9

Philosophy of Science

DAVID PAPINEAU

The philosophy of science can usefully be divided into two broad areas. On the one hand is the epistemology of science, which deals with issues relating to the justification of claims to scientific knowledge. Philosophers working in this area investigate such questions as whether science ever uncovers permanent truths, whether objective decisions between competing theories are possible and whether the results of experiment are clouded by prior theoretical expectations. On the other hand are topics in the metaphysics of science, topics relating to philosophically puzzling features of the natural world described by science. Here philosophers ask such questions as whether all events are determined by prior causes, whether everything can be reduced to physics and whether there are purposes in nature. You can think of the difference between the epistemologists and the metaphysicians of science in this way. The epistemologists wonder whether we should believe what the scientists tell us. The metaphysicians worry about what the world is like, if the scientists are right. Readers will wish to consult chapters on EPISTEMOLOGY *(chapter 1),* METAPHYSICS *(chapter 2),* PHILOSOPHY OF MATHEMATICS *(chapter 11),* PHILOSOPHY OF SOCIAL SCIENCE *(chapter 12) and* PRAGMATISM *(chapter 36).*

1 The Epistemology of Science

1.1 The problem of induction

Much recent work in the epistemology of science is a response to the problem of induction. Induction is the process whereby scientists decide, on the basis of various observations or experiments, that some theory is true. At its simplest, chemists may note, say, that on a number of occasions samples of sodium heated on a Bunsen burner have glowed bright orange, and on this basis conclude that in general *all* heated sodium will glow bright orange. In more complicated cases, scientists may move from the results of a series of complex experiments to the conclusion that some fundamental physical principle is true. What all such inductive inferences have in common,

however, is that they start with particular premises about a *finite* number of past observations, yet end up with a general conclusion about how nature will *always* behave. And this is where the problem lies. For it is unclear how any finite amount of information about what has happened in the past can guarantee that a natural pattern will *continue* for all time.

After all, what rules out the possibility that the course of nature may change, and that the patterns we have observed so far turn out to be a poor guide to the future? Even if all heated sodium has glowed orange up till now, who is to say it will not start glowing blue sometime in the next century?

In this respect induction contrasts with deduction. In deductive inferences the premises guarantee the conclusion. For example, if you know that *Either this substance is sodium or it is potassium*, and then learn further that *It is not sodium*, you can conclude with certainty that *It is potassium*. The truth of the premises leaves no room for the conclusion to be anything but true. But in an inductive inference this does not hold. To take the simplest case, if you are told, for properties *A* and *B*, that *Each of the As observed so far has been B*, this does not guarantee that *All As, including future ones, are Bs*. It is perfectly possible that the former claim may be true, but the latter false.

The problem of induction seems to pose a threat to all scientific knowledge. All scientific discoveries worth their name are in the form of general principles. Galileo's law of free fall says that '*All* bodies fall with constant acceleration'; Newton's law of gravitation says that '*All* bodies attract each other in proportion to their masses and in inverse proportion to the square of the distance between them'; Avogadro's law says that '*All* gases at the same temperature and pressure contain the same number of molecules per unit volume'; and so on. The problem of induction calls the authority of all these laws into question. For if our evidence is simply that these laws have worked so far, then how can we be sure that they will not be disproved by future occurrences?

1.2   Popper's falsificationism

One influential response to the problem of induction is due to Sir Karl Popper (1902–94). In Popper's (1959a, 1963, 1972) view, science does not rest on induction in the first place. Popper denies that scientists start with observations, and then infer a general theory. Rather, they first put forward a theory, as an initially uncorroborated conjecture, and then compare its predictions with observations to see whether it stands up to test. If such tests prove negative, then the theory is experimentally falsified, and the scientists will seek some new alternative. If, on the other hand, the tests fit the theory, then scientists will continue to uphold it – not as proven truth, admittedly, but nevertheless as an undefeated conjecture.

