WESLEY C. SALMON

Bayes's Theorem and the History of Science

0. Introduction

In his splendid introduction to this volume, Herbert Feigl rightly stresses the central importance of the distinction between the *context of discovery* and the *context of justification*. These terms were introduced by Hans Reichenbach to distinguish the social and psychological facts surrounding the discovery of a scientific hypothesis from the evidential considerations relevant to its justification.[1] The folklore of science is full of dramatic examples of the distinction; e.g., Kepler's mystical sense of celestial harmony[2] versus the confrontation of the postulated orbits with the observations of Tycho; Kekulé's drowsing vision of dancing snakes[3] versus the laboratory confirmation of the hexagonal structure of the benzine ring; Ramanujan's visitations in sleep by the Goddess of Namakkal[4] versus his waking demonstrations of the mathematical theorems. Each of these examples offers a fascinating insight into the personality of a working scientist, and each provides a vivid contrast between those psychological factors and the questions of evidence that must be taken into account in order to assess the truth or probability of the result. Moreover, as we all learned in our freshman logic courses, to confuse the source of a proposition with the evidence for it is to commit the genetic fallacy.

If one accepts the distinction between discovery and justification as viable, there is a strong temptation to maintain that this distinction marks the boundaries between history of science and philosophy of science. History is concerned with the *facts* surrounding the growth and development of science; philosophy is concerned with the *logical structure* of science, especially with the evidential relations between data and hypotheses or theories. As a matter of fact, Reichenbach described the transition from the context of discovery to the context of justification in terms of a *rational reconstruction*. On the one hand, the scientific innovator engages in thought processes that may be quite irrational or nonrational, manifesting no apparent logical structure: this is the road to discovery. On the other hand, when he wants to present his results to the community for judgment, he provides a reformulation in which the hypotheses and theories are shown in logical relation to the evidence that is offered in support of them: this is his rational reconstruction. The items in the context of discovery are *psychologically relevant* to the scientific conclusion; those in the context of justification are *logically relevant* to it. Since the philosopher of science is concerned with logical relations, not psychological ones, he is concerned with the rationally reconstructed theory, not with the actual process by which it came into being.

Views of the foregoing sort regarding the relations between the context of discovery and the context of justification have led to a conception of philosophy of science which might aptly be characterized as a "rational reconstructionist" or "logical reconstructionist" approach; this approach has been closely associated with the school of logical positivism, though by no means confined to it.[5] Critics of the reconstructionist view have suggested that it leaves the study of vital, living, growing science to the historian, while relegating philosophy of science to the dissection of scientific corpses—not the bodies of scientists, but of theories that have grown to the point of stagnation and ossification. According to such critics, the study of completed science is not the study of science at all. One cannot understand science unless he sees how it grows; to comprehend the logical structure of science, it is necessary to take account of scientific change and scientific revolution. Certain philosophers have claimed, consequently, that philosophy of science must deal with the logic of discovery as well as the logic of justification.[6] Philosophy of science, it has been said, cannot proceed apart from study of the history of science. Such arguments have

AUTHOR'S NOTE: This work was supported in part by a research grant from the National Science Foundation.

[1] *Experience and Prediction* (Chicago: University of Chicago Press, 1938), section 1. I have offered an elementary discussion of the distinction in *Logic* (Englewood Cliffs, N.J.: Prentice-Hall, 1963), sections 1–3.

[2] A. Pannekoek, *A History of Astronomy* (New York: Interscience, 1961), p. 235.

[3] J. R. Partington, *A History of Chemistry* (London: Macmillan, 1964), IV, 553ff.

[4] G. H. Hardy *et al.*, *Collected Papers of Srinivasa Ramanujan* (Cambridge: Cambridge University Press, 1927), p. xii.

[5] Reichenbach, for example, was not a logical positivist; indeed, he was one of the earliest and most influential critics of that school.

[6] See, for example, N. R. Hanson, "Is There a Logic of Discovery," in *Current Issues in the Philosophy of Science*, ed. Herbert Feigl and Grover Maxwell (New York: Holt, Rinehart and Winston, 1961), pp. 20–35.

led to a challenge of the very distinction between discovery and justification.[7] Application of this distinction, it is claimed, has led to the reconstructionist approach, which separates philosophy of science from real science, and makes philosophy of science into an unrealistic and uninteresting form of empty symbol manipulation.

The foregoing remarks make it clear, I hope, that the distinction between the context of discovery and the context of justification is a major focal point for any fundamental discussion of the relations between history of science and philosophy of science. As the dispute seems to shape up, the reconstructionists rely heavily upon the viability of a sharp distinction, and they apparently conclude that there is no very significant relation between the two disciplines. Such marriages as occur between them —e.g., the International Union of History and Philosophy of Science, the National Science Foundation Panel for History and Philosophy of Science, and the Departments of History and Philosophy of Science at Melbourne and Indiana—are all marriages of convenience. The anti-reconstructionists, who find a basic organic unity between the two disciplines, seem to regard a rejection of the distinction between discovery and justification as a cornerstone of their view. Whatever approach one takes, it appears that the distinction between the context of discovery and the context of justification is the first order of business.

