objective mark by which, in actual cases, analytically true hypotheses may be picked out. The body of hypotheses by which a theoretical term is introduced is generally lost in scientific prehistory: witness "force" and "atom." And even the hypothesis by which a theoretical term comes to have a standard and accepted use may be disconfirmed and rejected: witness Newtonian dynamics. The arguments against a class of analytic truths, let alone the sort of analytic truths Carnap required, are too familiar to require further repetition. Equally, there is no obvious temptation to take the second idea seriously. There are a great many different logical forms that would, in one or another kind of case, permit the deduction of a singular theoretical sentence from a singular observational sentence. For example,

$$\forall x(Tx \supset Ox)$$

permits us to infer that $\sim Ta$ from $\sim Oa$. Other empiricists, Hempel especially, were quick to point out the lack of grounds for supposing that the connections between observation sentences and theoretical sentences could be made only by Carnap's reduction sentences.

The first idea remains, and it remains whether or not one takes seriously the distinction between observation sentences and others. Whenever our evidence is stated in terms narrower than those of our theory, the first idea contains a strategem for making the connection between evidence and theory: use some of the hypotheses to deduce, from the evidence statements, instances of other hypotheses. This Carnapian principle will form the core of the theory of Chapter V, and we shall see there that in combination with other more traditional principles in the philosophy of science it has an extraordinary and surprising power to explain and to rationalize various aspects of scientific method.

CHAPTER III

Why I Am Not a Bayesian*

THE AIM of confirmation theory is to provide a true account of the principles that guide scientific argument insofar as that argument is not, and does not purport to be, of a deductive kind. A confirmation theory should serve as a critical and explanatory instrument quite as much as do theories of deductive inference. Any successful confirmation theory should, for example, reveal the structure and fallacies, if any, in Newton's argument for universal gravitation, in nineteenth-century arguments for and against the atomic theory, in Freud's arguments for psychoanalytic generalizations. Where scientific judgments are widely shared, and sociological factors cannot explain their ubiquity, and analysis through the lens provided by confirmation theory reveals no good explicit arguments for the judgments, confirmation theory ought at least sometimes to suggest some good arguments that may have been lurking misperceived. Theories of deductive inference do that much for scientific reasoning insofar as that reasoning is supposed to be demonstrative. We can apply quantification theory to assess the validity of scientific arguments, and although we must almost always treat such arguments as enthymematic, the premises we interpolate are not arbitrary; in many cases, as when the same subject matter is under discussion, there is a common set of suppressed premises. Again, there may be differences about the correct logical form of scientific claims; differences of this kind result in (or from) different formalizations, for example, of classical mechanics. But such differences often make no difference for the assessment of validity in

* "Who cares whether a pig-farmer is a Bayesian?"—R. C. Jeffrey.

actual arguments. Confirmation theory should do as well in its own domain. If it fails, then it may still be of interest for many purposes, but not for the purpose of understanding scientific reasoning.

The aim of confirmation theory ought not to be simply to provide precise replacements for informal methodological notions, that is, explications of them. It ought to do more; in particular, confirmation theory ought to *explain* both methodological truisms and particular judgments that have occurred within the history of science. By "explain" I mean at least that confirmation theory ought to provide a rationale for methodological truisms and ought to reveal some systematic connections among them and, further, ought without arbitrary or question-begging assumptions to reveal particular historical judgments as in conformity with its principles.

Almost everyone interested in confirmation theory today believes that confirmation relations ought to be analyzed in terms of *probability* relations. Confirmation theory is the theory of probability plus introductions and appendices. Moreover, almost everyone believes that confirmation proceeds through the formation of conditional probabilities of hypotheses on evidence. The basic tasks facing confirmation theory are thus just those of explicating and showing how to determine the probabilities that confirmation involves, developing explications of such metascientific notions as "confirmation," "explanatory power," "simplicity," and so on in terms of functions of probabilities and conditional probabilities, and showing that the canons and patterns of scientific inference result. It was not always so. Probabilistic accounts of confirmation really became dominant only after the publication of Carnap's *Logical Foundations of Probability*,[1] although of course many probabilistic accounts had preceded Carnap's. An eminent contemporary philosopher[2] has compared Carnap's achievement in inductive logic with Frege's in deductive logic: just as before Frege there was only a small and theoretically uninteresting collection of principles of deductive inference, but after him the foundation of a systematic and profound theory of demonstrative reasoning, so with Carnap and inductive reasoning. After Carnap's *Logical Foundations*, debates over confirmation theory seem to have focused chiefly on the interpretation of probability and on the appropriate probabilistic explications of various meta-scientific notions. The meta-scientific notions remain controversial, as does the interpretation of probability, although increasingly logical interpretations of probability are giving way to the doctrine that probability is degree of belief.[3] In very recent years a few philosophers have attempted to apply probabilistic analyses to derive and to explain particular methodological practices and precepts, and even to illucidate some historical cases.

I believe these efforts, ingenious and admirable as many of them are, are nonetheless misguided. For one thing, probabilistic analyses remain at too great a distance from the history of scientific practice to be really informative about that practice, and in part they do so exactly because they are probabilistic. Although considerations of probability have played an important part in the history of science, until very recently explicit probabilistic arguments for the confirmation of various theories, or probabilistic analyses of data, have been great rarities in the history of science. In the physical sciences at any rate, probabilistic arguments have rarely occurred. Copernicus, Newton, Kepler, none of them give probabilistic arguments for their theories, nor does Maxwell or Kelvin or Lavoisier or Dalton or Einstein or Schrodinger or. . . . There are exceptions. Jon Dorling has discussed a seventeenth-century Ptolemaic astronomer who apparently made an extended comparison of Ptolemaic and Copernican theories in

[1] R. Carnap, *The Logical Foundations of Probability*, Chicago: University of Chicago Press, 1950.

[2] See Hilary Putnam, "Probability and Confirmation," in S. Morgenbesser, *Philosophy of Science Today*, New York: Basic Books, 1967.

[3] A third view, that probabilities are to be understood exclusively as frequencies, has been most ably defended by Wesley Salmon. See his *Foundations of Scientific Inference*, Pittsburgh: University of Pittsburgh Press, 1969.

probabilistic terms; Laplace, of course, gave Bayesian arguments for astronomical theories. And there are people, Maxwell for example, who scarcely give a probabilistic argument when making a case for or against scientific hypotheses but who discuss *methodology* in probabilistic terms. This is not to deny that there are many areas of contemporary physical science where probability figures large in confirmation; regression analysis is not uncommon in discussions of the origins of cosmic rays, correlation and analysis of variance in experimental searches for gravitational waves, and so on. It *is* to say that, explicitly, probability is a distinctly minor note in the history of scientific argument.