If we look at science in this way, argues Popper, then we see that it does not need induction. According to Popper, the inferences which matter to science are refutations, which take some failed prediction as the premise, and conclude that the theory behind that prediction is false. These inferences are not inductive, but deductive. We see that some *A* is not-*B*, and conclude that it is not the case that *All As are Bs*. There is no room here for the premise to be true and the conclusion false. If we discover that some body falls with

increasing acceleration (say because it falls from a great height, and so is subject to a greater gravitational force as it nears the earth), then we know for sure that all bodies do not fall with constant acceleration. The point here is that it is much easier to disprove theories than to prove them. A single contrary example suffices for a conclusive disproof, but no number of supporting examples will constitute a conclusive proof.

So, according to Popper, science is a sequence of conjectures and refutations. Scientific theories are put forward as hypotheses, and they are replaced by new hypotheses when they are falsified. However, if scientific theories are always conjectural in this way, then what makes science better than astrology, or spirit worship, or any other form of unwarranted superstition? A non-Popperian would answer this question by saying that real science proves its claims on the basis of observational evidence, whereas superstition is nothing but guesswork. But on Popper's account, even scientific theories are guesswork – for they cannot be proved by the observations, but are themselves merely undefeated conjectures.

Popper calls this the 'problem of demarcation': what is the difference between science and other forms of belief? His answer is that science, unlike superstition, is at least *falsifiable*, even if it is not provable (Popper 1959a: ch. 2). Scientific theories are framed in precise terms, and so issue in definite predictions. For example, Newton's laws tell us exactly where certain planets will appear at certain times. And this means that if such predictions fail, we can be sure that the theory behind them is false. By contrast, belief systems like astrology are irredeemably vague, in a way which prevents their ever being shown definitely wrong. Astrology may predict that Scorpios will prosper in their personal relationships on Thursdays, but when faced with a Scorpio whose spouse walks out on a Thursday, defenders of astrology are likely to respond that the end of the marriage was probably for the best, all things considered. Because of this, nothing will ever force astrologists to admit their theory is wrong. The theory is phrased in such imprecise terms that no actual observations can possibly falsify it.

Popper himself uses the criterion of *falsifiability* to distinguish genuine science, not just from traditional belief systems like astrology and spirit worship, but also from Marxism, psychoanalysis and various other modern disciplines that he denigrates as 'pseudo-sciences'. According to Popper, the central claims of these theories are as unfalsifiable as those of astrology. Marxists predict that proletarian revolutions will be successful whenever capitalist regimes have been sufficiently weakened by their internal contradictions. But when faced with unsuccessful proletarian revolutions, they simply respond that the contradictions in those particular capitalist regimes have not yet weakened them sufficiently. Similarly, psychoanalytic theorists will claim that all adult neuroses are due to childhood traumas, but when faced by troubled adults with apparently undisturbed childhoods, they will say that those adults must nevertheless have undergone private psychological traumas when young. For Popper, such ploys are the antithesis of scientific seriousness. Genuine scientists will say beforehand what observational discoveries would make them change their minds, and will abandon their theories if these discoveries are made. But Marxists and psychoanalytic theorists frame their theories in such a way, argues Popper, that no possible observations need ever make them adjust their thinking.

1.3 The failings of falsificationism

At first sight Popper seems to offer an extremely attractive account of science. He explains its superiority over other forms of belief, while at the same time apparently freeing it from any problematic dependence on induction. Certainly his writings have struck a chord within the scientific community. Popper is one of the few philosophers ever to have become a Fellow of the Royal Society, an honour usually reserved for eminent scientists.

In the philosophical world, however, Popper's views are more controversial. This is because many philosophers feel that his account of science signally fails to solve the problem with which he begins, namely, the problem of induction (for example, see Ayer 1956: 71–5; Worrall 1989). The central objection to his position is that it only accounts for negative scientific knowledge, as opposed to positive knowledge. Popper points out that a single counter-example can show us that a scientific theory is wrong. But he says nothing about what can show us that a scientific theory is right. Yet it is positive knowledge of this latter kind that makes science important. We can cure diseases and send people to the moon because we know that certain causes *do* always have certain results, not because we know that they *do not*. Useful scientific knowledge comes in the form 'All *A*s are *B*s', not 'It's false that all *A*s are *B*s'. Since Popper only accounts for the latter kind of knowledge, he seems to leave out what is most interesting and important about science.

