1
Methodology

Introduction

This book is about scientific knowledge, particularly the concept of evidence. Its purpose is to explore scientific methodology in light of the obvious yet frequently neglected fact that belief is not an all-or-nothing matter, but is susceptible to varying degrees of intensity. More specifically, my main object is to exploit this fact to treat certain well-known puzzles in the philosophy of science, such as the problem of induction and the paradox of confirmation, as well as questions about *ad hoc* postulates, the tenability of realism, statistical testing, the relative merits of prediction and accommodation, a special quality of varied data, and the evidential value of further information. My second aim is to display the extent to which diverse elements of scientific method may be unified and justified by means of the concept of subjective probability. These two projects are intimately related. The probabilistic terms in which our evidential ideas will be formulated should promote clarity and accuracy, dispel confusion, and thereby facilitate the primary task. I should stress that this main goal is not to propound or defend a *theory* of the scientific method, either normative or descriptive, but rather to solve various paradoxical problems. I cannot now adequately describe the conception of philosophy which promotes this distinction and motivates my approach; but I think that some appreciation of that metaphilosophical perspective is needed to understand properly what is being attempted here and to pre-empt certain objections. I hope the following sketch will be better than nothing.

In a way, philosophy contains science and art. For there are philosophical research programmes whose methodology is scientific and others in which aesthetic standards prevail. Investigations into the semantics of natural languages, systematisations of basic ethical judgements, conceptual analyses – attempts to formulate necessary and sufficient conditions for S knows p, x causes y, and w refers to z – these typify scientific philosophy. Their object is a justified account of certain phenomena – a simple theory designed to accommodate the relevant data provided, usually, by intuition. On the other wing, we encounter the construction of metaphysical systems – symphonic flights, so far removed from testability that even the rubric of speculation would seem to distort their cognitive status. However, scientific and aesthetic philosophy do not exhaust the subject. What remain are the traditional puzzles and paradoxes, and an accumulating collection of modern ones: the problems of free-will, induction and scepticism, the paradoxes of Epimenides and Newcomb. Typically, we are confronted with apparently good reasons to accept both *p* and not-*p*: we have somehow blundered, become attached to some tempting misconception which must be located and exorcised. In such cases, the problem is not to formulate and defend a theory that will dictate which of the contradictory propositions is true; but rather to discover in ourselves the sources of our conflicting inclinations. This is Wittgenstein's idea of pure philosophy: no theories, only 'assembling reminders for a particular purpose': 'a battle against the bewitchment of our intelligence by means of language'. It should be emphasised that from this point of view there is no general reason to impugn the legitimacy or value of whatever else is done in the name of philosophy, and no need to fret about how that word should be used. True, there are bad projects – those resting upon mistaken presuppositions or governed by a foggy collection of adequacy conditions. These, through the confusion they produce, may engender the material for pure philosophy. Nonetheless, any theoretical project may be coherent, provided that there is a definite understanding of how successful accomplishment is to be recognised. And its results can sometimes illuminate a troublesome concept and constitute effective 'reminders' in the struggle to maintain an accurate view of it.

This sort of interaction is exemplified in what follows. The element of what I have called scientific philosophy will consist in the precise characterisation of a notion of degree of belief (designed to have the fruitful property that rational degrees of belief must conform to the probability calculus), and in the explication, within the framework of subjective probability, of various methodological concepts such as 'confirmation', '*ad hoc* postulate' and 'diverse evidence'. In order to deflect certain objections, the point of these constructions should be kept in mind. I do not claim that the foundations of subjective probability are absolutely secure, nor that my way is the only or the best way to capture our common-sense idea of degree of confidence. It may be that for other purposes such as psychology, decision theory, the history of science, or even an accurate description of scientific practice, we should prefer an explication which does not commit degrees of belief to precise numerical values. Different accounts of that notion could be appropriate and different analyses of the other relevant concepts may be called for. What is required here is just that the explications be sufficiently faithful and simple to enable a clear perception of the problems under discussion and to permit the confusion and misconceptions which produced them to be exposed and removed.

