as "A implies B," then at very least, if S knows that A, and S knows that A ⊦ B, he must also know that B as well. Put probabilistically, this amounts to adopting the following constraint over reasonable degree of belief functions on L*:

(K) P(B/A & A ⊦ B) = 1, when defined.

But since, Renyi and Popper aside, this conditional probability is undefined when P(A & A ⊦ B) = 0, we might replace (K) with the following slightly stronger condition:

(K*) P(A & B & A ⊦ B) = P(A & A ⊦ B).

(K*) clearly reduces to (K) when the conditional probability in (K) is defined.

(K*) is a stronger condition than it may appear on the surface. If, in addition to coherence, we impose (K*) on all "reasonable" probability functions defined on L*, then we get a number of interesting and desirable properties, as outlined in the following theorem:

(T1) If P is a probability function on L* and P satisfies (K*), then:
   (i) If P(A ⊦ B) = 1, then P(A ⊃ B) = 1 and P(B/A) = 1, when defined.
   (ii) P(~A/~B & A ⊦ B) = 1 when defined.
   (iii) If A and B are truth-functionally inconsistent in L, then P(A & A ⊦ B) = 0.
   (iv) P(B/(A ⊦ B) & (~A ⊦ B)) = 1, when defined.
   (v) If P(A & A ⊦ B) = 1, then P(A ⊦ ~B) = 0.
   (vi) If B and C are truth-functionally inconsistent in L, then P(A/(A ⊦ B) & (A ⊦ C)) = 0, when defined.
   (vii) As P(B) →0, P(A/A ⊦ B) →0 and P(A ⊦ B/A) →0.
   (viii) If A and ~B are both tautologies in L, then P(A ⊦ B) = 0.
   Proof: All of the arguments are trivial and left to the reader.

These properties are attractive, and seem appropriate when " ⊦ " is interpreted as a variety of implication.[22] Imposing (K*) guarantees that

when we learn that $A \vdash B$, our degrees of belief in "$A \supset B$" and our conditional degrees of belief in B given A will behave appropriately, by clause (i). It gives us a probabilistic version of *modus tollens* (clauses (ii) and (vii)). It also guarantees that S will be certain of the truth of anything that follows both from A and from ~A (clause (iv)), and that S will be certain of the falsity of anything that has truth-functionally inconsistent consequences (clause (vi)).

Now (K*) seems to be an appropriate constraint to impose on any probability function defined over L^*, if "$\vdash$" is to be interpreted as a variety of logical implication. Although it does not *guarantee* that we are dealing with a variety of implication,[23] it is certainly reasonable to require that any variety of implication should satisfy (K*). But now matters are not so trivial. Might adding (K*) as an extra constraint take away *all* of the freedom we had in assigning probabilities to sentences of the form "$A \vdash B$" in L^*? The coherence condition imposes no constraints on assigning probabilities to the atomic sentences of L^*, I have argued. Most importantly, it does not force us to logical omniscience, to the position in which all sentences of the form "$A \vdash B$" are forced to take on probabilities of 0 or 1. But might coherence in conjunction with (K*)? The surprising answer is that with one small exception (already given in (T1) (viii)), no! This result is set out in the following theorem:

(T2) There exists at least one probability function P on L^* such that P satisfies (K*) and such that every atomic sentence in L^* of the form "$A \vdash B$" where not both A and ~B are tautologies gets a value strictly between 0 and 1.

Proof: Consider L and L^* as above. Let P be any strictly positive probability on L. That is, for A in L, $P(A) = 0$ iff A is truth-functionally inconsistent in L. Then extend P to L^* as follows:

(i) Suppose that A in L is not a tautology. Then let C be any sentence in L which is nontautologous, noncontradictory, and inconsistent with A. If A is not truth-functionally inconsistent in L, then ~A will do; otherwise let C be any atomic sentence a_i in L. Then, for any B in L, let $P(A \vdash B) = P(C)$; and for any D in L^*, let $P([A \vdash B] \& D) = P(C \& D)$; $P([A \vdash B]vD) = P(CvD)$; etc.

(ii) Suppose that A in L is a tautology and B is not. Then let

$P(A \vdash B) = P(B)$; $P([A \vdash B] \& D) = P(B \& D)$; $P([A \vdash B]vD) = P(BvD)$; etc.

(iii) Suppose that A and B in L are both tautologies. Then let $P(A \vdash B) = P(a_i)$, where "a_i" is an arbitrary atomic sentence in L; $P([A \vdash B] \& D) = P(a_i \& D)$; $P([A \vdash B]vD) = P(a_ivD)$; etc.

P so extended is clearly a probability on L^*. Further, it can easily be shown that P so extended satisfies (K*). And finally, since P on L is strictly positive, $P(A \vdash B)$ will never have a value of either 0 or 1, except when both A and ~B are tautologies, in which case it will get a value of 0 by clause (ii).[24]

So it turns out that even if we add more structure, as we do when (K*) is introduced, we are not forced to logical omniscience. Even with (K*) *and* coherence, we are permitted to be uncertain of logical implications.[25]