The rarity of probability considerations in the history of science is more an embarrassment for some accounts of probability than for others. Logical theories, whether Carnap's or those developed by Hintikka and his students, seem to lie at a great distance from the history of science. Still, some of the people working in this tradition have made interesting steps toward accounting for methodological truisms. My own inclination is to believe that the interest such investigations have stems more from the insights they obtain into syntactic versions of structural connections among evidence and hypotheses than to the probability measures they mesh with these insights. Frequency interpretations suppose that for each hypothesis to be assessed there is an appropriate reference class of hypotheses to which to assign it, and the prior probability of the hypothesis is the frequency of true hypotheses in this reference class. The same is true for statements of evidence, whether they be singular or general. The matter of how such reference classes are to be determined, and determined so that the frequencies involved do not come out to be zero, is a question that has only been touched upon by frequentist writers. More to the point, for many of the suggested features that might determine reference classes, we have no statistics and cannot plausibly imagine those who figure in the history of our sciences to have had them. So conceived, the history of scientific argument must turn out to be largely a history of fanciful

guesses. Further, some of the properties that seem natural candidates for determining reference classes for hypotheses—simplicity, for example, seem likely to give perverse results. We prefer hypotheses that posit simple relations among observed quantities, and so on a frequentist view should give them high prior probabilities. Yet simple hypotheses, although often very useful approximations, have most often turned out to be literally false.

At present, perhaps the most philosophically influential view of probability understands it to be degree of belief. The subjectivist Bayesian (hereafter, for brevity, simply Bayesian) view of probability has a growing number of advocates who understand it to provide a general framework for understanding scientific reasoning. They are singularly unembarrassed by the rarity of explicit probabilistic arguments in the history of science, for scientific reasoning need not be explicitly probabilistic in order to be probabilistic in the Bayesian sense. Indeed, a number of Bayesians have discussed historical cases within their framework. Because of its influence and its apparent applicability, in what follows it is to the subjective Bayesian account that I shall give my full attention.

My thesis is several-fold. First, there are a number of attempts to demonstrate a priori the rationality of the restrictions on belief and inference that Bayesians advocate. These arguments are altogether admirable but ought, I shall maintain, to be unconvincing. My thesis in this instance is not a new one, and I think many Bayesians do regard these a priori arguments as insufficient. Second, there are a variety of methodological notions that an account of confirmation ought to explicate, and methodological truisms involving these notions that a confirmation theory ought to explain: for example, variety of evidence and why we desire it, ad hoc hypotheses and why we eschew them, what separates a hypothesis integral to a theory from one "tacked on" to the theory, simplicity and why it is so often admired, why "de-occamized" theories are so often disdained, what determines when a piece of evidence is relevant to a hypothesis, what, if anything, makes the confirmation of one bit of theory by one bit of evidence stronger than the confirmation of another bit of theory (or possibly the same bit) by another (or possibly the same) bit of evidence. Although there are plausible Bayesian explications of some of these notions, there are not plausible Bayesian explications of others. Bayesian accounts of methodological truisms and of particular historical cases are of one of two kinds: either they depend on general principles restricting prior probabilities or they don't. My claim is that many of the principles proposed by the first kind of Bayesian are either implausible or incoherent, and that, for want of such principles, the explanations the second kind of Bayesians provide for particular historical cases and for truisms of method are chimeras. Finally, I claim that there are elementary but perfectly common features of the relation of theory and evidence that the Bayesian scheme cannot capture at all without serious—and perhaps not very plausible—revision.

It is not that I think the Bayesian scheme or related probabilistic accounts capture nothing. On the contrary, they are clearly pertinent where the reasoning involved is explicitly statistical. Further, the accounts developed by Carnap, his predecessors, and his successors are impressive systematizations and generalizations, in a probabilistic framework, of certain principles of ordinary reasoning. But so far as understanding scientific reasoning goes, I think it is very wrong to consider our situation to be analogous to that of post-Fregean logicians, our subject matter transformed from a hodgepodge of principles by a powerful theory whose outlines are clear. We flatter ourselves that we possess even the hodgepodge. My opinions are outlandish, I know; few of the arguments I shall present in their favor are new, and perhaps none of them is decisive. Even so, they seem sufficient to warrant taking seriously entirely different approaches to the analysis of scientific reasoning.

The theories I shall consider share the following framework, more or less: There is a class of sentences that express all

hypotheses and all actual or possible evidence of interest; the class is closed under Boolean operations. For each ideally rational agent, there is a function defined on all sentences such that, under the relation of logical equivalence, the function is a probability measure on the collection of equivalence classes. The probability of any proposition represents the agent's degree of belief in that proposition. As new evidence accumulates, the probability of a proposition changes according to Bayes' rule: the posterior probability of a hypothesis on the new evidence is equal to the prior conditional probability of the hypothesis on the evidence. This is a scheme shared by diverse accounts of confirmation. I call such theories "Bayesian," or sometimes, "personalist."

We certainly have *grades* of belief. Some claims I more or less believe, some I find plausible and tend to believe, others I am agnostic about, some I find implausible and far-fetched, still others I regard as positively absurd. I think everyone admits some such gradations, although descriptions of them might be finer or cruder. The personalist school of probability theorists claim that we also have *degrees* of belief, degrees that can have any value between 0 and 1 and that ought, if we are rational, to be representable by a probability function. Presumably, the degrees of belief are to covary with everyday gradations of belief, so that one regards a proposition as preposterous and absurd just if his degree of belief in it is somewhere near zero, and he is agnostic just if his degree of belief is somewhere near a half, and so on. According to personalists, then, an ideally rational agent always has his degrees of belief distributed so as to satisfy the axioms of probability, and when he comes to accept a new belief he also forms new *degrees* of belief by conditionalizing on the newly accepted belief. There are any number of refinements, of course, but that is the basic view.