Aspects of the scientific method

The problems under consideration here stem from a number of very general and widely shared intuitions about evidence, which derive in turn from reflection upon scientific practice. I have divided these intuitions and problems into twelve topics, and what follows is a preliminary discussion of them. They comprise our subject matter; in later chapters each one is treated in more detail, and some explanations and answers are advanced.

(1) *Accommodation of data.* We are inclined to believe theories which make accurate predictions and accommodate experimental results, and disbelieve theories whose consequences are incompatible with our data. This should not require illustration. However, it is worth emphasising the well-known asymmetry between verification and falsification. If a theory is known to entail something false, it is conclusively refuted; but if it is known to entail something true the theory is *not* thereby taken to be conclusively confirmed, but, at best, merely supported. In other words, its correct predictions provide some evidence in favour of a theory, but do not establish it; whereas any mistaken predictions indicate that the theory should be abandoned or at least revised in part.

(2) *Statistical evidence.* The claim, about some experiment, that the probability of a certain outcome is x, is taken to be supported if roughly xn such outcomes, in a long sequence of n instances of the experiment, are obtained; and disconfirmed if the proportion of instances with that outcome differs substantially from x. In particular, if 100 consecutive tosses land heads up we begin to doubt that the coin is fair. In such cases the observed facts are neither entailed nor absolutely precluded by the hypothesis in question. But they are nonetheless of great evidential significance, and it would be desirable to have some explicit rationale for our practice and intuitions regarding the confirmation of empirical probabilistic hypotheses.

(3) *Severe tests*. Theories are tested, and confirmed to some extent if they pass; but they are well confirmed only if the tests are *severe*. For example, the use of highly accurate measuring instruments will tend to promote the severity of an experimental test, making it more difficult for the theory to pass and more impressive if it does. However, it remains to be seen what is meant in general by 'a severe test', and why survival through such things should give a special boost to the credibility of a theory.

(4) *Surprising predictions*. Particularly powerful support for a theory is conveyed by the verification of its relatively surprising predictions. In other words, a theory gets a lot of credit for predicting something quite unsuspected, or for explaining a bizarre and anomalous phenomenon; and it derives relatively little support from the prediction of something that we expected to occur anyway. So, for example, Einstein's special theory of relativity predicts that clocks moving with high velocity, near the speed of light, will run slowly compared with clocks at rest in our frame of reference, and that this so-called 'time dilation' is detectable. It also predicts that although the same effect is manifested by slowly moving clocks, its magnitude is too insubstantial to be measured, and so slowly moving clocks will *seem* to run at the same rate as stationary clocks. Now, both of these predictions – apparent time dilation in fast, but not in slowly moving clocks – have been verified; yet only the former is taken to provide us with striking confirmation of theory. Why is this? What is it to be surprising? And why is it that surprising, accurate predictions are of special evidential value?

(5) *Paradox of confirmation*. We feel that the hypothesis 'All ravens are black' is significantly confirmed by the observation that certain ravens are black; and not significantly confirmed by the observation that certain things which are not black are also not ravens. But it is peculiarly hard to come up with any rationale for this intuition; and for others like it. The general problem here is known as 'the paradox of confirmation'. It is natural to suppose that any scientific hypothesis of the form 'All *A*s are *B*' would be supported by evidence of the form '*k* is an *A* and *k* is a *B*'. This could well be the sort of principle we might propose as an example of a general canon of scientific methodology. In addition, it seems clear that if some datum supports, or confirms, or is evidence in favour of a scientific hypothesis, then it confirms every logically equivalent formulation of that hypothesis. The trouble – or paradox – is that these two very plausible principles lead to an extremely counterintuitive conclusion. For the first principle tells us that observation of a nonblack nonraven (say a white handkerchief) should be evidence in favour of the hypothesis 'All non-black things are nonravens'. Therefore, by the second principle, we are driven to the strange conclusion that the hypothesis 'All ravens are black' is confirmed by observation of a white handkerchief. This may be welcome on a rainy day, but it hardly squares with our intuitions about scientific methodology. I will try to explain and justify our intuitions, and show what is wrong with those plausible sounding principles which appear to be incompatible with them.