Popper's usual answer to this objection is that he is concerned with the logic of pure scientific research, not with practical questions about technological applications. Scientific research requires only that we formulate falsifiable conjectures, and reject them if we discover counter-examples. The further question of whether technologists should *believe* those conjectures, and rely on their predictions when, say, they administer some drug or build a dam, Popper regards as an essentially practical issue, and as such not part of the analysis of rational scientific practice.

But this will not do. After all, Popper claims to have solved the problem of induction. But the problem of induction is essentially the problem of how we can base judgements about the future on evidence about the past. In insisting that scientific theories are just conjectures, and that therefore we have no rational basis for *believing* their predictions, Popper is simply denying that we can make rational judgements about the future.

Consider these two predictions: (1) when I jump from this tenth-floor window I shall crash painfully into the ground; (2) when I jump from the window I will float like a feather to a gentle landing. Intuitively, it is more rational to believe (1), which assumes that the future will be like the past, than (2), which does not. But Popper, since he rejects induction, is committed to the view that past evidence does not make any beliefs about the future more rational than any others, and therefore that believing (2) is no less rational than believing (1).

Something has gone wrong. *Of course* believing (1) is more rational than believing (2). In saying this, I do not want to deny that there is a *problem* of induction. Indeed it is precisely *because* believing (1) is more rational than believing (2) that induction is problematic. Everybody, Popper aside, can see that believing (1) is more rational than believing (2). The problem is then to explain *why* believing (1) is more rational than believing (2), in the face of the apparent invalidity of induction. So Popper's denial of

the rational superiority of (1) over (2) is not so much a *solution* to the problem of induction, but simply a refusal to recognize the problem in the first place.

Even if it fails to deal with induction, Popper's philosophy of science does have some strengths as a description of pure scientific research. For it is certainly true that many scientific theories start life as conjectures, in just the way Popper describes. When Einstein's general theory of relativity was first proposed, for example, very few scientists actually *believed* it. Instead they regarded it as an interesting hypothesis, and were *curious* to see whether it was true. At this initial stage of a theory's life, Popper's recommendations make eminent sense. Obviously, if you are curious to see whether a theory is true, the next step is to put it to the observational test. And for this purpose it is important that the theory is framed in precise enough terms for scientists to work out what it implies about the observable world – that is, in precise enough terms for it to be falsifiable. And of course if the new theory does get falsified, then scientists will reject it and seek some alternative, whereas if its predictions are borne out, then scientists will continue to investigate it.

Where Popper's philosophy of science goes wrong, however, is in holding that scientific theories never progress beyond the level of conjecture. As I have just suggested, theories are often mere conjectures when they are first put forward, and they may remain conjectures as the initial evidence first comes in. But in many cases the accumulation of evidence in favour of a theory will move it beyond the status of conjecture to that of established truth. The general theory of relativity started life as a conjecture, and many scientists still regarded it as hypothetical even after Sir Arthur Eddington's famous initial observations in 1919 of light apparently bending near the sun. But by now this initial evidence has been supplemented with evidence in the form of gravitational red-shifts, time-dilation and black holes, and it would be an eccentric scientist who nowadays regarded the general theory as less than firmly established.

Such examples can be multiplied. The heliocentric theory of the solar system, the theory of evolution by natural selection and the theory of continental drift all started life as intriguing conjectures, with little evidence to favour them over their competitors. But in the period since they were first proposed these theories have all accumulated a great wealth of supporting evidence. It is only those philosophers who have been bemused by the problem of induction who view these theories as being no better than initial hypotheses. Everybody else who is acquainted with the evidence has no doubt that these theories are proven truths.