These technical conclusions call for some reflection. How can I say that I have gotten rid of logical omniscience if S is *still* required to know all tautologies of L^*? And if S is required to know all tautologies of L^*, mustn't the freedom he is given with respect to the sentences of the form "$A \vdash B$" inevitably lead to contradiction? As regards logical omniscience, that has been eliminated. Coherence still requires that we have some logical knowledge. But knowing the tautologies of L^* is a far cry from logical omniscience, since there are many logical truths that are not tautologies of L^*. The threat of internal contradiction is more subtle, though. Formally speaking, there is no contradiction. The key to seeing this lies in understanding the distinction between those logical truths that S is required to know and those that he is not. Let A, B, $A \vdash B$ be sentences in our local problem relative language L^*, where A and B are truth-functional combinations of atomic sentences of L, and "$A \vdash B$" is an atomic sentence of L^* interpreted (extrasystematically) as "A entails B." For the purposes of discussion, it does not matter whether the turnstile is interpreted as truth-functional entailment in L, or something weaker. Now suppose that, as a matter of fact, A *does* truth-functionally entail B. What precisely does coherence require? It clearly requires that $P(B/A) = 1$ and $P(A \supset B) = 1$. That is, it requires that S be certain of B conditional on A, and certain of the tautology "$A \supset B$." But if my argument is correct, S is not required to be certain of the *atomic* sentence "$A \vdash B$," which can get a degree of belief strictly between 0 and 1. That is, in requiring that S be certain of "$A \supset B$,"

coherence requires that S be certain that a particular truth-functional combination of atomic sentences of L is *true*. But at the same time, in allowing uncertainty with respect to "A ⊢ B," coherence allows that S might be uncertain as to whether or not that truth-functional combination of atomic sentences is *valid*. And insofar as truth and validity are distinct, there is no formal contradiction in asserting that S may be certain that "A ⊃ B" is *true* without necessarily being certain that it is *valid*, i.e., without being certain that "A ⊢ B" is true.

But even if there is no formal contradiction, there does appear to be a kind of *informal* contradiction in requiring that S be certain of A ⊃ B when A truth-functionally entails B in L, while at the same time allowing him to be uncertain of A ⊢ B. But this informal contradiction can be resolved easily enough by adopting a new constraint on reasonable probability functions on L*:

(*)   If "A ⊃ B" is a tautology in L, then $P(A ⊢ B) = 1$.

This would require that S know not only the *truth* of all tautologies of L, but also their *validity*.[26] Although I see no particular reason to adopt (*), doing so would resolve the informal appearance of contradiction without doing much damage to the formalism or its applicability to scientific reasoning. For truth-functional implication is not the only variety of implication. In fact, when we are interested in the logical relations between hypotheses and evidentiary sentences, the kind of implicatory relations we are interested in will most likely be not truth-functional implication, but quantification-theoretic implication in some background language richer than L* in which the hypotheses and evidence receive their (extrasystematic) interpretation. So, in any realistic application of the formalism developed in this section, adding (*) as a constraint will fix only a small number of sentences of the form "A ⊢ B," and leave all of the rest unaffected. (*) will fix *all* such sentences *only* in the case in which "A ⊢ B" is interpreted rather narrowly as "A truth-functionally entails B in L," a case that is not likely to prove of much use in the analysis of scientific reasoning.

5. The Conditionalization Model and Old Evidence Redux

After this rather lengthy argument, it might help to review where we have been and gauge how much farther we have to go. Starting with the problem of old evidence, I argued that a fully adequate Bayesian account of scientific reasoning must include some account of the learning of logical

truths; in particular, it must allow for the fact that the logical and mathematical relations between hypotheses and evidence must be discovered, just as the empirical evidence itself must be. I then presented two Bayesian models of logical learning, the evolving probability model and the conditionalization model, argued that the evolving probability model has serious weaknesses, and suggested that we explore the conditionalization model. In the previous two sections I showed that the central problem with the conditionalization model, the widely held conviction that coherence requires that all logical truths get probability one, turns out not to be a problem at all. I showed that if we think of the Bayesian framework as problem-relative, a hand-held calculator rather than a scientific robot, then we can make perfectly good sense of assigning probabilities of less than one to the logical truths we are interested in, without even violating coherence! This conclusion enables us to return to the conditionalization model for learning logical truth, and discuss its adequacy, particularly in regard to the problem of old evidence.

On the conditionalization model, when S learns a logical truth, like "h ⊢ e," he should change his beliefs as follows:

$$P_1(-) = P_0(- / h ⊢ e)$$

The investigations of the previous sections have shown that this does not necessarily reduce to triviality, nor does it force us to give up the requirement of coherence. But is it an otherwise attractive way to think about the consequences of learning a logical truth? In discussing the evolving probability model, I noted three problems: (a) the evolving probability model does not uniquely determine a new probability function upon learning that h ⊢ e; (b) the evolving probability model maintains an asymmetry between logical and empirical learning; and (c) the evolving probability model offers no solution to the (historical) problem of old evidence. It is clear that the conditionalization model deals admirably with the first two of these problems. Since "P(−/ h ⊢ e)" is uniquely determined for all sentences in the language over which P is defined, the conditionalization model gives us a unique new value for all sentences of that language, upon learning that h ⊢ e. And there is obviously no asymmetry between logical and empirical learning: *both* can proceed by conditionalization. The third question, then, remains: how does the conditionalization model do with respect to the problem of old evidence? Unlike the previous two questions, the answer to this one is not obvious at all.