Why should we think that we really do have *degrees* of belief? Personalists have an ingenious answer: people have them because we can measure the degrees of belief that people have. Assume that no one (rational) will accept a wager on which he expects a loss, but anyone (rational) will accept any wager on which he expects a gain. Then we can measure a person's degree of belief in proposition P by finding, for fixed amount v, the highest amount u such that the person will pay u in order to receive $u + v$ if P is true but receive nothing if P is not true. If u is the greatest amount the agent is willing to pay for the wager, his expected gain on paying u must be zero. The agent's gain if P is the case is v; his gain if P is not the case is $-u$. Thus

$$v \cdot \text{prob}(P) + (-u) \cdot \text{prob}(\sim P) = 0$$

Since $\text{prob}(\sim P) = 1 - \text{prob}(P)$ we have

$$\text{prob}(P) = u/(u + v)$$

The reasoning is clear: any sensible person will act so as to maximize his expected gain; thus, presented with a decision whether or not to purchase a bet, he will make the purchase just if his expected gain is greater than zero. So the betting odds he will accept determine his degree of belief.[4]

I think that this device really does provide evidence that we have, or can produce, degrees of belief, in at least some propositions, but at the same time it is evident that betting odds are not an unobjectionable device for the measurement of degrees of belief. Betting odds could fail to measure degrees of belief for a variety of reasons: the subject may not believe that the bet will be paid off if he wins, or he may doubt that it is clear what constitutes winning, even though it is clear what constitutes losing. Things he values other than monetary gain (or whatever) may enter into his determination of the expected utility of purchasing the bet: for example, he may place either a positive or a negative value on risk itself. And the very fact

[4] More detailed accounts of means for determining degrees of belief may be found in R. Jeffrey, *The Logic of Decision*, New York: McGraw-Hill, 1965. It is a curious fact that the procedures that Bayesians use for determining subjective degrees of belief empirically are an instance of the general strategy described in Chapter V below. Indeed, the strategy typically used to determine whether or not actual people behave as rational Bayesians involves the bootstrap strategy described in that chapter.

that he is offered a wager on P may somehow change his degree of belief in P.

Let us suppose, then, that we do have degrees of belief in at least some propositions, and that in some cases they can be at least approximately measured on an interval from 0 to 1. There are two questions: why should we think that, for rationality, one's degrees of belief must satisfy the axioms of probability, and why should we think that, again for rationality, changes in degrees of belief ought to proceed by conditionalization? One question at a time. In using betting quotients to measure degrees of belief it was assumed that the subject would act so as to maximize *expected* gain. The betting quotient determined the degree of belief by determining the coefficient by which the gain is multiplied in case that P is true in the expression for the expected gain. So the betting quotient determines a degree of belief, as it were, in the *role* of a probability. But why should the things, degrees of belief, that play this role, be probabilities? Supposing that we do choose those actions that maximize the sum of the product of our degrees of belief in each possible outcome of the action and the gain (or loss) to us of that outcome. Why must the degrees of belief that enter into this sum be probabilities? Again there is an ingenious argument: if one acts so as to maximize his expected gain using a degree-of-belief function that is not a probability function, and if for every proposition there were a possible wager (which, if it is offered, one believes will be paid off if it is accepted and won), then there is a circumstance, a combination of wagers, that one would enter into if they were offered, and in which one would suffer a net loss whatever the outcome. That is what the Dutch Book argument shows; what it councils is prudence.

Some of the reasons why it is not clear that betting quotients are accurate measures of degrees of belief are also reasons why the Dutch Book argument is not conclusive: there are many cases of propositions in which we may have degrees of belief, but on which, we may be sure, no acceptable wager will be offered us; again, we may have values other than the value we place on the stakes and these other values may enter into our determination whether or not to gamble; and we may not have adopted the policy of acting so as to maximize our expected gain or our expected utility: that is, we may save ourselves from having book made against us by refusing to make certain wagers, or combinations of wagers, even though we judge the odds to be in our favor.

The Dutch Book argument does not succeed in showing that in order to avoid absurd commitments, or even the possibility of such commitments, one must have degrees of belief that are probabilities. But it does provide a kind of justification for the personalist viewpoint, for it shows that if one's degrees of belief are probabilities, then a certain kind of absurdity is avoided. There are other ways of avoiding that kind of absurdity, but at least the personalist way is one such.[5]

One of the common objections to Bayesian theory is that it fails to provide any connection between what is inferred and what is the case. The Bayesian reply is that the method guarantees that, in the long run, everyone will agree on the truth. Suppose that B_i are a set of mutually exclusive, jointly exhaustive hypotheses, each with probability $B(i)$. Let $\bar{x}_r$ be a sequence of random variables with a finite set of values and conditional distribution given by $P(\bar{x}_r = x_r | B_i) = \varepsilon(x_r | B_i)$; then we can think of the values x_r as the outcomes of experiments, each hypothesis determining a likelihood for each outcome. Suppose that no two hypotheses have the same likelihood distribution, that is, for $i \neq j$ it is not the case that for all values x_r of $\bar{x}_r$, $\varepsilon(x_r | B_i) = \varepsilon(x_r | B_j)$, where the ε's are defined as above. Let $\bar{x}$ denote the first n of these variables, where x is a value of $\bar{x}$. Now imagine an observation of these n random variables. In Savage's words:

> Before the observation, the probability that the probability given x of whichever element of the partition actually obtains will be greater than α is
>
> $$\sum_i B(i) P(P(B_i | x) > \alpha | B_i)$$
>
> where summation is confined to those i's for which $B(i) \neq 0$.[6]

[5] Fur further criticisms of the Dutch Book argument see H. Kyburg, "Subjective Probability: Criticisms, Reflections and Problems," *Journal of Philosophical Logic* 7, 1978, pp. 157–180.

In the limit as n approaches infinity, the probability that the probability given x of whichever element of the partition actually obtains is greater than α is 1. That is the theorem. What is its significance? According to Savage, "with the observation of an abundance of relevant data, the person is almost certain to become highly convinced of the truth, and it has also been shown that he himself knows this to be the case."[7] That is a little misleading. The result involves second-order probabilities, but these too, according to personalists, are degrees of belief. So what has been shown seems to be this: In the limit as n approaches infinity an ideally rational Bayesian has degree of belief 1 that an ideally rational Bayesian (with degrees of belief as in the theorem) has degree of belief, given x, greater than α in whichever element of the partition actually obtains. The theorem does not tell us that in the limit any rational Bayesian will assign probability 1 to the true hypothesis and probability 0 to the rest; it only tells us that rational Bayesians are certain that he will. It may reassure those who are already Bayesians, but it is hardly grounds for conversion. Even the reassurance is slim. Mary Hesse[8] points out, entirely correctly I believe, that the assumptions of the theorem do not seem to apply even approximately in actual scientific contexts. Finally, some of the assumptions of stable estimation theorems can be dispensed with if one assumes instead that all of the initial distributions considered must agree regarding which evidence is relevant to which hypotheses. But there is no evident a priori reason why there should be such agreement.

[6] L. Savage, *The Foundations of Statistics*, New York: Dover, 1972, p. 49.

[7] Ibid., p. 50.

[8] See M. Hesse, *The Structure of Scientific Inference*, Berkeley: University of California Press, 1974, pp. 117–119.