1.4   Bayesianism

If we insist, against Popper, that we are fully entitled to believe at least some scientific theories on the basis of past evidence, then we are committed to finding some solution to the problem of induction. One currently popular account of the legitimacy of induction is found within Bayesianism, named after Thomas Bayes (*c*.1701–61) (Horwich 1982; Howson and Urbach 1989).

Bayesians are philosophers who hold that our beliefs, including our beliefs in scientific theories, come in degrees. Thus, for example, I can believe to degree 0.5 that it will rain today, in the sense that I think there is a 50 per cent likelihood of rain today. Similarly, I might attach a 0.1 degree of belief to the theory that the strong nuclear and

electro-weak forces are the same force – I think it unlikely, but allow that there is a one-in-ten possibility it may turn out true.

As these examples indicate, Bayesians think of degrees of belief as the extent to which you subjectively take something to be PROBABLE (pp. 167–8). Accordingly, they argue that your degrees of belief ought to satisfy the axioms of the probability calculus. (See the box below for the Dutch Book Argument for this thesis.) It is important to realize, however, that while Bayesians think of degrees of belief as probabilities in this mathematical sense, they still think of them as *subjective* probabilities. In particular, they allow that it can be perfectly rational for different people to attach *different* subjective probabilities to the same proposition – you can believe that it will rain today to degree 0.2, while I believe this to degree 0.5. What rationality does require, according to the Bayesians, is only that if you have a subjective probability of 0.2 for rain, then you must have one of 0.8 for its not raining, while if I have 0.5 for rain, then I must have 0.5 for its not raining. That is, both of us must accord, in our different ways, with the theorem of the probability calculus that Prob(p) = 1 − Prob(not-p).

At first sight, this element of subjectivity might seem to disqualify Bayesianism as a possible basis for scientific rationality. If we are all free to attach whatever degrees of belief we like to scientific theories, provided only that we are faithful to the structure of the probability calculus, then what is to stop each of us from supporting different theories, depending only on individual fads or prejudices? But Bayesians have an answer; namely, that it does not matter what prejudices you start with, as long as you revise your degrees of belief in a rational way.

Bayesians derive their account of how to revise degrees of belief, as well as their name, from Bayes's theorem, originally proved by Thomas Bayes in a paper published in 1763. Bayes's theorem states:

$$\text{Prob}(H/E) = \text{Prob}(H) \times \text{Prob}(E/H)/\text{Prob}(E).$$

The simple proof of this theorem is given in the box. But the philosophical significance of the theorem is that it suggests a certain procedure for revising your degrees of belief in response to new evidence. Suppose that H is some hypothesis, and E is some newly discovered evidence. Then Bayesians argue that, when you discover E, you should adjust your degree of belief in H in line with the right-hand side of the above equation: that is, you should increase it to the extent that you think E is likely given H, but unlikely otherwise. In other words, if E is in itself very surprising (like light bending in the vicinity of the sun) but at the same time just what you would expect given your theory H (the general theory of relativity), then E should make you increase your degree of belief in H a great deal. On the other hand, if E is no more likely given H than it would be on any other theory, then observing E provides no extra support for H. The movement of the tides, for example, is no great argument for general relativity, even though it is predicted by it, since it is also predicted by the alternative Newtonian theory of gravitation.

Note in particular that this strategy for updating degrees of belief in response to evidence can be applied to inductive inferences. Consider the special case where H is some universal generalization – all bodies fall with constant acceleration, say – and the evidence E is that some particular falling body has been observed to accelerate constantly.

If this observation was something you did not expect at all, then Bayesianism tells you that you should increase your degree of belief in Galileo's law significantly, for it is just what Galileo's law predicts. Of course, once you have seen a number of such observations, and become reasonably convinced of Galileo's law, then you will cease to find new instances surprising, and to that extent will cease to increase your degree of belief in the law. But that is as it should be. Once you are reasonably convinced of a law, then there is indeed little point in gathering further supporting instances, and so it is to the credit of Bayesianism that it explains this.

The Bayesian account of how to revise degrees of belief seems to make good sense. In addition, it promises a solution to the problem of induction, since it implies that positive instances give us reason to believe scientific generalizations.