(6) *The 'grue' problem*. There is a further objection to the natural idea, already threatened by the above paradox of confirmation, that scientific reasoning may be codified by some such rule as

All sampled A s have been B

∴ Probably, all A s are B

Nelson Goodman (1955) has devised instances of this schema that constitute intuitively bad arguments. For example, define the predicate 'grue' as follows:

$$x \text{ is grue} \overset{\text{definition}}{\equiv} \begin{array}{l} x \text{ is sampled and green} \\ \text{or unsampled and blue} \end{array}$$

Now, the argument

All sampled emeralds have been grue

∴ Probably, all emeralds are grue

conforms to the alleged rule of induction. However, that reasoning is definitionally equivalent to

All sampled emeralds have been green

∴ Probably, all sampled emeralds are green

and unsampled emeralds are blue

which we would surely reject. Instead, it would be our inductive practice to infer from the given information that unsampled emeralds are green. Thus, the schema is not accurate in general, although certain instances (for example, A = emerald, B = green) do produce acceptable reasoning. Therefore, we are left with the question: how to specify the class of predicates (so-called projectible predicates) whose substitution in the inductive schema will yield acceptable arguments.

(7) *Simplicity*. Given two incompatible theories which both accommodate our data, we feel that the simpler one is more likely to be true. The need for some such intuition arises in the first place because it is always possible to find various incompatible theories, all of which fit the evidence that we have already accumulated. A typical case of this phenomenon is the possibility of drawing many curves through our set of data points. Thus, suppose a scientist wants to know the functional relationship between two parameters *X* and *Y* (for example the temperature and pressure of a fixed quantity of some gas, confined to a chamber whose volume is constant). Let us say he can vary the value of *X* and measure the corresponding value of *Y*. Now, suppose that in this way he has obtained, for six values of *X*, the corresponding values of *Y*, and plotted these points on graph paper. The points turn out to lie upon a straight line; nevertheless, many other functions are compatible with them, as shown in Fig. 1.

Another example, familiar from the grue problem, involves conflicting hypotheses, each of which could accommodate the evidence:

All sampled emeralds are green

In this case the alternatives are as shown in Fig. 2. These equally account for our observations of green emeralds, though they diverge concerning the colour of future, as yet unexamined, emeralds.

This phenomenon – the prevalence of competing observationally adequate hypotheses – gives rise to three questions:

(A) How do we choose between the alternatives? On what is our preference based? Simplicity? If so, how is simplicity to be recognised? The