I think relatively few Bayesians are actually persuaded of the correctness of Bayesian doctrine by Dutch Book arguments, stable estimation theorems, or other a priori arguments. Their frailty is too palpable. I think that the appeal of Bayesian doctrine derives from two other features. First, with only very weak or very natural assumptions about prior probabilities, or none at all, the Bayesian scheme generates principles that seem to accord well with common sense. Thus with minor restrictions one obtains the principle that hypotheses are confirmed by positive instances of them, and, again, one obtains the result that if an event that actually occurs is, on some hypothesis, very unlikely to occur, then that occurrence renders the hypothesis less likely than it would otherwise have been. These principles, and others, can claim something like the authority of common sense, and Bayesian doctrine provides a systematic explication of them. Second, the restrictions placed a priori on rational degrees of belief are so mild, and the device of probability theory at once so precise and so flexible, that Bayesian philosophers of science may reasonably hope to explain the subtleties and vagaries of scientific reasoning and inference by applying their scheme together with plausible assumptions about the distribution of degrees of belief. This seems, for instance, to be Professor Hesse's line of argument. After admitting the insufficiency of the standard arguments for Bayesianism, she sets out to show that the view can account for a host of alleged features of scientific reasoning and inference. My own view is different: particular *inferences* can almost always be brought into accord with the Bayesian scheme by assigning degrees of belief more or less ad hoc, but we learn nothing from this agreement. What we want is an explanation of scientific argument; what the Bayesians give us is a theory of learning, indeed a theory of personal learning. But arguments are more or less impersonal; I make an argument to persuade anyone informed of the premisses, and in doing so I am not reporting any bit of autobiography. To ascribe to me degrees of belief that make my slide from my premisses to my conclusion a plausible one

fails to explain anything not only because the ascription may be arbitrary, but also because, even if it is a correct assignment of my degrees of belief, it does not explain why what I am doing is *arguing*—why, that is, what I say should have the least influence on others, or why I might hope that it should. Now, Bayesians might bridge the gap between personal inference and argument in either of two ways. In the first place, one might give arguments in order to change others' beliefs because of the respect they have for his opinion. This is not very plausible; if that were the point of giving arguments, one would not bother with them but would simply state one's opinion. Alternatively, and more hopefully, Bayesians may suggest that we give arguments exactly because there are general principles restricting belief, principles that are widely subscribed to, and in giving arguments we are attempting to show that, supposing our audience has certain beliefs, they must in view of these principles have other beliefs, those we are trying to establish. There is nothing controversial about this suggestion, and I endorse it. What is controversial is that the general principles required for argument can best be understood as conditions restricting prior probabilities in a Bayesian framework. Sometimes they can, perhaps, but I think that when arguments turn on relating evidence to theory, it is very difficult to explicate them in a plausible way within the Bayesian framework. At any rate, it is worth seeing in more detail what the difficulties may be.

There is very little Bayesian literature about the hodgepodge of claims and notions that are usually canonized as scientific method; very little seems to have been written, from a Bayesian point of view, about what makes a hypothesis ad hoc, about what makes one body of evidence more various than another body of evidence, and why we should prefer a variety of evidence, about why, in some circumstances, we should prefer simpler theories, and what it is that we are preferring when we do. And so on. There is little to nothing of this in Carnap, and more recent, and more personalist, statements of the Bayesian position are almost as disappointing. In a lengthy discussion of what he calls "tempered personalism," Abner Shimony[9] discusses only how his version of Bayesianism generalizes and qualifies hypothetico-deductive arguments. (Shimony does discuss simplicity, but only to argue that it is overvalued.) Mary Hesse devotes the later chapters of her book to an attempt to show that certain features of scientific method do result when the Bayesian scheme is supplemented with a postulate that restricts assignments of prior probabilities. Unfortunately, as we shall see, her restrictive principle is incoherent.[10]

One aspect of the demand for a variety of evidence arises when there is some definite set of alternative hypotheses between which we are trying to decide. In such cases we naturally prefer the body of evidence that will be most helpful in eliminating false competitors. This aspect of variety is an easy and natural one for Bayesians to take account of, and within an account such as Shimony's it is taken care of so directly as hardly to require comment. But there is more to variety. In some situations we have some reason to suspect that if a theory is false, its falsity will show up when evidence of certain kinds is obtained and compared. For example, given the tradition of Aristotelian distinctions, there was some reason to demand both terrestrial and celestial evidence for seventeenth-century theories of motion that subjected all matter to the same dynamical laws. Once again, I see no special reason why this kind of demand for a variety of evidence cannot be fitted into the Bayesian scheme. But there is still more. A complex theory may contain a great many logically independent hypotheses, and particular bodies of evidence may provide grounds for some of those hypotheses but not for others. Surely part of the demand for a variety of evidence, and an important part, derives from a desire to see to it that the various independent

[9] See A. Shimony, "Scientific Inference," in R. Colodny, *The Nature and Function of Scientific Theories*, Pittsburgh: University of Pittsburgh Press, 1970.

[10] Moreover, I believe that much of her discussion of methodological principles has only the loosest relation to Bayesian principles.

parts of our theories are tested. Taking account of this aspect of the demand for a variety of evidence is just taking account of the relevance of evidence to pieces of theory. How Bayesians may do this we shall consider later.

Simplicity is another feature of scientific method for which some Bayesians have attempted to account. There is one aspect of the scientific preference for the simple that seems beyond Bayesian capacities, and that is the disdain for "deoccamized" hypotheses, for theories that postulate the operation of a number of properties, determinable only in combination, when a single property would do. Such theories can be generated by taking any ordinary theory and replacing some single quantity, wherever it occurs in the statement of the theory, by an algebraic combination of new quantities. If the original quantity was not one that occurs in the statement of some body of evidence for the theory, then the new, deoccamized theory will have the same entailment relations with that body of evidence as did the original theory. If the old theory entailed the evidence, so will the new, deoccamized one. Now, it follows from Bayesian principles that if two theories both entail e, then (provided the prior probability of each hypothesis is neither 1 nor 0) if e confirms one of them it confirms the other. How then is the fact (for so I take it to be) that pieces of evidence just don't seem to *count* for deoccamized theories to be explained? Not by supposing that deoccamized theories have lower prior probabilities than undeoccamized theories, for being "deoccamized" is a feature that a theory has only with respect to a certain body of evidence, and it is not hard to imagine artificially restricted bodies of evidence with respect to which perfectly good theories might count as deoccamized. Having extra wheels is a feature a theory has only in relation to a body of evidence; the only Bayesian relation that appears available and relevant to scientific preference is the likelihood of the evidence on the theory, and unfortunately the likelihood is the same for a theory and for its deoccamized counterparts whenever the theory entails the evidence.