There are, however, problems facing this account. For a start, a number of philosophers have queried whether Bayes's theorem, which after all is little more than an arithmetical truth, can constrain what degrees of belief we adopt in the future (see the box below). And even if we put this relatively technical issue to one side, it is unclear how far the Bayesian account really answers the worry raised above, that the subjectivity of degrees of belief will allow different scientists to commit themselves arbitrarily to different theories. The Bayesian answer to this worry was that Bayes's theorem will at least constrain these different scientists to revise their degrees of belief in response to the evidence in similar ways. But, even so, it still seems possible that the scientists will remain on different tracks, if they start at different places. If two scientists are free to attach different prior degrees of belief in Galileo's law, and both update those degrees of belief according to Bayes's theorem when they learn the evidence, will they not still end up with different posterior degrees of beliefs?

The standard answer to the objection is to appeal to *convergence* of opinion. The idea is that, given enough evidence, everybody will *eventually* end up in the same place, even if they have different starting-points. There are a number of theorems of probability theory showing that, within limits, differences in initial probabilities will be 'washed out', in the sense that sufficient evidence and Bayesian updating will lead to effectively identical final degrees of belief. So in the end, argue Bayesians, it does not matter if you start with a high or low degree of belief in Galileo's law – for after 1,000 observations of constantly falling bodies you will end up believing it to a degree close to 1 anyway.

However, interesting as these results are, they do not satisfactorily answer the fundamental philosophical questions about inductive reasoning. For they do not work for all possible initial degrees of belief. Rather, they assume that the scientists at issue, while differing among themselves, all draw their initial degrees of belief from a certain range. While this range includes all the initial degrees of belief that seem at all intuitively plausible, there are nevertheless other possible initial degrees of belief that are consistent with the axioms of probability, but which will not lead to eventual convergence. So, for example, the Bayesians do not in fact explain what is wrong with people who never end up believing Galileo's law because they are always convinced that the course of nature is going to change tomorrow. Of course, Bayesians are right to regard such people as irrational. But they do not explain why they are irrational. So they fail to show why all thinkers must end up with the same attitude to scientific theories. And in particular they fail to solve the problem of induction, since they do not show why all rational thinkers must expect the future to be like the past.

Bayesianism

The Dutch Book Argument

The axioms of probability require that

(1) $0 \leq \text{Prob}(P) \leq 1$, for any proposition P
(2) $\text{Prob}(P) = 1$, if P is a necessary truth
(3) $\text{Prob}(P) = 0$, if P is impossible
(4) $\text{Prob}(P \text{ or } Q) = \text{Prob}(P) + \text{Prob}(Q)$, if P and Q are mutually exclusive.

Bayesians appeal to the Dutch Book Argument to show why subjective degrees of belief should conform to these axioms. Imagine that your degrees of belief did not so conform. You believe proposition P to degree y, say, and yet do *not* believe not-P to degree $1 - y$. (You thus violate the conjunction of axioms (2) and (4), because P or not-P is a necessary truth.) Then it will be possible for somebody to induce you to make bets on P and not-P in such a way that you will lose whatever happens. A set of bets that guarantee that you will lose whatever happens is called a 'Dutch book'. The undesirability of such a set of bets thus provides an argument that any rational person's subjective degrees of belief should satisfy the axioms of the probability calculus.

Bayes's Theorem

The conditional probability of P given Q – Prob(P/Q) – is defined as Prob(P and Q)/Prob(Q). Intuitively, Prob(P/Q) signifies the probability of P on the assumption that Q is true. It immediately follows from this definition that

$$\text{Prob}(H/E) = \text{Prob}(H) \times \text{Prob}(E/H)/\text{Prob}(E)$$

This is Bayes's theorem. As you can see, it says that the conditional probability Prob(H/E) of some hypothesis H given evidence E is greater than Prob(H) to the extent that E is improbable in itself, but probable given H.