I must confess at this point, if it is not already apparent, that I am an unreconstructed reconstructionist, and I believe that the distinction between the context of discovery and the context of justification is viable, significant, and fundamental to the philosophy of science. I do not believe, however, that this view commits me to an intellectual divorce from my historical colleagues; in the balance of this essay I should like to explain why. I shall not be concerned to argue in favor of the distinction, but shall instead try (1) to clarify the distinction, and repudiate certain common misconceptions of it, (2) to show that a clear analysis of the nature of scientific confirmation is essential to an understanding of the distinction, and that a failure to deal adequately with the logic of confirmation can lead to serious historical misinterpretations, and (3) to argue that an adequate conception of the logic of confirmation leads to basic, and largely unnoticed, logical functions of historical information. In other words, I shall be attempting to show how certain aspects of the relations between history

[7] E.g., Thomas S. Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1962), p. 9.

and philosophy of science can be explicated within the reconstructionist framework. Some of my conclusions may appear idiosyncratic, but I shall take some pains along the way to argue that many of these views are widely shared.

1. The Distinction between Discovery and Justification

When one presents a distinction, it is natural to emphasize the differences between the two sorts of things, and to make the distinction appear more dichotomous than is actually intended. In the present instance, some commentators have apparently construed the distinction to imply that, first of all, a creative scientist goes through a great succession of irrational (or nonrational) processes, e.g., dreaming, being hit on the head, pacing the floor, or having dyspepsia, until a full-blown hypothesis is born. Only after these processes have terminated does the scientist go through the logical process of mustering and presenting his evidence so as to justify his hypothesis. Such a conception would, of course, be factually absurd; discovery and justification simply do not occur in that way. A more realistic account might go somewhat as follows. A scientist, searching for a hypothesis to explain some phenomenon, hits upon an idea, but soon casts it aside because he sees that it is inconsistent with some theory he accepts, or because it does not really explain the phenomenon in question. This phase undoubtedly involves considerable logical inference; it might, for instance, involve a mathematical calculation which shows that the explanation in question would not account for a result of the correct order of magnitude. After more searching around—in the meantime perhaps he attends a cocktail party and spends a restless night—he hits upon another idea, which also proves to be inadequate, but he sees that it can be improved by some modification or other. Again, by logical inference, he determines that his new hypothesis bears certain relations to the available evidence. He further realizes, however, that although his present hypothesis squares with the known facts, further modification would make it simpler and give it a wider explanatory range. Perhaps he devises and executes additional tests to check the applicability of his latest revision in new domains. And so it goes. What I am trying to suggest, by such science fiction, is that the processes of discovery and justification are intimately intertwined, with steps of one type alternating with steps of the other. There is no reason to conclude, from a distinction between the context of discovery and the context of justification, that the entire process of discovery must be completed before the process of justification can begin, and that the rational reconstruction can be undertaken only after the creative work has ended. Such conclusions are by no means warranted by the reconstructionist approach.

There is, moreover, no reason to suppose that the two contexts must be mutually exclusive. Not only may elements of the context of justification be temporally intercalated between elements of the context of discovery, but the two contexts may have items in common. The supposition that this cannot happen is perhaps the most widespread misunderstanding of the distinction between the two contexts. The most obvious example is the case in which a person or a machine discovers the answer to a problem by applying an algorithm, e.g., doing a sum, differentiating a polynomial, or finding a greatest common divisor. Empirical science also contains routine methods for finding answers to problems—which is to say, for discovering correct hypotheses. These are often the kinds of procedures that can be delegated to a technician or a machine, e.g., chemical analyses, ballistic testing, or determination of physical constants of a new compound. In such cases, the process of discovery and the process of justification may be nearly identical, though the fact that the machine blew a fuse, or the technician took a coffee break, could hardly qualify for inclusion in the latter context. Even though the two contexts are not mutually exclusive, the distinction does not vanish. The context of discovery consists of a number of items related to one another by psychological relevance, while the context of justification contains a number of items related to one another by (inductive and deductive) logical relevance. There is no reason at all why one and the same item cannot be both psychologically and logically relevant to some given hypothesis. Each context is a complex of entities all of which are interrelated in particular ways. The contexts are contrasted with one another, not on the ground that they can have no members in common, but rather on the basis of differences in the types of relations they incorporate. The fact that the two contexts can have items in common does not mean that the distinction is useless or perverse, for there are differences between logical and psychological relevance relations which are important for the understanding of science.

The problem of scientific discovery does not end with the thinking up of a hypothesis. One has also to discover evidence and the logical connections between the evidence and the hypothesis. The process of discovery is, therefore, involved in the very construction of the rational reconstruction. When the scientist publishes his hypothesis as acceptable, confirmed, or corroborated, along with the evidence and arguments upon which the claim is based, he is offering *his* rational reconstruction (the one he has *discovered*), and is presumably claiming that it is logically sound. This is a fact about the scientist; his evidence and arguments satisfy him. A critic—scientist or philosopher—might, of course, show that he has committed a logical or methodological error, and consequently, that his rational reconstruction is unsound. Such an occurrence would belong to the context of justification. However, even if the argument seems compelling to an entire scientific community, it may still be logically faulty. The convincing character of an argument is quite distinct from its validity; the former is a *psychological* characteristic, the latter is *logical*. Once more, even though there may be extensive overlap between the contexts of discovery and justification, it is important not to confuse them.