Earlier I argued that the (historical) problem of old evidence derives from the assumption of logical omniscience. For the logically omniscient S, old evidence can never be appealed to in order to *increase* his degree of belief because, as soon as h is proposed, S can immediately see all of the logical consequences of h, and thus his *initial* probability for h will be *based on* a complete knowledge of what it entails. If old evidence can be used to *raise* the probability of a new hypothesis, then, it must be by way of the *discovery* of previously unknown logical relations. In the cases that give rise to the problem of old evidence, we are thus dealing with circumstances in which hypotheses are confirmed not by the empirical evidence itself, but by the discovery of some logical relation between hypothesis and evidence, by the discovery that h ⊢ e. Now the evolving probability model of logical learning failed to deal with the problem of old evidence because on that model, when P(e) = 1, learning that h ⊢ e has no effects on S's degrees of belief. The evolving probability model thus breaks down in precisely the cases that are of interest to us here. But, one might ask, does the conditionalization model do any better? That is, is it possible on the conditionalization model for the discovery that h ⊢ e to change S's beliefs when e is previously known, for P(h/h ⊢ e) to be greater than P(h) when P(e) = 1? Unfortunately, (T2) will not help us very much here. (T2) does have the consequence that P(h ⊢ e) can be less than one when P(e) = 1, which is certainly *necessary* if P(h/h ⊢ e) is to be greater than P(h). But because of the assumption of a strictly positive probability on L in the proof of (T2), the probability function constructed there, in which (almost) all implications get probability strictly between 0 and 1 will be such that for any e, P(e) = 1 if and only if e is a tautology. Thus (T2) does not assure us that P(h ⊢ e) can be less than one when S is certain of a *non*tautologous e. This is not very convenient, since the old evidence we are interested in is not likely to be tautologous! Furthermore, although (T2) assures us that (K*) does not require extreme values on all logical implications, it does not assure us that that strong constraint ever allows for probability functions in which P(h/h ⊢ e) > P(h) for *any* e at all, tautologous or not. But luckily it is fairly easy to show that under appropriate circumstances, there is always a probability function on L* (in fact, an infinite number of them) that satisfies (K*) in which, for *any* noncontradictory e, and for *any* nonextreme values that might be assigned to P(h) and P(h ⊢ e), P(e) = 1 and P(h/h ⊢ e) > P(h). This is the content of the following theorem:

(T3) For L and L* constructed as above, for any atomic sentence of L*

of the form "A ⊢ B" where B is not a truth-functional contradiction in L and where A does not truth-functionally entail ~B in L and B does not truth-functionally entail A in L, and for any r, s in (0, 1), there exist an infinite number of probability functions on L* that satisfy (K*) and are such that P(B) = 1, P(A ⊢ B) = r, P(A) = s, and P(A/A ⊢ B) > P(A).

Proof: Consider all sentences s_i in L* of the following form (Carnapian state descriptions):

$(\pm)a_1\&$. . .$\&(\pm)a_n\&(\pm)[A ⊢ B]$

where $a_1, . . . , a_n$ are the atomic sentences of L that appear in every sentence of L equivalent to either A or B, if B is not a tautology, or those that appear in every equivalent of A, if it is, and "$(\pm)$" is replaced by either a negation sign or a blank. Define a function P over the s_i as follows. First of all, assign a P-value of 0 to any s_i that truth-functionally entails ~B in L*. Since B is not truth-functionally inconsistent, there will be some s_i that remain after the initial assignment. Divide the remaining s_i into the following classes:

Class 1: s_i that truth-functionally entail A&[A ⊢ B]
Class 2: s_i that truth-functionally entail A&~[A ⊢ B]
Class 3: s_i that truth-functionally entail ~A&[A ⊢ B]
Class 4: s_i that truth-functionally entail ~A& ~[A ⊢ B]

Each s_i truth-functionally entails either [A ⊢ B] or ~[A ⊢ B], but not both, and since each s_i fixes the truth values of all of the atomic sentences in A, each s_i truth-functionally entails either A or ~A, but not both. Thus every remaining s_i fits into one and only one of these classes. Also, since A does not truth-functionally entail ~B, there will be some s_i that remain which entail A. And while every remaining s_i truth-functionally entails B, since B does not truth-functionally entail A, there will be some that remain which entail ~A. Thus, it is obvious that none of these classes will be empty. Now, let $\delta = \min(r(1-s), s(1-r))$, and let ε be an arbitrarily chosen number in $(0, \delta]$. Because of the constraints imposed on r and s, $\delta > 0$ and $(0, \delta]$ is nontrivial. Given the constraints imposed on r, s, and ε it can be shown that each of the following quantities is in [0, 1]:

$rs + \varepsilon$, $s(1 - r) - \varepsilon$, $r(1 - s) - \varepsilon$, $(1 - r)(1 - s) + \varepsilon$

So, we can extend P to the remaining s_i, those that do not truth-functionally entail $\sim$B, as follows:

Class 1: Let P assign any values in [0, 1] to the s_i in class 1 that sum to $rs + \varepsilon$

Class 2: Let P assign any values in [0, 1] to the s_i in class 2 that sum to $s(1 - r) - \varepsilon$

Class 3: Let P assign any values in [0, 1] to the s_i in class 3 that sum to $r(1 - s) - \varepsilon$

Class 4: Let P assign any values in [0, 1] to the s_i in class 4 that sum to $(1 - r)(1 - s) + \varepsilon$

This completes the definition of P on the s_i. Since the values assigned sum to 1, P defines a unique probability function on the sublanguage of L* generated by the s_i. This can be further extended to the whole of L* by assigning a P-value of 0 to all atomic sentences of L* that do not appear in the s_i. P so defined clearly satisfies (K*), and is such that P(B) = 1. Also:

$$P(A \vdash B) = rs + \varepsilon + r(1 - s) - \varepsilon = r$$
$$P(A) = rs + \varepsilon + s(1 - r) - \varepsilon = s$$

Furthermore, $P(A \& [A \vdash B]) = rs + \varepsilon > rs$, so, $P(A \& [A \vdash B]) > P(A)P(A \vdash B)$ and thus $P(A/A \vdash B) > P(A)$. Since ε was arbitrarily chosen from $(0, \delta]$, there are an infinite number of probability functions on L* that have the required properties.[27]