Bayesian Updating

Bayesians recommend that if you observe some evidence E, then you should *revise* your degree of belief in H, and set your new $\text{Prob}_{t'}(H)$ equal to your *previous* conditional degree of belief in H given E, $\text{Prob}_t(H/E)$, where t is the time before you learn E, and t' after. Bayes's theorem, applied to your subjective probabilities at t, then indicates that this will *increase* your degree of belief in H to the extent that you previously thought E to be subjectively improbable in itself, but subjectively probable given H.

　　This Bayesian *recommendation*, that you revise your degree of belief in H by setting it equal to your old conditional degree of belief in H given E, should be distinguished from Bayes's *theorem*. Bayes's theorem is a trivial consequence of the definition of conditional probability, and constrains your degrees of belief at *a given time*. The Bayesian recommendation, by contrast, specifies how your degrees of belief should change *over time*. Bayes's theorem is uncontentious, but it is a matter of active controversy whether there is any satisfactory way of defending the Bayesian recommendation (Hacking 1967; Teller 1973).

1.5   Instrumentalism versus realism

At this stage let us leave the problem of induction for a while and turn to a different difficulty facing scientific knowledge. Much of science consists of claims about *unobservable* entities like viruses, radio waves, electrons and quarks. But if these entities are unobservable, how are scientists supposed to have found out about them? If they cannot see or touch them, does it not follow that their claims about them are at best speculative guesses, rather than firm knowledge?

It is worth distinguishing this problem of unobservability from the problem of induction. Both problems can be viewed as difficulties facing *theoretical* knowledge in science. But whereas the problem of induction arises because scientific theories make *general* claims, the problem of unobservability is due to our *lack of sensory access* to the subject matter of many scientific theories. (So the problem of induction arises for general claims even if they are not about unobservables, such as 'All sodium burns bright orange'. Conversely, the problem of unobservability arises for claims about unobservables even if they are not general, such as 'One free electron is attached to this oil drop'. In this section and the next, however, it will be convenient to use the term 'theory' specifically for claims about unobservables, rather than for general claims of any kind.)

There are two general lines of response to the problem of unobservability. On the one hand are *realists*, who think that the problem can be solved. Realists argue that the observable facts provide good indirect evidence for the existence of unobservable entities, and so conclude that scientific theories can be regarded as accurate descriptions of the unobservable world. On the other hand are *instrumentalists*, who hold that we are in no position to make firm judgements about imperceptible mechanisms. Instrumentalists allow that theories about such mechanisms may be useful 'instruments' for simplifying our calculations and generating predictions. But they argue that these theories are no more true descriptions of the world than the 'theory' that all the matter in a stone is concentrated at its centre of mass (which is also an extremely useful assumption for doing certain calculations, but clearly false).

Earlier this century instrumentalists used to argue that we should not even *interpret* theoretical claims literally, on the grounds that we cannot so much as meaningfully talk about entities we have never directly experienced. But nowadays this kind of semantic instrumentalism is out of favour. Contemporary instrumentalists allow that scientists can meaningfully *postulate*, say, that matter is made of tiny atoms containing nuclei orbited by electrons. But they then take a SCEPTICAL (pp. 45–56) attitude to such postulates, saying that we have no entitlement to *believe* them (as opposed to using them as an instrument for calculations).

An initial line of argument open to realism is to identify some feature of scientific practice and then argue that instrumentalism is unable to account for it. One aspect of scientific practice invoked in this connection has been the unification of different kinds of theories in pursuit of a single 'theory of everything' (Friedman 1984); other features of science appealed to by realists have included the use of theories to *explain* observable phenomena (Boyd 1980), and the reliance on theories to make novel *predictions* (Smart 1963). For, so the realist argues, these aspects of scientific practice only make sense on the assumption that scientific theories are *true* descriptions of reality. After all, says the realist, if theories are simply convenient calculating devices, then why

expect different theories to be unifiable into one consistent story? Unification is clearly desirable if our theories all aim to contribute to the overall truth, but there seems no parallel reason why a bunch of instruments should be unifiable into one big 'instrument of everything'. And similarly, the realist will argue, there seems no reason to expect a mere calculating instrument, as opposed to a true description of an underlying reality, to yield a genuine explanation of some past occurrence, or a reliable prediction of a future one.