Considerations of the foregoing sort have led to serious controversy over the appropriate role of philosophy of science. On the other hand, it is sometimes claimed that philosophy of science must necessarily be a historically oriented empirical study of the methods scientists of the past and present have actually used, and of the canons they have accepted. On the other hand, it is sometimes maintained that such factual studies of the methods of science belong to the domain of the historian, and that the philosopher of science is concerned exclusively with logical and epistemological questions. Proponents of the latter view—which is essentially the reconstructionist approach—may appear quite open to the accusation that they are engaged in some sort of scholastic symbol mongering which has no connection whatever with actual science. To avoid this undesirable state of affairs, it may be suggested, we ought to break down the distinctions between history and philosophy, between psychology and logic, and, ultimately, between discovery and justification.

There is, I believe, a better alternative. While the philosopher of science may be basically concerned with abstract logical relations, he can hardly afford to ignore the actual methods that scientists have found acceptable. If a philosopher expounds a theory of the logical structure of science according to which almost all of modern physical science is methodologically unsound, it would be far more reasonable to conclude that the philosophical reasoning had gone astray than to suppose that modern science is logically misconceived. Just as certain empirical facts, such as geometrical diagrams or soap film experiments, may have great heuristic value for mathematics, so too may the historical facts of scientific development provide indispensable guidance for the formal studies of the logician. In spite of this, the philosopher of science is properly concerned with issues of logical correctness which cannot finally be answered by appeal to the history of science. One of the problems with which the philosopher of science might grapple is the question of what grounds we have for supposing scientific knowledge to be superior to other alleged types of knowledge, e.g., alchemy, astrology, or divination. The historian may be quick to reply that *he* has the means to answer that question, in terms of the relative success of physics, chemistry, and astronomy. It required the philosophical subtlety of David Hume to realize that such an answer involves a circular argument.[8] The philosopher of science, consequently, finds himself attempting to cope with problems on which the historical data may provide enormously useful guidance, but the solutions, if they are possible at all, must be logical, not historical, in character. The reason, ultimately, is that justification is a normative concept, while history provides only the facts.

I have been attempting to explain and defend the distinction between discovery and justification largely by answering objections to it, rather than by offering positive arguments. My attitude, roughly, is that it is such a plausible distinction to begin with, and its application yields such rich rewards in understanding, that it can well stand without any further justification. Like any useful tool, however, it must be wielded with some finesse; otherwise the damage it does may far outweigh its utility.

2. Bayes's Theorem and the Context of Justification

It would be a travesty to maintain, in any simpleminded way, that the historian of science is concerned only with matters of discovery, and not with matters of justification. In dealing with any significant case, say the replacement of an old theory by a new hypothesis, the historian will be deeply interested in such questions as whether, to what extent, and in what manner the old theory has been disconfirmed; and similarly, what evidence is offered in support of the new hypothesis, and how adequate it is. How strongly, he may ask, are factors such as national rivalry among scientists, esthetic disgust with certain types of theories, personal idiosyncrasies of influential figures, and other nonevidential factors operative? Since science aspires to provide objective knowledge of the world, it cannot be under-

[8] For further elaboration of this point see my *Foundations of Scientific Inference* (Pittsburgh: University of Pittsburgh Press, 1967), pp. 5–17.

stood historically without taking very seriously the role of evidence in scientific development and change. Such historical judgments—whether a particular historical development was or was not rationally justified on the basis of the evidence available at the time—depend crucially upon the historian's understanding of the logic of confirmation and disconfirmation. If the historian seriously misunderstands the logic of confirmation, he runs the risk of serious historical misevaluation. And to the possible rejoinder that any historian worth his salt has a sufficiently clear intuitive sense of what constitutes relevant scientific evidence and what does not, I must simply reply that I am not convinced.

Perhaps the most widely held picture of scientific confirmation is one that had great currency in the nineteenth century; it is known as the hypothetico-deductive (H-D) method. According to this view, a scientific hypothesis is tested by deducing observational consequences from it, and seeing whether these consequences actually do transpire. If a given consequence does occur, it constitutes a confirming instance for the hypothesis; if it does not occur, it is a disconfirming instance. There are two rather immediate difficulties with this characterization, and they are easily repaired. First, a scientific hypothesis, by itself, ordinarily does not have any observational consequences; it is usually necessary to supply some empirically determined initial conditions to make it possible validly to deduce any observational consequences at all. For example, from Kepler's law of planetary motion alone, it is impossible to deduce the position of Mars at some future time, but with initial conditions on the motion of Mars at some earlier time, a prediction of the position is possible. Similarly, from Hooke's law alone it is impossible to predict the elongation of a spring under a given weight, but with an empirically determined coefficient of elasticity, the prediction can be deduced. Second, it is frequently, if not always, necessary to make use of auxiliary hypotheses in order to connect the observations with the hypothesis that is being tested. For example, if a medical experimenter predicts that a certain bacillus will be found in the blood of a certain organism, he must conjoin to his medical hypothesis auxiliary hypotheses of optics which pertain to the operation of his microscope, for only in that way can he establish a deductive connection between what he observes under the microscope and the actual presence of the microorganism. With these additions, the H-D method can be schematized as follows:

$$\begin{array}{l} H \text{ (hypothesis being tested)} \\ A \text{ (auxiliary hypotheses)} \\ I \text{ (initial conditions)} \\ \hline O \text{ (observational consequence)} \end{array}$$