To take a simple numerical example as an illustration of (T3), let us suppose that h and e are both atomic sentences of L, say a_1 and a_2, and let us suppose that we want to build a probability function on L* in which $P(a_1) = .4$, $P(a_2) = 1$, and $P(a_1 \vdash a_2) = .4$, and in which $P(a_1/a_1 \vdash a_2) > P(a_1)$. One such function can be constructed by assigning the following probabilities to the appropriate state descriptions, and extending the function to L* as in the proof of (T3):

$P(a_1 \& a_2 \& [a_1 \vdash a_2]) = .3$	$P(a_1 \& a_2 \& \sim[a_1 \vdash a_2]) = .1$
$P(\sim a_1 \& a_2 \& [a_1 \vdash a_2]) = .1$	$P(\sim a_1 \& a_2 \sim[a_1 \vdash a_2]) = .5$
$P(a_1 \& \sim a_2 \& [a_1 \vdash a_2]) = 0$	$P(a_1 \& \sim a_2 \& \sim[a_1 \vdash a_2]) = 0$
$P(\sim a_1 \& \sim a_2 \& [a_1 \vdash a_2]) = 0$	$P(\sim a_1 \& \sim a_2 \& \sim[a_1 \vdash a_2]) = 0$

(Using the notation of the proof of (T3), $r = s = .4$, and $\delta = .24$, allowing ε to be any number in $(0, .24]$. The ε chosen in the example is .14). The

extension of these probabilities on the state descriptions clearly satisfies (K*), and clearly assigns the specified values to $P(a_1)$, $P(a_2)$, and $P(a_1 \vdash a_2)$. Furthermore, one can easily calculate that $P(a_1/a_1 \vdash a_2) = .3/.4 = .75$, which is clearly greater than $P(a_1)$. Thus, on my construction, it is *not* trivially the case that $P(h/h \vdash e) = P(h)$ when $P(e) = 1$, and the discovery that $h \vdash e$ *can* raise S's confidence in h. That is to say, unlike the evolving probability model, the conditionalization model of logical learning does not break down over the case of the problem of old evidence, even when (K*) is assumed to hold.

With this last feature of the conditionalization model in place, we have completed our solution to the problem of old evidence. I have shown how old evidence e can contribute to the confirmation of a more recently proposed h through the discovery that $h \vdash e$, and I have shown how this can be done in a way consistent with Bayesian first principles. Or, perhaps more accurately, I have shown *one* way in which the Bayesian can explain how, on his view of things, old evidence can confirm new hypotheses. This takes the sting out of Glymour's critique. With a bit of ingenuity the Bayesian *can* accommodate the kinds of cases that Glymour finds so damaging. But work remains before one can make a final judgment on the particular proposal that I have advanced, the particular way in which I have proposed to deal with the problem of old evidence. In particular, one must examine with great care the cases that Glymour cites—the case of Copernican astronomy and the ancient evidence on which it rested, Newton's theory of gravitation and Kepler's laws, and Einstein's field equations and the perihelion of Mercury—along with other cases like them, in order to determine whether or not my analysis of the reasoning fits the cases at hand. We must show that the scientists in question were initially uncertain that $h \vdash e$ for the appropriate h and e, that their prior degrees of belief were such that $P(h/h \vdash e) > P(h)$,[28] and that it was, indeed, the discovery that $h \vdash e$ that was, as a matter of fact, instrumental in increasing their confidence in h. Such investigations go far beyond the scope of this paper. My *intuition* is that when we look carefully at such cases, the details will work out in favor of the account that I propose.[29] But this is just an intuition.

6. Postscript: Bayesianizing the Bootstrap

I should point out that Clark Glymour was fully aware of the general lines of the solution to the problem of old evidence offered here at the time

Theory and Evidence was published. I proposed it to him while his book was still in manuscript, and we discussed it at some length. In the published version, Glymour gives a crude and early version of this line of argument, along with some remarks on why he does not believe it saves the Bayesian position. Glymour says:

> Now, in a sense, I believe this solution to the old evidence/new theory problem to be the correct one; what matters is the discovery of a certain logical or structural connection between a piece of evidence and piece of theory . . . . [The] suggestion is at least correct in sensing that our judgement of the relevance of evidence to theory depends on the perception of a structural connection between the two, and that the degree of belief is, at best, epiphenomenal. In the determination of the bearing of evidence on theory, there seem to be mechanisms and strategems that have no apparent connection with degrees of belief, which are shared alike by people advocating different theories . . . . But if this is correct, what is really important and really interesting is what these structural features may be. The condition of positive relevance [i.e., q confirms p iff P(p/q) > P(p)], even if it were correct, would simply be the least interesting part of what makes evidence relevant to theory.[30]

As I understand it, Glymour's point is that what should be of interest to confirmation theory is not degrees of belief and their relations, but the precise nature of the structural or logical or mathematical relations between hypothesis and evidence by virtue of which the evidence confirms the hypothesis. Put in terms I used earlier, Glymour is arguing that what confirmation theory should interest itself in is the precise nature of the "⊢" necessary to make the above given formalism applicable to the analysis of scientific contexts, rather than in the fine details of how the discovery that h ⊢ e may, in some particular situation, raise (or lower) some scientist's degree of belief in h. Now the most difficult kind of criticism to answer is the one that says that a certain project is just not very interesting or important. I shall not attempt to defend the interest of my investigations; but I shall argue that they should be of some importance even to Glymour's own program by showing that the account of confirmation through the discovery of logical truth that I offered in the body of this paper can be used to fill in a large gap in Glymour's theory of confirmation.