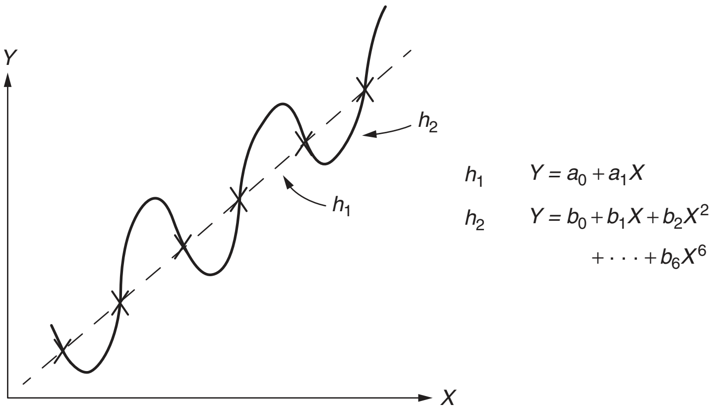

Fig. 1

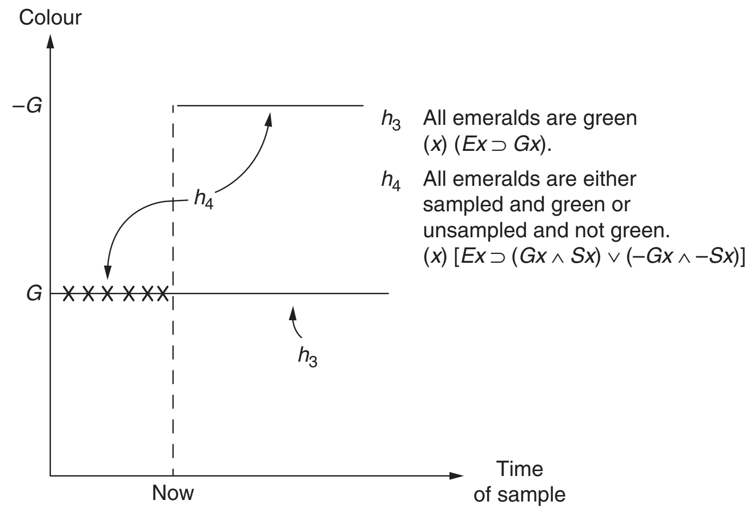

Fig. 2

grue problem is an element in this general problem of devising a description of our inductive practice.

(B) How is this measure of preferability to be combined with the knowledge that some theory fits certain data, to produce an overall assessment of its plausibility?

(C) What justifies our method of theory evaluation? Given an answer to question (A) – an account of the characteristics of a hypothesis which we take to recommend it above others which equally fit the data – what reason do we have to think that our procedure will tend to lead us towards the truth? In particular, why should we conclude, as we would given the evidence cited above, that probably all emeralds are green? This is the traditional problem of induction.

(8) *Ad hoc hypotheses*. The postulation of *ad hoc* hypotheses is thought to be somewhat disreputable. When an established theory is in danger of falsification by the discovery of facts it cannot explain, its proponents may patch up the theory in such a way as to reconcile it with the data. Such a manoeuvre is sometimes said to be *ad hoc* and scientists take a dim view of it. Consider, for example, the *ad hoc* claim that phlogiston has negative weight. This was proposed solely to rescue the theory that combustion of metals involves the escape of phlogiston, from the embarrassing observation that the ashes weigh more than the original metal. By reference to this and other examples we might hope to extract a definition of ad hocness, and a justification, based upon that characterisation, of the fact that we regard it as undesirable to postulate *ad hoc* hypotheses.

(9) *Diverse evidence*. We think that theories are better confirmed by a broad spectrum of different kinds of evidence than by a narrow repetitive set of data. Thus, intuitively, E_2 is better evidence for H than E_1 (see Fig. 3). Or, consider Snell's law of refraction: for any pair of media M_1 and M_2 there is a constant $\mu_{1,2}$ such that, for any i and r, if light is incident on the boundary between M_1 and M_2 at an angle i and is refracted at an angle r, then $\sin i/\sin r = \mu_{1,2}$ (Fig. 4). Some evidence for this claim is provided by 100 experiments in which r is measured for various different values of i, using the same pair of media M_1 and M_2, and found to be in accord with the claim. But stronger evidence is provided if the 100 experiments involve, not only a variation in i and r, but also a variation in the media used and a variation of the temperature at which the experiments are conducted. Again, what we want is a precise characterisation of breadth and an explanation of its evidential value.

(10) *Prediction versus accommodation.* There is an inclination to assume that a set of data provides better support for a theory if it was *predicted* by the theory before it was obtained, than if the theory was formulated after the data were obtained and was designed specifically to *accommodate* that information. This attitude is sometimes expressed in criticism of psycho-analytic theory. It is recognised that Freudians can concoct explanations afterwards of someone's behaviour, but it is felt that this is too easy – no