Since we are not primarily interested in epistemological problems about the reliability of the senses, let us assume for the purposes of the present discussion that the initial conditions *I* have been established as true by observation and, in addition, that we can ascertain by observation whether the observational consequence O is true or false. Let us assume, moreover, that for purposes of the present test of our hypothesis *H*, the auxiliary hypotheses A are accepted as unproblematic.[9] With these simplifying idealizations, we can say that *H* implies O; consequently if O turns out to be false, it follows that *H* must be false—this is the deductively valid *modus tollens*. Given the truth of O, however, nothing follows deductively about the truth of *H*. To infer the truth of *H* from the truth of O in these circumstances is obviously the elementary deductive fallacy of *affirming the consequent*. According to the H-D view the truth of O does, nevertheless, tend to confirm or lend probability to *H*. Presumably, if enough of the right kinds of observational consequences are deduced and found by observation to be true—i.e., if enough observational predictions are borne out by experience—the hypothesis can become quite highly confirmed. Scientific hypotheses can never be completely and irrefutably verified in this manner, but they can become sufficiently confirmed to be scientifically acceptable. According to this H-D conception, induction—the logical relation involved in the confirmation of scientific hypotheses—is a kind of inverse of deduction. The fact that a true observational prediction follows deductively from a given hypothesis (in conjunction with initial conditions and auxiliary hypotheses) means, according to the H-D view, that a relation of inductive support runs in the reverse direction from O to *H*.

The H-D account of scientific confirmation is, it seems to me, woefully inadequate. The situation is nicely expressed in a quip attributed to Morris R. Cohen: A logic text is a book that contains two parts; in the first (on deduction) the fallacies are explained, and in the second (on induction) they are committed. Quite clearly, we need a more satisfactory account of

[9] This is, of course, an unrealistic assumption, for as Pierre Duhem pointed out, in many cases the appropriate move upon encountering a disconfirming case is rejection or modification of an auxiliary hypothesis, rather than rejection of the principal hypothesis. This point does not affect the present discussion.

scientific confirmation. Automatically transforming a deductive fallacy into a correct inductive schema may offer an appealing way to account for scientific inference, but certainly our forms of inductive inference ought to have better credentials than that. The main shortcomings of the H-D method are strongly suggested by the fact that, given *any* finite body of observational evidence, there are infinitely many hypotheses which are confirmed by it in exactly the same manner; that is, there are infinitely many alternative hypotheses that could replace our hypothesis *H* in the schema above and still yield a valid deduction. This point is obvious if one considers the number of curves that can be drawn through a finite set of points on a graph. Hence, Hooke's law, which says that a certain function is a straight line, and Kepler's first law, which says that a planetary orbit is an ellipse, could each be replaced by infinitely many alternatives that would give rise to precisely the same observational consequences as Hooke's and Kepler's laws respectively. As it stands, the H-D method gives us no basis whatever for claiming that either of these laws is any better confirmed by the available evidence than is any one of the infinitude of alternatives. Clearly it stands in dire need of supplementation.

When we look around for a more adequate account of scientific confirmation, it is natural to see whether the mathematical calculus of probability can offer any resources. If we claim that the process of confirmation is one of lending probability to a hypothesis in the light of evidence, it is reasonable to see whether there are any theorems on probability that characterize confirmation. If so, such a theorem would provide some sort of valid schema for formal confirmation relations. Theorems do not, of course, come labeled for their specific applications, but Bayes's theorem does seem well suited for this role.[10]

In order to illuminate the use of Bayes's theorem, let us introduce a simple game. This game is played with two decks of cards made up as follows: deck I contains eight red and four black cards; deck II contains four red and eight black cards. A turn begins with the toss of an ordinary die; if the side six appears the player draws from deck I, and if any other side comes up he draws from deck II. The draw of a red card constitutes a win. There is a simple way to calculate the probability of a win in this

[10] I have discussed the Bayesian conception of confirmation at some length in *Foundations of Scientific Inference*, pp. 108–131 (Bayes's theorem is deduced within the formal calculus on pp. 58–62), and in "Inquiries into the Foundations of Science," in *Vistas in Science*, ed. David L. Arm (Albuquerque: University of New Mexico Press, 1968), pp. 1–24. The latter article is the less technical of the two.

game. Letting $P(A,B)$ stand for the probability *from* A *to* B (i.e., the probability of B, given A), and letting A stand for tosses of the die, B for draws from deck I, and C for draws resulting in red, the following formula yields the desired probability:

$$(1) \qquad P(A,C) = P(A,B)P(A \,\&\, B,C) + P(A,\sim B)P(A \,\&\, \sim B,C)$$

where the ampersand stands for "and" and the tilde preceding a symbol negates it. Probability expressions appearing in the formula are $P(A,C)$, probability of a red card on a play of this game; $P(A,B)$, probability of drawing from deck I ($= 1/6$); $P(A,\sim B)$, probability of drawing from deck II ($= 5/6$); $P(A \,\&\, B,C)$, probability of getting red on a draw from deck I ($= 2/3$); $P(A \,\&\, \sim B,C)$, probability of getting red on a draw from deck II ($= 1/3$). The probability of a win on any given play is 7/18.