It is common practice in fitting curves to experimental data, in the absence of an established theory relating the quantities measured, to choose the "simplest" curve that will fit the data. Thus linear relations are preferred to polynomial relations of higher degree, and exponential functions of measured quantities are preferred to exponential functions of algebraic combinations of measured quantities, and so on. The problem is to account for this preference. Harold Jeffreys,[11] a Bayesian of sorts, offered an explanation along the following lines. Algebraic and differential equations may be ordered by simplicity; the simpler the hypothetical relation between two or more quantities, the greater is its prior probability. If measurement error has a known probability distribution, we can then compute the likelihood of any set of measurement results given an equation relating the measured quantities. It should be clear, then, that with these priors and likelihoods, ratios of posterior probabilities may be computed from measurement results. Jeffreys constructed a Bayesian significance test for the introduction of higher-degree terms in the equation relating the measured quantities. Roughly, if one's equation fits the data *too* well, then the equation has too many terms and too many arbitrary parameters; and if the equation does not fit the data well enough, then one has not included enough terms and parameters in the equation. The whole business depends, of course, entirely on the ordering of prior probabilities. In this *Theory of Probability*[12] Jeffreys proposed that the prior probability of a hypothesis decreases as the number of arbitrary parameters increases, but hypotheses having the same number of arbitrary parameters have the same prior probability. This leads immediately to the conclusion that the prior probability of every hypothesis is zero. Earlier, Jeffreys proposed a slightly more complex assignment of priors that did not suffer this difficulty. The problem is not really one of finding a way to assign finite probabilities to an infinite number of incompatible hypotheses, for there are

[11] See H. Jeffreys, *Scientific Inference*, Cambridge: Cambridge University Press, 1973.

[12] H. Jeffreys, *Theory of Probability*, Oxford: Clarendon Press, 1967.

plenty of ways to do that. The trouble is that it is just very implausible that scientists typically have their prior degrees of belief distributed according to any plausible simplicity ordering, and still less plausible that they would be rational to do so. I can think of very few simple relations between experimentally determined quantities that have withstood continued investigation, and often simple relations are replaced by relations that are infinitely complex: consider the fate of Kepler's laws. Surely it would be naive for anyone to suppose that a set of newly measured quantities will truly stand in a simple relation, especially in the absence of a well-confirmed theory of the matter. Jeffreys' strategy requires that we proceed in ignorance of our scientific experience, and that can hardly be a rational requirement.

Consider another Bayesian attempt, this one due to Mary Hesse. Hesse puts a "clustering" constraint on prior probabilities: for any positive r, the conjunction of $r + 1$ positive instances of a hypothesis is more probable than a conjunction of r positive instances with one negative instance. This postulate, she claims, will lead us to choose, ceteris paribus, the most economical, the simplest, hypotheses compatible with the evidence. Here is the argument:

> Consider first evidence consisting of individuals a_1, a_2, . . . , a_n, all of which have properties P and Q. Now consider an individual a_{n+1} with property P. Does a_{n+1} have Q or not? If nothing else is known the clustering postulate will direct us to predict Q_{a+1} since, *ceteris paribus*, the universe is to be postulated to be as homogeneous as possible consistently with the data. . . . But this is also the prediction that would be made by taking the most economical general law which is both confirmed by the data and of sufficient content to make a prediction about the application of Q to a_{n+1}. For h = "All P are Q" is certainly more economical than the "gruified" conflicting hypothesis of equal content h': "All x up to a_n that are P and Q, and all other x that are P are $\sim Q$."

> It follows in the [case] considered that if a rule is adopted to choose the prediction resulting from the most probable hypothesis on grounds of content, or, in case of a tie in content, the most economical hypothesis on those of equal content, this rule will yield the same predictions as the clustering postulate.

Here is the argument applied to curve fitting:

> Let f be the assertion that two data points (x_1, y_1), (x_2, y_2) are obtained from experiments. . . . The two points are consistent with the hypothesis $y = a + bx$, and also of course with an indefinite number of other hypotheses of the form $y = a_0 + a_1 + \cdots + a_n x$, where the values of $a_0, \ldots, a_n$ are not determined by (x_1, y_1), (x_2, y_2). What is the most economical prediction of the y-value of a further point g, where the x-value of g is x_3? Clearly it is the prediction which uses only the information already contained in f, that is, the calculable values of a, b rather than a prediction which assigns arbitrary values to the parameters of a higher-order hypothesis. Hence the most economical prediction is about the point $g = (x_3, a + bx_3)$, which is also the prediction given by the "simplest" hypothesis on almost all accounts of the simplicity of curves. Translated into probabilistic language, this is to say that to conform to intuitions about economy we should assign higher initial probability to the assertion that points $(x_1, a + bx_1)$, $(x_2, a + bx_2)$, $(x_3, a + bx_3)$ are satisfied by the experiment, than to that in which the third point is inexpressible in terms of a and b alone. In this formulation economy is a function of finite descriptive lists of points rather than general hypotheses, and the relevant initial probability is that of a universe containing these particular points rather than that of a universe in which the corresponding general law is true. . . . Description in terms of a minimum number of parameters may therefore be regarded as another aspect of homogeneity or clustering of the universe.[13]

[13] Hesse, *The Structure of Scientific Inference*, pp. 230–232.

Hesse's clustering postulate applies directly to the curve-fitting case, for hèr clustering postulate then requires that if two paired values of x and y satisfy the predicate $y = ax + b$, then it is more probable than not that a third pair of values will satisfy the predicate. So the preference for the linear hypothesis in the next instance results from Hesse's clustering postulate and the probability axioms. Unfortunately, with trivial additional assumptions, everything results. For surely if $y = a + bx$ is a legitemate predicate, then so is $y = a_1 + b_1 x^2$, for any definite values of a_1 and b_1. Now Hesse's first two data points can be equally well described by $(x_1, a_1 + b_1 x_1^2)$ and $(x_2, a_1 + b_1 x_2^2)$, where

$$b_1 = \frac{y_1 - y_2}{x_1^2 - x_2^2} \qquad a_1 = y_1 - x_1^2\left(\frac{y_1 - y_2}{x_1^2 - x_2^2}\right)$$

Hence her first two data points satisfy both the predicate $y = a + bx$ and the predicate $y = a_1 + b_1 x^2$. So by the clustering postulate the probability that the third point satisfies the quadratic expression must be greater than one-half, and the probability that the third point satisfies the linear expression must also be greater than one-half, which is impossible.