The structural relation, which, Glymour argues, should be what is of interest to the confirmation theorist, is the main focus of *Theory and Evidence*. What he offers is a version of instance confirmation, but with an

important and novel twist. Unlike previous writers, Glymour allows the use of *auxiliary theories* in the arguments used to establish that a given piece of evidence is a positive instance of a given hypotheses. Glymour summarizes his account as follows:

> [N]eglecting anomalous cases, hypotheses are supported by positive instances, disconfirmed by negative ones; instances of a hypothesis in a theory, whether positive or negative, are obtained by "bootstrapping," that is, by using the hypotheses of that theory itself (or, conceivably, some other) to make computations from values obtained from experiment, observation, or independent theoretical considerations; the computations must be carried out in such a way as to admit the possibility that the resulting instance of the hypothesis tested will be negative. Hypotheses, on this account, are not generally tested or supported or confirmed absolutely, but only *relative to a theory*.[31]

Glymour's intuitive sketch could be filled out in a number of ways. But since the idea is clear enough, I shall pass over the details here. With Glymour's bootstrap analogy in mind, I shall say that e BS confirms h with respect to T when the structural relation in question holds, and will symbolize it by "[h ⊢ e]$_T$."

Glymour tells us a great deal about BS confirmation. But one thing that he doesn't say very much about is how we can compare different BS confirmations. The discovery that [h ⊢ e]$_T$ is supposed to confirm h; it is supposed to support h and give us some reason for believing h. But when does one BS confirmation support h better than another? This is a general question, one that could be asked in the context of any confirmation theory. But it has special importance for Glymour. A distinctive feature of Glymour's theory of confirmation, one that he takes great pains to emphasize, is the fact that BS confirmations are *explicitly* relativized to auxiliary theories or hypotheses. By itself, this feature is unobjectionable. But it leads to a bit of a problem when we realize that for virtually any hypothesis h and any evidence e, there will be *some* auxiliary T such that [h ⊢ e]$_T$. I shall not give a general argument for this, but the grounds for such a claim are evident enough when we examine how Glymour's BS method applies to systems of equations relating observational and theoretical quantities.[32] Let the hypothesis h be the following equation:

$$X(q_1, \ldots, q_j) = 0$$

where $q_1, \ldots, q_j$ are taken to be theoretical quantities; and let our

evidence e consist of an n-tuple $e_1, \ldots, e_n$ of data points. The hypothesis h and evidence e may be entirely unrelated intuitively; h might be some equation relating physical magnitudes to one another, and e might be some quantities derived from a sociological study. Yet, as long as h is not itself a mathematical identity (i.e., not every j-tuple of numbers is a positive instance of h), we can always construct an auxiliary hypothesis with respect to which e BS confirms h. Let $c_1, \ldots, c_j$ be a j-tuple of numbers that satisfies h, and $d_1, \ldots, d_j$ be one that does not. The auxiliary appropriate to the data points $e = \{e_1, \ldots, e_n\}$ can then be constructed as follows. Let F be a function which takes e onto $c_1, \ldots, c_j$ and all other n-tuples onto $d_1 \ldots, d_j$. Then consider the auxiliary T:

$$F(p) = q$$

where "p" is an n-tuple of "observational" quantities, and $q = \{q_1, \ldots, q_j\}$ the j-tuple of theoretical terms appearing in h. Clearly, e BS confirms h with respect to T, since, on the assumption of T, e constitutes a positive instance of h.

Given the ease with which we can come by BS confirmations, the question of comparative confirmation becomes quite crucial: why is it that some BS confirmations count for more than others? Why is it that we take BS confirmations with respect to some auxiliaries as seriously reflecting on the acceptability of the hypothesis, whereas we ignore the great mass of trivial BS confirmations, those relativized to ad-hoc auxiliaries? Glymour attempts to offer something of an answer:

> The distinctions that the strategy of testing makes with regard to what is tested by what with respect to what else are of use despite the fact that if a hypothesis is not tested by a piece of evidence with respect to a theory, there is always some *other* theory with respect to which the evidence confirms or disconfirms the hypothesis. It is important that the bearing of evidence is sensitive to the changes of theory, but the significance of that fact is not that the distinctions regarding evidential relevance are unimportant. For in considering the relevance of evidence to hypothesis, one is ordinarily concerned either with how the evidence bears on a hypothesis with respect to some *accepted* theory or theories, or else one is concerned with the bearing of the evidence on a hypothesis with respect to definite theory containing that hypothesis.[33]

Glymour is surely correct in his intuitions about what we ordinarily do. But this just rephrases the problem. Why *should* we do what we ordinarily do?

Why *should* we take some BS confirmations, those that use the "appropriate" auxiliaries more seriously than we take others? If it is permissible to take seriously a BS confirmation relative to an untested auxiliary or relative to the hypothesis itself being tested, as Glymour often insists, how can he disregard *any* BS confirmations?

What is missing from Glymour's theory of confirmation seems obvious. Glymour gives us *no* way of mediating the gap between any one BS confirmation of h, and our increased confidence in h; he gives us no way to gauge how *much* any one BS confirmation supports, h, and the factors that go into that determination. Although there may be a number of different ways of filling in this gap in Glymour's program, the earlier sections of this paper suggest one attractive solution. Earlier I offered a Bayesian response to the problem of old evidence, in which the problem is resolved by showing how confirmation in the cases at hand can be understood as proceeding by conditionalization on the discovery of some logical relation between the hypothesis and the evidence in question. Now the logical relation I talked about most explicitly was logical implication. But almost everything I said holds good for whatever conception of the logical relation we like: *and this includes the logical relation that Glymour explicates*, $[h \vdash e]_T$. This framework is ready-made to fill in the gap in Glymour's program. Within this framework, we can show how the discovery that a given e BS confirms h with respect to T may increase our confidence in h, given one group of priors, and how, given other priors, the discovery that e BS confirms h with respect to T may have little or no effect on our confidence in H. The Bayesian framework, as interpreted above, thus gives us the tools needed to distinguish between the effects that different BS confirmations may have on our confidence in h, and gives us a way of resolving the problem of the ad-hoc auxiliary. To those of us of the Bayesian persuasion, the conclusion is obvious: Glymour's theory of confirmation can be fully adequate only if it is integrated into a Bayesian theory of reasoning.[34]