Suppose, now, that a player has just drawn a red card, but you failed to notice from which deck he drew. We ask, what is the probability that it was drawn from deck I? The probability we wish to ascertain is $P(A \,\&\, C, B)$, the probability that a play which resulted in a red card was one on which the die turned up six, and the draw was made from deck I. Bayes's theorem

$$(2) \qquad P(A \,\&\, C,B) = \frac{P(A,B)P(A \,\&\, B,C)}{P(A,B)P(A \,\&\, B,C) + P(A,\sim B)P(A \,\&\, \sim B,C)}$$

supplies the answer. Substituting the available values on the right-hand side of the equation yields the value 2/7 for the desired probability. Note that although the probability of getting a red card if you draw from deck I is much greater than the probability of getting a red card if you draw from deck II, the probability that a given red draw came from deck I is much less than the probability that it came from deck II. This is because the vast majority of draws are made from deck II.

There is nothing controversial about either of the foregoing formulas, or about their application to games of the kind just described. The only difficulty concerns the legitimacy of extending the application of Bayes's theorem, formula (2), to the problem of confirmation of hypotheses. In order to see how that might go, let me redescribe the game, with some admitted stretching of usage. We can take the draw of a red card as an effect that can be produced in either of two ways, by throwing a six and drawing from deck I, or by tossing some other number and drawing from the other deck. There are, correspondingly, two causal hypotheses. When we ask for the probability that the red draw came from deck I, we are ask-

ing for the probability of the first of these hypotheses, given the evidence that a red card had been drawn. Looking now at the probability expressions that appear in Bayes's theorem, we have: $P(A,B)$, the *prior probability* of the first hypothesis; $P(A,\sim B)$, the *prior probability* that the first hypothesis does not hold; $P(A \,\&\, B,C)$, the probability of the effect (red card drawn) if the first hypothesis is correct; $P(A \,\&\, \sim B,C)$, the probability of the effect if the first hypothesis is incorrect; $P(A \,\&\, C,B)$, the *posterior probability* of the first hypothesis on the evidence that the effect has occurred. The probabilities $P(A \,\&\, B,C)$ and $P(A \,\&\, \sim B,C)$ are called *likelihoods* of the two hypotheses, but it is important to note clearly that they are not probabilities *of hypotheses* but, rather, probabilities *of the effect*. It is the posterior probability that we seek when we wish to determine the probability of the hypothesis in terms of the given evidence.

In order to apply Bayes's theorem, we must have three probabilities to plug into the right-hand side of (2). Since the two prior probabilities must add up to one, it is sufficient to know one of them, but the likelihoods are independent, so we just have both of them. Thus, in order to compute the posterior probability of our hypothesis, we need its prior probability, the probability that we would get the evidence we have if it is true, and the probability that we would get the evidence we have if it were false. None of these three is dispensable, except in a few obvious special cases.[11]

When the H-D schema was presented, we stipulated that the hypothesis being tested implied the evidence, so in that case $P(A \,\&\, B,C) = 1$. This value of one of the likelihoods does not determine a value for the posterior probability, and, indeed, the posterior probability can be arbitrarily small even in the case supplied by the H-D method. This fact shows the inadequacy of the H-D schema quite dramatically: even though the data confirm the hypothesis according to the H-D view, the posterior probability of the hypothesis in the light of the available evidence may be as small as you like—even zero in the limiting special case in which the prior probability of the hypothesis is zero.

If Bayes's theorem provides a correct formal schema for the logic of confirmation and disconfirmation of scientific hypotheses, it tells us that we need to take account of three factors in attempting to assess the degree to which a hypothesis is rendered probable by the evidence. Roughly, it says,

[11] Viz., if $P(A,B) = 0$ or $P(A \,\&\, B,C) = 0$, then $P(A \,\&\, C,B) = 0$; if $P(A,\sim B) = 0$ or $P(A \,\&\, \sim B,C) = 0$, then $P(A \,\&\, C,B) = 1$. Also, if $P(A,C) = 0$, the fraction becomes indeterminate, for, by (1), that is the denominator in (2).

we must consider how well our hypothesis explains the evidence we have (this is what the H-D schema requires), how well an alternative hypothesis might explain the same evidence, and the prior probability of the hypothesis. The philosophical obstacle that has always stood in the way of using Bayes's theorem to account for confirmation is the severe difficulty in understanding what a prior probability could be. I have argued elsewhere that it is essentially an assessment of what one might call the *plausibility* of the hypothesis, prior to, or apart from, the results of directly testing that hypothesis.[12] Without attempting to analyze what is meant by plausibility, I shall offer a few plausibility judgments of my own, just to illustrate the sort of thing I am talking about. For instance, I regard as quite implausible Velikovsky's hypotheses about the origin of Venus, any ESP theory that postulates transfer of information at a speed greater than that of light, and any teleological biological theory. Hypotheses of these kinds strike me as implausible because, in one way or another, they do not fit well with currently accepted scientific theory. I regard it as quite plausible that life originated on the face of the earth in accordance with straightforward physicochemical principles governing the formation of large "organic" molecules out of simpler inorganic ones. This does seem to fit well with what we know. You need not accept my plausibility judgments; you can supply your own. The only crucial issue is the existence of such prior probabilities for use in connection with Bayes's theorem.