Another Bayesian account of our preference for simple theories has recently been offered by Roger Rosencrantz.[14] Suppose that we have some criterion for "goodness of fit" of a hypothesis to data, for example, confidence regions based on the χ^2 distribution for categorical data, or in curve fitting perhaps that the average sum of squared deviations is less than some figure. Where the number of possible outcomes is finite we can compare the number of such possible outcomes that meet the goodness of fit criterion with the number that do not. This ratio Rosencrantz calls the "observed sample coverage" of the hypothesis. Where the possible outcomes are infinite, if the region of possible outcomes meeting the goodness-of-fit criterion is always bounded for all relevant hypotheses, we

[14] See R. Rosencrantz, "Simplicity," in W. Harper and C. Hooker, *Foundations and Philosophy of Statistical Inference*, Boston: Reidel, 1976.

can compare the volumes of such regions for different hypotheses and thus obtain a measure of comparative sample coverage.

It seems plausible enough that the smaller the observed sample coverage of a hypothesis, the more severely it is tested by observing outcomes. Rosencrantz' first proposal is this: the smaller the observed sample coverage, the simpler the hypothesis. But further, he proves the following for hypotheses about categorical data: If H_1 and H_2 are hypotheses with parameters and H_1 is a special case of H_2 obtained by letting a free parameter in H_2 take its maximum likelihood value, then if we average the likelihood of getting evidence that fits each hypothesis well enough over all the possible parameter values, the average likelihood of H_1 will be greater than the average likelihood of H_2. The conclusion Rosencrantz suggests is that the simpler the theory the greater the average likelihood of data that fits it sufficiently well. Hence, even if a simple theory has a lower prior probability than more complex theories, because the average likelihood is higher for the simple theory, its posterior probability will increase more rapidly than that of more complex theories. When sufficient evidence has accumulated the simple theory will be preferred. Rosencrantz proposes to identify average likelihood with support.

Rosencrantz' approach has many virtues; I shall concentrate on its vices. First, observed sample coverage does not correlate neatly with simplicity. If H is a hypothesis, T another utterly irrelevant to H and to the phenomena about which H makes predictions, then $H \& T$ will have the same observed sample coverage as does H. Further, if H^* is a deoccamization of H, then H^* and H will have the same observed sample coverage. Second, Rosencrantz' theorem does not establish nearly enough. It does not establish, for example, that in curve fitting the average likelihood of a linear hypothesis is greater than the average likelihood of a quadratic or higher-degree hypothesis. We cannot explicate support in terms of average likelihood unless we are willing to allow that evidence supports a deoccamized hypothesis as much as undeoccamized ones,

and a hypothesis with tacked-on parts as much as one without such superfluous parts.

Finally, we come to the question of the relevance of evidence to theory. When does a piece of evidence confirm a hypothesis according to the Bayesian scheme of things? The natural answer is that it does so when the posterior probability of the hypothesis is greater than its prior probability, that is, if the conditional probability of the hypothesis on the evidence is greater than the probability of the hypothesis. That is what the condition of positive relevance requires, and that condition is the one most commonly advanced by philosophical Bayesians. The picture is a kinematic one: a Bayesian agent moves along in time having at each moment a coherent set of degrees of belief; at discrete intervals he learns new facts and each time he learns a new fact, e, he revises his degrees of belief by conditionalizing on e. The discovery that e is the case has confirmed those hypotheses whose probability after the discovery is higher than their probability before. For several reasons, I think this account is unsatisfactory; moreover, I doubt that its difficulties are remediable without considerable changes in the theory.

The first difficulty is a familiar one. Let us suppose that we can divide the consequences of a theory into sentences consisting of reports of actual or possible observations, and simple generalizations of such observations, on the one hand; and on the other hand, sentences that are theoretical. Then the collection of "observational" consequences of the theory will always be at least as probable as the theory itself; generally, the theory will be less probable than its observational consequences. A theory is never any better established than is the collection of its observational consequences. Why, then, should we entertain theories at all? On the probabilist view, it seems, they are a gratuitous risk. The natural answer is that theories have some special function that their collection of observational consequences cannot serve; the function most frequently suggested is explanation—theories explain, their collection of observational consequences do not. But however sage this suggestion may be, it only makes more vivid the difficulty of the Bayesian way of seeing things. For whatever explanatory power may be, we should certainly expect that goodness of explanation will go hand in hand with warrant for belief, yet if theories explain and their observational consequences do not, the Bayesian must deny the linkage. The difficulty has to do both with the assumption that rational degrees of belief are generated by probability measures and with the Bayesian account of evidential relevance. Making degrees of belief probability measures in the Bayesian way already guarantees that a theory can be no more credible than any collection of its consequences. The Bayesian account of confirmation makes it impossible for a piece of evidence to give us more total credence in a theory than in its observational consequences. The Bayesian way of setting things up is a natural one, but it is not inevitable, and wherever a distinction between theory and evidence is plausible, it leads to trouble.

A second difficulty has to do with how praise and blame are distributed among the hypotheses of a theory. Recall the case of Kepler's laws, discussed in Chapter II. It seems that observations of a single planet (and, of course, the sun) might provide evidence for or against Kepler's first law (all planets move on ellipses) and for or against Kepler's second law (all planets move according to the area rule), but no observations of a single planet would constitute evidence for or against Kepler's third law (for any two planets, the ratio of their periods equals the $^3/_2$ power of the ratio of their distances). In Chapter II we saw that hypothetico-deductive accounts of confirmation have great difficulty explaining this elementary judgment. Can the Bayesians do any better? One thing that Bayesians can say (and some have said) is that our degrees of belief are distributed—and historically were distributed—so that conditionalizing on evidence about one planet may change our degrees of belief in the first and second laws, but not our degree of belief in the third law.[15] I don't see that this is an

[15] This is the account suggested by P. Horwich in "An Appraisal of Glymour's Confirmation Theory," *Journal of Philosophy* 75, 1978, pp. 98–113.

explanation for our intuition at all; on the contrary, it seems merely to restate (with some additional claims) what it is that we want to be explained. Are there any reasons why people had their degrees of belief so distributed? If their beliefs had been different, would it have been equally rational for them to view observations of Mars as a test of the third law, but not of the first? It seems to me that we never succeed in explaining a widely shared judgment about the relevance or irrelevance of some piece of evidence merely by asserting that degrees of belief happened to be so distributed as to generate those judgments according to the Bayesian scheme. Bayesians may instead try to explain the case by appeal to some structural difference among the hypotheses; the only gadget that appears to be available is the likelihood of the evidence about a single planet on various combinations of hypotheses. If it is supposed that the observations are such that Kepler's first and second laws entail their description, but Kepler's third law does not, then it follows that the likelihood of the evidence on the first and second laws—that is, the conditional probability of the evidence given those hypotheses—is unity, but the likelihood of the evidence on the third law may be less than unity. But any attempt to found an account of the case on these facts alone is simply an attempt at a hypothetico-deductive account. The problem is reduced to one already unsolved. What is needed to provide a genuine Bayesian explanation of the case in question (as well as of many others that could be adduced) is a *general* principle restricting conditional probabilities and having the effect that the distinctions about the bearing of evidence that have been noted here do result. Presumably any such principles will have to make use of relations of content or structure between evidence and hypothesis. The case does nothing to establish that no such principles exist; it does, I believe, make it plain that without them the Bayesian scheme does not *explain* even very elementary features of the bearing of evidence on theory.