Notes

1. The criticisms are widespread, but the following are representative of the literature: Henry Kyburg, "Subjective Probability: Criticisms, Reflections and Problems," *Journal of Philosophical Logic* 7 (1978): 157-180; Isaac Levi, "Indeterminate Probabilities," *Journal of Philosophy* 71 (1974); 391-418; and Glenn Shafer, *A Mathematical Theory of Evidence* (Princeton: Princeton University Press, 1976).

2. Clark Glymour, *Theory and Evidence* (Princeton: Princeton University Press, 1980), hereafter referred to as *T & E*.

3. See I. J. Good, "46656 Varieties of Bayesians." *The American Statistician* 25, 5 (Dec. 1971): 62-63.

4. Following Kolmogorov's influential systematization, *Foundations of the Theory of Probability* (New York: Chelsea Publishing Co., 1950), most mathematical treatments of the theory of probability take probability functions to be defined over structured collections of sets, (σ-) rings or (σ-) fields, or over Boolean algebras. For obvious reasons philosophers of the Bayesian persuasian have often chosen to define probabilities over sentences in formal languages. I shall follow this practice. Because of the structural similarities among the different approaches, though, many of the theorems carry over from one domain to another, and in what follows, I shall not make use of the mathematically special features of probability functions defined on languages. Although I talk of probability functions defined on sentences rather than propositions or statements, no philosophical point is intended. Any of these objects would do as well.

5. For a fuller treatment of the coherence theorem, originally due to de Finetti, see Abner Shimony, "Coherence and the Axioms of Probability." *Journal of Symbolic Logic* 20 (1955): 1-28 or John Kemeny, "Fair Bets and Inductive Probabilities." *Journal of Symbolic Logic* 20 (1955): 263-273. The coherence theorem is not the only argument Bayesians appeal to to argue that degrees of belief ought to be probabilities. See, e.g., the arguments given in L. J. Savage, *The Foundations of Statistics* (New York: John Wiley and Dover, 1954 and 1972), chapter 3, and R. T. Cox, *The Algebra of Probable Inference* (Baltimore: Johns Hopkins University Press, 1961). However, the coherence argument is often cited and well accepted by Bayesians. Moreover, the coherence condition is closely connected with the requirement of logical omniscience, which will be one of the central foci of this paper.

6. For justifications of conditionalization, see Bruno de Finetti, *Theory of Probability* vol. 1 (New York: John Wiley, 1974), section 4.5; and Paul Teller, "Conditionalization, Observation, and Change of Preference," in W. Harper and C. Hooker, *Foundations and Philosophy of Epistemic Applications of Probability Theory* (Dordrecht: Reidel, 1976), pp. 205-259. Among other rational ways of changing one's beliefs I would include the extension of conditionalization proposed by Richard Jeffrey in chapter 11 of his *The Logic of Decision* (New York: McGraw-Hill, 1965), and the sorts of changes that one makes upon discovering an incoherence in one's beliefs. The former is appropriate when changing one's beliefs on the basis of uncertain observation, and the latter when one discovers, e.g., that one attributes probability .5 to heads on a given coin, yet attributes probability .25 to a run of three heads on the same coin. There may be other alternatives to conditionalization, but we shall not consider them here.

7. Both of these conceptions of Bayesianism are widespread. For a statement of the thought-police model, see L. J. Savage, *The Foundations of Statistics* (New York: John Wiley and Dover, 1954 and 1972), p. 57, and for a statement of the ideal learning machine model, see Rudolph Carnap, The Aim of Inductive Logic, in E. Nagel, P. Suppes, and A. Tarski, *Logic, Methodology and Philosophy of Science* (Stanford: Stanford University Press, 1962), pp. 303-318.

8. *T & E*, pp. 85-6.

9. The historical problem, as posed, appears to presuppose that evidence for an hypothesis must somehow serve to increase the scientist's degree of belief in that hypothesis. This *may* not hold for *everything* that we want to call evidence. Peter Achinstein argues that the evidence for an hypothesis may not only fail to raise the scientist's degree of belief in that hypothesis, but might actually *lower* it! See his "Concepts of Evidence," *Mind* 87 (1978): 22-45, and "On Evidence: A Reply to Bar-Hillel and Margalit," *Mind* 90 (1981): 108-112. But be that as it may, it seems clear to me that in the sorts of cases Glymour cites in this connection, we are dealing with circumstances in which considerations relating to the evidence *do* increase the scientist's degree of belief in his hypothesis. Whatever more general account of the notion of evidence we might want to adopt, there is an important question as to how the Bayesian can account for that. Closely related to what I call the historical problem of old

evidence, the question as to how old evidence can increase the scientist's degree of belief in a new hypothesis, is the question of how the Bayesian is to deal with the introduction of new theories at all. This is especially difficult for what I shall later call global Bayesianism, where the enterprise is to trace out the changes that would occur in an ideally rational individual's degrees of belief as he acquires more and more experience, and where it is assumed that the degree of belief function is defined over some maximally rich global language capable of expressing all possible evidence and hypotheses. Since I shall reject global Bayesianism, I won't speculate on how a global Bayesian might respond. I shall assume that at any time, a new hypothesis can be introduced into the collection of sentences over which S's degree of belief function is defined, and his previous degree of belief function extended to include that new hypothesis, as well as all truth-functional combinations of that hypothesis with elements already in the domain of S's beliefs. The new degrees of belief will, of course, reflect S's confidence in the new hypothesis. Although these new degrees of belief will be prior probabilities in the strictest sense of the term, they will not be without ground, so to speak, since they may be based on the relations that the new hypothesis is known to bear to past evidence, other hypotheses already considered, and so on.