Let us now return to the problems of the historian. I claim above that the analysis of the logic of confirmation could have a crucial bearing upon historical judgments. Having compared the H-D account of confirmation with the Bayesian analysis, we can see an obvious way in which this problem could arise. If a historian accepts the H-D analysis of confirmation, then there is no place for plausibility judgments in the logic of science—at least not in the context of justification. If such a historian finds plausibility considerations playing an important role historically in the judgments scientists render upon hypotheses, he will be forced to exclude them from the context of justification, and he may conclude that the course of scientific development is massively influenced by nonrational or nonevidential considerations. Such an "H-D historian" might well decide, along with the editors of *Harper's Magazine*, that it was scientific prejudice, not objective evaluation, that made the scientific community largely ignore Velikovsky's

[12] *Foundations of Scientific Inference* and "Inquiries into the Foundations of Science."

views.[13] He might similarly conclude that Einstein's commitment to the "principle of relativity" on the basis of plausibility arguments shows his views to have been based more upon preconceptions than upon objective evidence.[14] A "Bayesian historian," in contrast, will see these plausibility considerations as essential parts of the logic of confirmation, and he will place them squarely within the context of justification. The consequence is, I would say, that the historian of science who regards the H-D schema as a fully adequate characterization of the logical structure of scientific inference is in serious danger of erroneously excluding from the context of justification items that genuinely belong within it. The moral for the historian should be clear. There are considerations relating to the acceptance or rejection of scientific hypotheses which, on the H-D account, must be judged *evidentially* irrelevant to the truth or falsity of the hypothesis, but which are, nevertheless, used by scientists in making decisions about such acceptance or rejection. These same items, on the Bayesian account, become evidentially relevant. Hence, the judgment of whether scientists are making decisions on the basis of evidence, or on the basis of various psychological or social factors that are evidentially irrelevant, hinges crucially upon the question of whether the H-D or the Bayesian account of scientific inference is more nearly correct. It is entirely conceivable that one historian might attribute acceptance of a given hypothesis to nonrational considerations, while another might judge the same decision to have an entirely adequate rational basis. Which account is historically more satisfactory will depend mainly upon which account of scientific inference is more adequate. The historian can hardly be taken to be unconcerned with the context of justification, and with its differences from the context of discovery; indeed, if he is to do his job properly he must understand them very well.

3. The Status of Prior Probabilities

It would be rather easy, I imagine, for the historian, and others who are not intimately familiar with the technicalities of inductive logic and confirmation theory, to suppose that the H-D account of scientific inference is the correct one. This view is frequently expressed in the opening pages of introductory science texts, and in elementary logic books.[15] At the same

[13] This case is discussed in "Inquiries into the Foundations of Science."

[14] See Albert Einstein, "Autobiographical Notes," in *Albert Einstein: Philosopher-Scientist*, ed. Paul Arthur Schilpp (New York: Tudor, 1949), pp. 2–95.

[15] In my *Logic*, section 23, I have tried, without introducing any technicalities of

time, it is important to raise the question of whether scientists in general—including the authors of the aforementioned introductory texts—actually comply with the H-D method in practice, or whether in fact they use something similar to the Bayesian approach sketched in the preceding section. I am strongly inclined to believe that the Bayesian schema comes closer than the H-D schema to capturing actual scientific practice, for it seems to me that scientists do make substantial use of plausibility considerations, even though they may feel somewhat embarrassed to admit it. I believe also that practicing scientists have excellent intuitions regarding what constitutes sound scientific methodology, but that they may not always be especially adept at fully articulating them. If we want the soundest guidance on the nature of scientific inference, we should look carefully at scientific practice, rather than the methodological pronouncements of scientists.

It is, moreover, the almost universal judgment of contemporary inductive logicians—the experts who concern themselves explicitly with the problems of confirmation of scientific hypotheses—that the simple H-D schema presented above is incomplete and inadequate. Acknowledging the well-known fact that there is very little agreement on which particular formulation among many available ones is most nearly a correct inductive logic, we can still see that among a wide variety of influential current approaches to the problems of confirmation, there is at least agreement in rejecting the H-D method. This is not the place to go into detailed discussions of the alternative theories, but I should like to mention five leading candidates, indicating how each demands something beyond what is contained in the H-D schema. In each case, I think, what needs to be added is closely akin to the plausibility considerations mentioned in the preceding section.

1. The most fully developed explicit confirmation theory available is Rudolf Carnap's theory of logical probability (degree of confirmation) contained in his monumental *Logical Foundations of Probability*.[16] In the systems of inductive logic he elaborated in that book, one begins with a formalized language and assigns a priori weights to all statements in that language, including, of course, all hypotheses. It is very easy to show that Carnap's theory of confirmation is thoroughly Bayesian, with the a priori

the probability calculus, to offer an introductory Bayesian account of the confirmation of hypotheses.