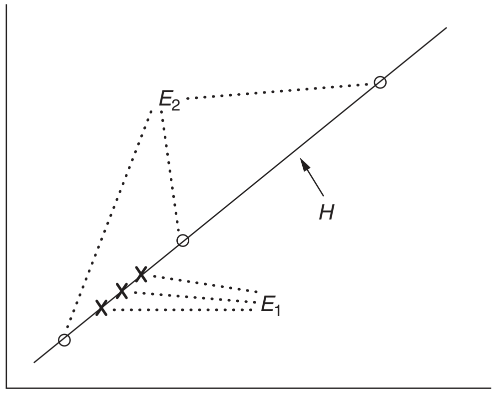

Fig. 3

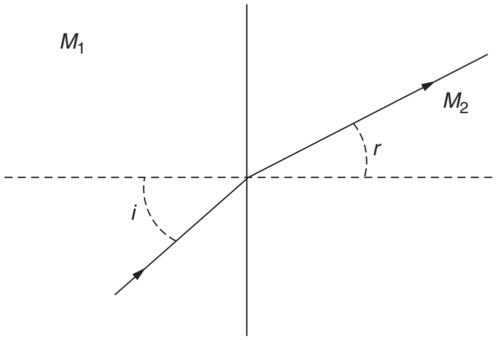

Fig. 4

real test of the theory. A real test would involve a definite prediction of what someone will do. Only then, and if the predictions are accurate, would we have good reason to believe the theory. We will explore the justification for this idea that accurate prediction has greater evidential value than the mere ability of a theory to accommodate, or subsume, or explain the data.

(11) *Desirability of further evidence*. When in doubt, it is prima facie desirable to acquire further evidence, for extra data tend to permit better assessments of the hypotheses under consideration. Imagine that two perfectly competent investigators are working independently of one another, but on the same problem – to decide between a set of alternative hypotheses in some area. Both scientists assess the plausibility of each hypothesis in the light of the data they have collected, and continually revise their assessments as new evidence accumulates. Now, suppose there is some time at which *A* has more information than *B*. *A* possesses all the facts known to *B*, but has some further results not known to *B*. And as a consequence of this difference they disagree about which is the most plausible hypothesis. In these circumstances we feel that *A* is in a superior position. His plausibility estimates, we suppose, are in some sense 'better' by virtue of the fact that they are based upon more information. Furthermore, this intuition seems to be intimately related to our thirst for data. Faced with a sufficiently interesting problem, we are disposed to go to a great deal of trouble to gather as much relevant evidence as possible. The problem here is to explain these banalities. Why is extra information so desirable? Why is it that plausibility assessments based on more data are better? And what precisely is the meaning of 'better' in this context?

(12) *Realism versus instrumentalism*. One might naturally think that the point of science is to discover those fundamental, general and hidden, features of the world, which are responsible for observable phenomena. This is the so–called realist conception of scientific theories: that they are proposed as true descriptions of reality. On the other hand, it has been argued that realism fails to appreciate the inexorable pattern of scientific revolutions, the inevitable eventual renunciation of our current theoretical constructions and, therefore, the unsuitability of truth as the proper object of scientific theorising. Recognition of these points leads to instrumentalism – the idea that theories are mere devices for the efficient organisation of data. But this view is also plagued with difficulties, of which the most serious is a need to distinguish sharply between the data statements, which correspond to facts in the world, and the theoretical formalism which is said to be designed merely to systematise them. I will suggest that the concept of subjective probability permits a sort of reconciliation between these views.

A taste of Bayesianism

As I said at the outset, one of our aims is a perspicuous representation of scientific methodology, and this project can be split into three intimately related components. First, we want clarification: precise formulations of the various intuitions and practices I have just described. Second, a systematisation of these elements: we want to be able to formulate a set of fundamental principles and show that various, apparently independent, features of scientific methodology may be derived from them and reflect no more than an implicit commitment to those principles. Third, we would like to justify the methods and assumptions which underlie the way in which science is conducted. If we have succeeded in the first two tasks of clarifying and systematising our practice, then this third component will involve an attempt to justify those basic principles of which all other features of scientific methodology are consequences.

The orthodox, and most widely held, theory of scientific methodology is Hempel's hypothetico-deductive model. It is so-called because it supposes that the method of scientific investigation involves the stages shown in Fig. 5.

I think that this is approximately right, as far as it goes. But it is silent on a wide range of important matters. It fails to account for the testing and adoption of statistical hypotheses – for they don't entail observable predictions. It provides no measure of the *degree* of confirmation conferred by