10. I am indebted to Brian Skyrms for pointing out the ambiguity in Glymour's problem.

11. See *T & E*, p. 87-91 for a development of this line of argument, along with Glymour's criticisms.

12. The logical probabilist, like Carnap, does not have to go to counterfactual degrees of belief to solve the ahistorical problem of old evidence. Since a logical c-function is taken to measure the degree of logical overlap between its arguments, we can always appeal to the value of "c(h/e)" as a measure of the extent to which e confirms h, regardless of whether or not we, as a mater of fact, happen to believe that e. But, as far as I can see, the logical probabilist will be in no better shape than his subjectivist comrade is with respect to the *historical* problem of old evidence. Even for Carnap's logically perfect learning machine, once e has been acquired as evidence, it is difficult to see how it could be used to increase the degree of confirmation of a new hypothesis. I would like to thank James Hawthorne for this observation.

13. e confirms h in the relevance sense iff learning that e would increase S's confidence or degree of belief in h. On the relations among the various senses of confirmation and the importance of the notion of relevance, see Wesley Salmon, Confirmation and Relevance, in Maxwell and Anderson, eds., *Minnesota Studies in Philosophy of Science*, vol. 6 (Minneapolis: University of Minnesota Press, 1975).

14. These worries are eloquently pressed by Ian Hacking in "Slightly More Realistic Personal Probability," *Philosophy of Science* 34 (1967): 311-325. Much of my own solution to the problem of logical omniscience is very much in the spirit of Hacking's, although the details of our two accounts differ significantly.

15. The evolving probability model is suggested by I. J. Good in a number of places, though I know of no place where he develops it systematically. See, e.g., "Corroboration, Explanation, Evolving Probability, Simplicity and a Sharpened Razor," *British Journal for the Philosophy of Science* 19 (1968): 123-143, esp. 125, 129; "Explicativity, Corroboration, and the Relative Odds of Hypotheses", *Synthèse* 30 (1975): 39-73, esp. 46, 57; and "Dynamic Probability, Computer Chess, and the Measurement of Knowledge," *Machine Intelligence* 8 (1977), pp. 139-150. Good's preferred name for the position is now "dynamic probability." A similar position is expressed by Richard Jeffrey in a short note, "Why I am a Born-Again Bayesian," dated 5 Feb. 1978 and circulated through the Bayesian samizdat. To the best of my knowledge, the conditionalization model does not appear in the literature, although it is consistent with the sort of approach taken in I. Hacking, "Slightly More Realistic Personal Probability."

16. My own intuition is that in any *actual* case, the *way* we change our beliefs upon discovering that h ⊢ e will be determined by the strength of our prior belief that h ⊢ e, and that it is because the evolving probability model leaves out any considerations of these prior beliefs that it suffers from radical indeterminacy. This, it seems to me, is where the evolving probability model differs most clearly from the conditionalization model, which does, of course, take into acount the relevant prior beliefs as prior probabilities.

17. See, e.g., Hacking, "Slightly More Realistic Personal Probability," and I. J. Good, *Probability and the Weighing of Evidence* (London: C. Griffin, 1950).

18. Carnap, in The Aim of Inductive Logic, is an example of such an approach.

19. The local approach is by far the dominant one among practicing Bayesian statisticians and decision theorists, although it is often ignored by philosophers. One exception to this is Abner Shimony, who takes locality to be of central importance to his own version of the Bayesian program. See his Scientific Inference, in R. C. Colodny, ed., *The Nature and Function of Scientific Theories* (Pittsburgh: University of Pittsburgh Press, 1970), pp. 79-172, esp. pp. 99-101.

20. Glymour's bootstrap theory of confirmation will be discussed below in section 6.

21. What this formalism does *not* allow is the embedding of the turnstile. So sentences like "[A ⊢ B] ⊢ C" and "A ⊢ [B ⊢ C]" will not be well formed. An extension of the language to include such sentences may be needed if we want to talk about the confirmation of sentences of the form "A ⊢ B" and the problem of old evidence as it arises at that level.

22. Not everything of interest can be derived from (K*). The following interesting properties are *not* derivable from (K*) and the axioms of the probability calculus alone:

(a) $P(A \vdash B \;\&\; A \vdash C) = P(A \vdash B \;\&\; C)$
(b) If A truth-functionally entails B in L, $P(A \vdash B) = 1$
(c) If A and B are truth functionally inconsistent in L, then $P(A \vdash B) = 0$
(d) $P(A \vdash B \vee A \vdash C) = P(A \vdash B \vee C)$
(e) $P(A \vdash B \;\&\; B \vdash C) \leq P(A \vdash C)$

Later we shall discuss adding (b). But any of these properties could be added as additional constraints. The more constraints we add, however, the less freedom S has in assigning probabilities, and the closer we get to the specter of logical omniscience.

23. (K*) will be satisfied if "A ⊢ B" is interpreted as "A & B," say.