[16] Chicago: University of Chicago Press, 1950.

weights functioning precisely as the prior probabilities in Bayes's theorem. Although these systems had the awkward feature that general hypotheses all have prior probabilities, and consequently, posterior probabilities on any finite amount of evidence, equal to zero, Jaakko Hintikka has shown how this difficulty can be circumvented without fundamentally altering Carnap's conception of confirmation as a logical probability.[17]

2. Although not many exponents of the frequency theory of probability will agree that it even makes sense to talk about the probability of scientific hypotheses, those who do explicitly invoke Bayes's theorem for that purpose. Reichenbach is the leading figure in this school, although his treatment of the probability of hypotheses is unfortunately quite obscure in many important respects.[18] I have tried to clarify some of the basic points of misunderstanding.[19]

3. The important "Bayesian" approach to the foundations of statistics has become increasingly influential since the publication in 1954 of L. J. Savage's *The Foundations of Statistics*.[20] It has gained many adherents among philosophers as well as statisticians. This view is based upon a subjective interpretation of probability ("personal probability," as Savage prefers to say, in order to avoid confusion with earlier subjective interpretations), and it makes extensive use of Bayes's theorem. The prior probabilities are simply degrees of prior belief in the hypothesis, before the concrete evidence is available. The fact that the prior probabilities are so easily interpreted on this view means that Bayes's theorem is always available for use. Savage, himself, is not especially concerned with probabilities of general hypotheses, but those who are interested in such matters have a ready-made Bayesian theory of confirmation.[21] On this view, the prior probabilities are subjective plausibility judgments.

4. Nelson Goodman, whose influential *Fact, Fiction, and Forecast* poses and attempts to resolve "the new riddle of induction," clearly recognizes

[17] Jaakko Hintikka, "A Two-Dimensional Continuum of Inductive Methods," in *Aspects of Inductive Logic*, ed. Jaakko Hintikka and Patrick Suppes (Amsterdam: North-Holland, 1966), pp. 113–132.

[18] Hans Reichenbach, *The Theory of Probability* (Berkeley and Los Angeles: University of California Press, 1949), section 85.

[19] *Foundations of Scientific Inference*, pp. 115ff.

[20] New York: Wiley, 1954. An excellent exposition is found in Ward Edwards, Harold Lindman, and Leonard J. Savage, "Bayesian Statistical Inference for Psychological Research," *Psychological Review*, 70 (1963), 193–242.

[21] Sir Harold Jeffreys illustrates an explicitly Bayesian approach to the probability of hypotheses; see his *Scientific Inference* (Cambridge: Cambridge University Press, 1957), and *Theory of Probability* (Oxford: Clarendon Press, 1939).

that there is more to confirmation than mere confirming instances.[22] He attempts to circumvent the difficulties, which are essentially those connected with the H-D schema, by introducing the notion of "entrenchment" of terms that occur in hypotheses. Recognizing that a good deal of the experience of the human race becomes embedded in the languages we use, he brings this information to bear upon hypotheses that are candidates for confirmation. Although he never mentions Bayes's theorem or prior probabilities, the chapter in which he presents his solution can be read as a tract on the Bayesian approach to confirmation.

5. Sir Karl Popper rejects entirely the notions of confirmation and inductive logic.[23] His concept of *corroboration*, however, plays a central role in his theory of scientific methodology. Although corroboration is explicitly regarded as nonprobabilistic, it does offer a measure of how well a scientific hypothesis has stood up to tests. The measure of corroboration involves such factors as simplicity, content, and testability of hypotheses, as well as the seriousness of the attempts made to falsify them by experiment. Although Popper denies that a highly corroborated hypothesis is highly probable, the highly corroborated hypothesis does enjoy a certain status: it may be chosen over its less corroborated fellows for further testing, and if practical needs arise, it may be used for purposes of prediction. The important point, for the present discussion, is that Popper rejects the H-D schema, and introduces additional factors into his methodology that play a role somewhat analogous to our plausibility considerations.

The foregoing survey of major contemporary schools of thought on the logic of scientific confirmation strongly suggests not only that the naive H-D schema is *not universally accepted* nowadays by inductive logicians as an adequate characterization of the logic of scientific inference, but also that it is *not even a serious candidate* for that role. Given the wide popular acceptance of the H-D method, it seems entirely possible that significant numbers of historians of science may be accepting a view of confirmation that is known to be inadequate, and one which differs from the current serious contending views in ways that can have a profound influence upon historical judgments. It seems, therefore, that the branch of contemporary philosophy of science that deals with inductive logic and confirmation theory may have some substantive material that is highly relevant to the professional activities of the historian of science.

[22] First edition (Cambridge, Mass.: Harvard University Press, 1955); second edition (Indianapolis: Bobbs-Merrill, 1965).

[23] *The Logic of Scientific Discovery* (New York: Basic Books, 1959). It is to be noted that, in spite of the title of his book, Popper accepts the distinction between discovery and justification, and explicitly declares that he is concerned with the latter but not the former.