A third difficulty has to do with Bayesian kinematics. 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, not on the basis of any startling new predictions derived from the theory, and presumably it was on the basis of such arguments that he won the adherence of his early disciples. 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 than half a century earlier. Other physicists found the argument enormously forceful, and it is a fair conjecture that without it the British would not have mounted the famous eclipse expedition of 1919. 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 in virtue of the positive relevance condition nor in virtue of the likelihood of e on T. None of the Bayesian mechanisms apply, and if we are strictly limited to them, we have the absurdity that old evidence cannot confirm new theory. The result is fairly stable. If the probability of e is very high but not unity, $\text{Prob}_t(e, T)$ will still be unity if T entails e, and so $\text{Prob}_t(T, e)$ will be very close to $\text{Prob}_t(T)$. How might Bayesians deal with the old evidence/ new theory problem?[16] Red herrings abound:

[16] All of the defenses sketched below were suggested to me by one or another philosopher sympathetic to the Bayesian view; I have not attributed the arguments to anyone for fear of misrepresenting them. Nonetheless, I thank Jon Dorling, Paul Teller, Daniel Garber, Ian Hacking, Patrick Suppes, Richard Jeffrey, and Roger Rosencrantz for valuable discussions and correspondence on the point at issue.

the prior probability of the evidence, Bayesians may object, is not really unity; when the evidence is stated as measured or observed values, the theory does not really entail that those exact values obtain; an ideal Bayesian would never suffer the embarrassment of a novel theory. None of these replies will do: the acceptance of old evidence may make the degree of belief in it as close to unity as our degree of belief in some bit of evidence ever is; although the exact measured value (of, for example, the perihelion advance) may not be entailed by the theory and known initial conditions, that the value of the measured quantity lies in a certain interval may very well be entailed, and that is what is believed anyway; and, finally, it is beside the point that an ideal Bayesian would never face a novel theory, for the idea of Bayesian confirmation theory is to explain scientific inference and argument by means of the assumption that good scientists are, about science at least, approximately ideal Bayesians, and we have before us a feature of scientific argument that seems incompatible with that assumption.

A natural line of defense lies through the introduction of counterfactual degrees of belief. When using Bayes' rule to determine the posterior probability of a new theory on old evidence, one ought not to use one's actual degree of belief in the old evidence, which is unity or nearly so; one ought instead to use the degree of belief one would have had in e if.... The problem is to fill in the blanks in such a way that it is both plausible that we have the needed counterfactual degrees of belief and that they do serve to determine how old evidence bears on new theory. I tend to doubt that there is such a completion. We cannot merely throw e and whatever entails e out of the body of accepted beliefs; we need some rule for determining a counterfactual degree of belief in e and a counterfactual likelihood of e on T. To simplify, let us suppose that T does logically entail e, so that the likelihood is fixed.

If one flips a coin three times and it turns up heads twice and tails once, in using this evidence to confirm hypotheses (e.g., of the fairness of the coin) one does not take the probability of two heads and one tail to be what it is after the flipping—namely, unity—but what it was before the flipping. In this case there is an immediate and natural counterfactual degree of belief that is used in conditionalizing by Bayes' rule. The trouble with the scientific cases is that no such immediate and natural alternative distribution of degree of belief is available. Consider someone trying, in a Bayesian way, to determine in 1915 how much Einstein's derivation of the perihelion advance confirmed general relativity. There is no single event, like the coin flipping, that makes the perihelion anomaly virtually certain; rather, Leverrier first computed the anomaly in the middle of the nineteenth century, Simon Newcomb calculated it again around 1890 using Leverrier's method but new values for planetary masses, and obtained a substantially higher value than had Leverrier. Both Newcomb and Leverrier had, in their calculations, approximated an infinite series by its first terms without any proof of convergence, thus leaving open the possibility that the entire anomaly was the result of a mathematical error; in 1912 Eric Doolittle calculated the anomaly by a wholly different method, free of any such assumption, and obtained virtually the same value as had Newcomb.[17] For actual historical cases, unlike the coin-flipping case, there is no single counterfactual degree of belief in the evidence ready to hand, for belief in the evidence sentence may have grown gradually—in some cases it may have even waxed, waned, and waxed again. So the old evidence/new theory problem cannot be assimilated to coin flipping.

The suggestion that what is required is a counterfactual degree of belief is tempting nonetheless; but there are other problems with it besides the absence of any unique historical degree of belief. A chief one is that various ways of manufacturing counterfactual degrees of belief in the evidence threaten

[17] The actual history is still more complicated. Newcomb and Doolittle obtained values for the anomaly differing by about two seconds of arc per century. Early in the 1920s, Grossmann discovered that Newcomb had made an error in calculation of about that magnitude.

us with incoherence. One suggestion, for example, is the following, used implicitly by some Bayesian writers. At about the time T is introduced, there will be a number of alternative competing theories available; call them $T_1, T_2, \ldots, T_k$, and suppose that they are mutually exclusive of T and of each other. Then $P(e)$ is equal to

$$P(T_1)P(e, T_1) + P(T_2)P(e, T_2) + \cdots + P(T_k)P(e, T_k) \\ + P(\sim(T_1 \mathrm{v} \cdots \mathrm{v} T_k)P(e, T_1 \mathrm{v} \cdots \mathrm{v} T_k)$$

and we may try to use this formula to evaluate the counterfactual degree of belief in e. The problem is with the last term. Of course, one could suggest that this term just be ignored when evaluating $P(e)$, but it is difficult to see within a Bayesian framework any rationale at all for doing so. For if one does ignore this term, then the collection of prior probabilities used to evaluate the posterior probability of T will not be coherent unless either the likelihood of e on T is zero or the prior probability of T is zero. One could remedy this objection by replacing the last term by