24. I would like to thank William Tait for pointing out a mistake in an earlier and stronger but, unfortunately, false version of (T2), and for suggesting the method of proof used here. On the existence of strictly positive probabilities, see, e.g., A. Horn and A. Tarski, "Measures in Boolean Algebras," *Transactions of the American Mathematical Society* 64 (1948): 467-497; or J. L. Kelley, "Measures on Boolean Algebras," *Pacific Journal of Mathematics* 9 (1959): 1165-1177. The proof of Theorem 2.5 in Horn and Tarski suggests a simple way of actually constructing an infinite number of different strictly positive probabilities on L, one corresponding to each countably infinite ordered set of numbers in (0, 1) that sum to 1. Consequently, there are an infinite number of probabilities on L* having the properties specified in (T2).

25. Although the recent literature on probability and conditionals, both indicative and subjunctive, is vast, something should be said about the relation between my results here and what others have done on conditionals. Two constraints on probabilities of conditionals have been toyed with in the literature, Stalnaker's thesis and Harper's constraint:

(C1) $P(h \rightarrow e) = P(e/h)$
(C2) $P(h \rightarrow e) = 1$ iff $P(e/h) = 1$

Unfortunately both constraints seem too strong, and lead to triviality results. David Lewis has shown that if (C1) is satisfied, then P can take on at most four different values. See his Probabilities of Conditionals and Conditional Probabilities. *Philosophical Review* 85 (1976): 297-315. Similarly, Stalnaker has shown that if (C2) is satisfied, then $P(h \rightarrow e) = P(h \supset e)$. See Letter by Robert Stalnaker to W. L. Harper, in Harper and Hooker, *Foundations and Philosophy*, pp. 113-115. Neither of these arguments has gone without challenge. See, e.g., Bas van Fraassen's answer to Lewis, Probabilities of Conditionals, in Harper and Hooker, *Foundations and Philosophy*, pp. 261-308 and Harper's answer to Stalnaker in Ramsey Test Conditionals and Iterated Belief Change, in Harper and Hooker, *Foundations and Philoso-phy*, pp. 117-135. But (C1) and (C2) are obviously strong conditions that introduce substantial complications. Luckily I don't have to worry about the complications or the triviality proofs. (C1) and (C2) fail in my formalism when "→" is replaced by " ⊢ ." Instead, I am committed only to the following more modest constraint:

(C3) If $P(h \vdash e) = 1$ then $P(e/h) = 1$.

26. It would be unwise to adopt the slightly stronger constraint:

(**) If $P(A \supset B) = 1$, then $P(A \vdash B) = 1$.

(**) is certainly unnatural if " ⊢ " is interpreted as implication, since S could be certain of "A ⊃ B" because he was certain that A is false, say. Adopting (**) would also block our ability to use the formalism in the solution of the problem of old evidence, since (**) has the consequence that if $P(e) = 1$, then $P(h \vdash e) = 1$, no matter what h or e we are dealing with.

27. The same basic technique can be used to construct other probability functions of interest. If the conditions of the theorem are satisfied, and $\varepsilon = 0$, then P(A), P(B), and $P(A \vdash B)$ will all have the required values and $P(A/A \vdash B) = P(A)$. If ε is chosen to be in the interval $[-\delta', 0)$, where $\delta' = \min(rs, (1-r)(1-s))$, then $P(A/A \vdash B) < P(A)$.

28. It certainly will not be the case that *every* configuration of priors is such that the discovery that h ⊢ e will increase S's degree of belief that h. It can easily be shown that $P(h/h \vdash e) > P(h)$ if and only if $P(h \vdash e/h) > P(h \vdash e/\sim h)$. That is, the discovery that h ⊢ e will increase S's degree of belief in h if and only if S believes that it is more likely that h entails e if h is true than if it is false. (This has an obvious parallel in the case of e confirming h: the discovery that e confirms h if and only if e is more likely given h than it is given ~h.) It is obvious that this condition will not always be satisfied. For example, when e is known to be false it is clear that $P(h \vdash e/h)$ ought to be 0. Even when $P(e) = 1$, one would not always expect $P(h \vdash e/h)$ to be greater than $P(h \vdash e/\sim h)$ (let h be an arbitrary hypothesis in biology and e be Kepler's laws). I have found it impossible to specify in any illuminating way a set of circumstances in which it is always reasonable to expect that $P(h \vdash e/h) > P(h \vdash e/\sim h)$.

29. In the discussion period following this paper when it was presented at the Minnesota Center for Philosophy of Science, Clark Glymour suggested that the historical facts of the Einstein case do indeed agree with my analysis.

30. *T & E*, pp. 92-3.

31. *T & E*, p. 122.

32. I am appealing here to the formulation of bootstrap confirmation that Glymour outlines in *T & E*, pp. 116-117

33. *T & E*, pp. 120-121. Glymour elsewhere discusses how his method can distinguish between the confirmation afforded to whole theories, i.e., *collections* of hypotheses. See *T & E* pp. 152-155, 182, 352-353. But nothing Glymour says there touches on the problem that concerns me here, so far as I can see.

34. For a very different attempt to combine the bootstrap idea with Bayesian probability, see a paper that Glymour wrote after publishing *T & E*, "Bootstraps and Probabilities," *Journal of Philosophy* 77 (1980): 691-699. In that essay, Glymour uses the tools of subjective probability directly in the explication of the relation, "e BS confirms h with respect to T," rather than considering the probability function defined over instances of that relation, itself defined independently of probabilistic notions. I am inclined to agree with Paul Horwich in thinking that "Glymour's proposal may reduce under pressure to a trivial modification of probabilistic confirmation theories" ("The Dispensability of Bootstrap Conditions," *Journal of Philosophy* 77 (1980): 699-702, esp. 700), and I am inclined to think that my way of combining bootstraps with probability yields a much richer and more palatable mixture than does Glymour's.