It is fair to say, I believe, that one of the most basic points on which the leading contemporary theories of confirmation differ from one another is with regard to the nature of the prior probabilities. As already indicated, the logical theorist takes the prior probability as an a priori assessment of the hypothesis, the personalist takes the prior probability as a measure of subjective plausibility, the frequentist must look at the prior probability as some sort of success frequency for a certain type of hypothesis, Goodman would regard the prior probability as somehow based upon linguistic usage, and Popper (though he violently objects to regarding it as a prior probability) needs something like the potential explanatory value of the hypothesis. In addition, I should remark, N. R. Hanson held plausibility arguments to belong to the logic of discovery, but I have argued that, on his own analysis, they have an indispensable role in the logic of justification.[24]

This is not the place to go into a lengthy analysis of the virtues and shortcomings of the various views on the nature of prior probabilities.[25] Rather, I should like merely to point out a consequence of my view that is quite germane to the topic of the conference. If one adopts a frequency view of probability, and attempts to deal with the logic of confirmation by way of Bayes's theorem (as I do), then he is committed to regarding the prior probability as some sort of frequency—e.g., the frequency with which hypotheses relevantly similar to the one under consideration have enjoyed significant scientific success. Surely no one would claim that we have reliable statistics on such matters, or that we can come anywhere near assigning precise numerical values in a meaningful way. Fortunately, that turns out to be unnecessary; it is enough to have very, very rough estimates. But this approach does suggest that the question of the plausibility of a scientific hypothesis has something to do with our experience in dealing with scientific hypotheses of similar types. Thus, I should say, the reason I would place a rather low plausibility value on teleological hypotheses is closely related to our experience in the transitions from teleological to mechanical explanations in the physical and biological sciences and, to some extent, in the social sciences. To turn back toward telelogical hypotheses

[24] *Foundations of Scientific Inference*, pp. 111–114, 118.

[25] This is done in the items mentioned in footnote 10.

would be to go against a great deal of scientific experience about what kinds of hypotheses work well scientifically. Similarly, when Watson and Crick were enraptured with the beauty of the double helix hypothesis for the structure of the DNA molecule, I believe their reaction was more than purely esthetic.[26] Experience indicated that hypotheses of that degree of simplicity tend to be successful, and they were inferring that it had not only beauty, but a good chance of being correct. Additional examples could easily be exhibited.

If I am right in claiming not only that prior probabilities constitute an indispensable ingredient in the confirmation of hypotheses and the context of justification, but also that our estimates of them are based upon empirical experience with scientific hypothesizing, then it is evident that the history of science plays a crucial, but largely unheralded, role in the current scientific enterprise. The history of science is, after all, a chronicle of our past experience with scientific hypothesizing and theorizing—with learning what sorts of hypotheses work and what sorts do not. Without the Bayesian analysis, one could say that the study of the history of science might have some (at least marginal) heuristic value for the scientist and philosopher of science, but on the Bayesian analysis, the data provided by the history of science constitute, *in addition*, an essential segment of the evidence relevant to the confirmation or disconfirmation of hypotheses. Philosophers of science and creative scientists ignore this fact at their peril.

[26] James D. Watson, *The Double Helix* (New York: New American Library, 1969). This book provides a fascinating account of the *discovery* of an important scientific hypothesis, and it illustrates many of the points I have been making. Perhaps if literary reviewers had had a clearer grasp of the distinction between the context of discovery and the context of justification they would have been less shocked at the emotions reported in the narrative.

PETER ACHINSTEIN

Inference to Scientific Laws

The topic of inference to scientific laws is one to which, I believe, both philosophers of science and historians of science can contribute to their mutual benefit. This is by no means self-evident, and indeed has been denied by philosophers as well as historians. There is even a view that there is no such topic at all to discuss. The view, held by proponents of the hypothetico-deductive (H-D) picture of science, is that there are no inferences *to* laws, only *from* them. The scientist does not infer a law from the data. He invents it, guesses it, imagines it, and then derives consequences from it which he tests. For example, Popper, one of the foremost proponents of this view, speaks of theories, including laws, as "free creations of our own minds, the result of an almost poetic intuition," and he rejects the idea that they are inferred in any way from observations.[1] Again, in a recent work, Hempel writes: "The transition from data to theory requires creative imagination. Scientific hypotheses are not *derived* from observed facts, but invented in order to account for them. They constitute guesses at the connections that might obtain between the phenomena under study. . . ."[2] The physicist Feynman agrees. He writes: "In general we look for a new law by the following process. First we guess it. Then we compute the consequences of the guess to see what would be implied if this law that we guessed is right. Then we compare the result of the computation to nature . . . to see if it works. If it disagrees with experiment it is wrong. In that simple statement is the key to science."[3]

According to the H-D view scientists do not make inferences to laws,

AUTHOR'S NOTE: This paper contains in abbreviated form some material from chapters 6 and 7 of my *Law and Explanation in Science*, to be published.

[1] Karl Popper, *Conjectures and Refutations* (London: Routledge and Kegan Paul, 1965), p. 192.

[2] Carl G. Hempel, *Philosophy of Natural Science* (Englewood Cliffs, N.J.: Prentice-Hall, 1966), p. 15.

[3] Richard Feynman, *The Character of Physical Law* (Cambridge, Mass.: MIT Press, 1967), p. 156.