$$P(T)P(e, T)$$

but this will not do either, for if one's degree of belief in

$$P(T_1 \mathrm{v} T_2 \mathrm{v} \cdots \mathrm{v} T_k \mathrm{v} T)$$

is not unity, then the set of prior degrees of belief will still be incoherent. Moreover, not only will it be the case that if the actual degree of belief in e is replaced by a counterfactual degree of belief in e according to either of these proposals then the resulting set of priors will be incoherent, it will further be the case that if we conditionalize on e the resulting conditional probabilities will be incoherent. For example, if we simply delete the last term, one readily calculates that

$$P(T_1 \mathrm{v} \cdots \mathrm{v} T_k, e) = \frac{P(T_1 \mathrm{v} \cdots \mathrm{v} T_k)P(e, T_1 \mathrm{v} \cdots \mathrm{v} T_k)}{P(e, T_1 \mathrm{v} \cdots \mathrm{v} T_k)P(T_1 \mathrm{v} \cdots \mathrm{v} T_k)} = 1$$

and further that

$$P(T, e) = \frac{P(T)P(e, T)}{P(e, T_1 \vee \cdots \vee T_k)P(T_1 \vee \cdots \vee T_k)}$$

But because T is supposed inconsistent with $T_1 \vee \cdots \vee T_k$ and $P(T, e)$ is not zero, this is incoherent.

Let us return to the proposal that when new theory confronts old evidence we should look backwards to the time when the old evidence e had not yet been established and use for the prior probability of e whatever degree of belief we would have had at that time. We cannot just stick in such a counterfactual value for the prior probability of e and change nothing else without, as before, often making both prior and conditionalized probabilities incoherent. If we give all of our sentences the degree of belief they would have had in the relevant historical period (supposing we somehow know what period that is) and then conditionalize on e, incoherency presumably will not arise, but it is not at all clear how to combine the resulting completely counterfactual conditional probabilities with our actual degrees of belief. It does seem to me that the following rather elaborate procedure will work when a new theory is introduced. Starting with your actual degree of belief function P, consider the degree of belief you would have had in e in the relevant historical period, call it $H(e)$. Now change P by regarding $H(e)$ as an arbitrary change in degree of belief in e and using Richard Jeffrey's[18] rule

$$P'(S) = H(e)P(S, e) + (1 - H(e))P(S, \sim e)$$

Jeffrey's rule guarantees that P' is a probability function. Finally, conditionalize on e

$$P''(S) = P'(S, e)$$

and let P'' be your new actual degree of belief function. (Alternatively, P'' can be formed by using Jeffrey's rule a second time.)

[18] See Jeffrey, *The Logic of Decision*.

There remain a number of objections to the historical proposal. It is not obvious that there are, for each of us, degrees of belief we personally would have had in some historical period. It is not at all clear which historical period is the relevant one. Suppose, for example, that the gravitational deflection of sunlight had been determined experimentally around 1900, well before the introduction of general relativity.[19] In trying to assess the confirmation of general relativity, how far back in time should a twentieth-century physicist go under this supposition? If only to the nineteenth, then if he would have shared the theoretical prejudices of the period, gravitational deflection of light would have seemed quite probable. Where ought he to stop and why? But laying aside these difficulties, it is implausible indeed that such a historical Bayesianism, however intriguing a proposal, is an accurate account of the principles by which scientific judgments of confirmation are made. For if it were, then we should have to condemn a great mass of scientific judgments on the grounds that those making them had not studied the history of science with sufficient closeness to make a judgment as to what their their degrees of belief would have been in relevant historical periods. Combined with the delicacy that is required to make counterfactual degrees of belief fit coherently with actual ones, these considerations make me doubt that we should look to counterfactual degrees of belief for a plausible Bayesian account of how old evidence bears on new theory.

Finally, consider a quite different Bayesian response to the old evidence/new theory problem. Whereas the ideal Bayesian

[19] Around 1900 is fanciful, before general relativity is not. In 1914 E. Freundlich mounted an expedition to Russia to photograph the eclipse of that year in order to determine the gravitational deflection of starlight. At that time, Einstein had predicted an angular deflection for light passing near the limb of the sun that was equal in value to that derived from Newtonian principles by Soldner in 1801. Einstein did not obtain the field equations that imply a value for the deflection equal to twice the Newtonian value until late in 1915. Freundlich was caught in Russia by the outbreak of World War I and was interned there. Measurement of the deflection had to wait until 1919.

agent is a perfect logician, none of us are, and there are always consequences of our hypotheses that we do not know to be consequences. In the situation in which old evidence is taken to confirm a new theory, it may be argued that there is *something* new that is learned, and typically, what is learned is that the old evidence is entailed by the new theory. Some old anomalous result is lying about, and it is not this old result that confirms a new theory, but rather the new discovery that the new theory entails (and thus explains) the old anomaly. If we suppose that semirational agents have degrees of belief about the entailment relations among sentences in their language and that

$$P(h \vdash e) = 1 \quad \text{implies} \quad P(e, h) = 1$$

this makes a certain amount of sense. We imagine the semirational Bayesian changing his degree of belief in hypothesis h in light of his new discovery that h entails e by moving from his prior degree of belief in h to his conditional degree of belief in h given that e, that $h \vdash e$, and whatever background beliefs there may be. Old evidence can, in this vicarious way, confirm a new theory, then, provided that

$$P(h, b \,\&\, e \,\&\, (h \vdash e)) > P(h, b \,\&\, e)$$

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 a piece of theory, and it is in virtue of that connection that the evidence, if believed to be true, is thought to be evidence for the bit of theory. What I do not believe is that the relation that matters is simply the entailment relation between the theory, on the one hand, and the evidence on the other. The reasons that the relation cannot be simply that of entailment are exactly the reasons why the hypothetico-deductive account, discussed in the preceding chapter, is inaccurate; but the suggestion is at least correct in sensing that our judgment of the relevance of evidence to theory depends

on the perception of a structural connection between the two, and that 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. Save for the most radical innovations, scientists seem to be in close agreement regarding what would or would not be evidence relevant to a novel theory; claims as to the relevance to some hypothesis of some observation or experiment are frequently buttressed by detailed calculations and arguments. All of these features of the determination of evidential relevance suggest that that relation depends somehow on structural, objective features connecting statements of evidence and statements of theory. But if that is correct, what is really important and really interesting is what these structural features may be. The condition of positive relevance, even if it were correct, would simply be the least interesting part of what makes evidence relevant to theory.

None of these arguments is decisive against the Bayesian scheme of things, nor should they be, for in important respects that scheme is undoubtedly correct. But taken together, I think they do at least strongly suggest that there must be relations between evidence and hypotheses that are important to scientific argument and to confirmation but to which the Bayesian scheme has not yet penetrated.