However, this form of argument tends to be inconclusive. There are two possible lines of response open to instrumentalists. They can offer an instrumentalist account of the relevant feature of scientific practice. Alternatively, they can deny that this feature really is part of scientific practice in the first place. As an example of the first response, they could argue that the unification of science is motivated, not by the pursuit of one underlying truth, but simply by the desirability of having a single all-purpose calculating instrument rather than a rag-bag of different instruments for different problems. The second kind of response would be to deny that unification is essential to science to start with. Thus Nancy Cartwright argues that science really *is* a rag-bag of different instruments. She maintains that scientists faced with a given kind of problem will standardly deploy simplifying techniques and rules of thumb which owe nothing to general theory, but which have shown themselves to deliver the right answer to the kind of problem at hand (Cartwright 1983).

Similar responses can be made by instrumentalists to the arguments from explanation and prediction. Instrumentalists can either retort that there is no reason why the status of theories as calculating instruments should preclude them from giving rise to predictions and explanations; or they can query whether scientific theories really do add to our ability to predict and explain to start with. Not all these lines of response are equally convincing. But between them they give instrumentalism plenty of room to counter the initial realist challenge.

1.6 Theory, observation and incommensurability

A different line of argument against instrumentalism focuses on the distinction between what is observable and what is not. This distinction is crucial to instrumentalism, in that instrumentalists argue that claims about observable phenomena are unproblematic, but claims about unobservables are not. However, a number of writers have queried this distinction, arguing that observation reports are not essentially different from claims about unobservables, since they too depend on theoretical assumptions about the underlying structure of reality. Norwood Hanson (1958) has argued, for example, that scientists before and after Copernicus *saw* different things when they looked at the Sun: whereas pre-Copernicans regarded the Earth as stationary and so saw the Sun revolving round it, post-Copernican scientists saw the Sun as stationary and the Earth as rotating. Similarly, Hanson (1963) argues that the photographic plate which looks like a squiggly mess to a lay observer is seen as displaying a well-defined electron–positron pair by an experienced particle physicist. Examples like these undermine the distinction between what is observable and what is not, since they show that even judgements made in immediate response to sensory stimulation are influenced by fallible theories about reality.
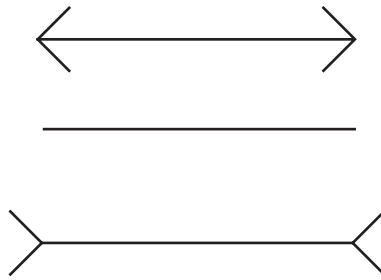
Figure 9.1 The Müller–Lyer. Although all three lines are the same length, the top line, with inward-pointing arrowheads, appears to be shorter, and the bottom line, with outward-pointing arrowheads, appears to be longer, than the 'neutral' middle line.

Nor is the point restricted to recherché observations of astronomical bodies or sub-atomic particles. Even immediate perceptual judgements about the colour, shape and size of medium-sized physical objects can be shown to depend on theoretical assumptions implicit in our visual systems. Perhaps the best-known illustration is the Müller–Lyer illusion (see figure 9.1), which shows how our visual system uses complex assumptions about the normal causes of certain kinds of retinal patterns to draw conclusions about the geometry of physical figures. And analogous illusions can be used to demonstrate the presence of other theoretical presuppositions in our visual and other sensory systems.

As I said, in the first instance the unclarity of the observable–unobservable distinction counts against instrumentalism rather than realism. After all, it is instrumentalism, not realism, which *needs* the distinction, since instrumentalism says that we should be sceptical about unobservable claims, but not observable ones, whereas realism is happy to regard both kinds of claims as belief-worthy, so does not mind if they cannot be sharply distinguished.

However, there is another way of responding to doubts about the theory–observation distinction. For note that the arguments against the observable–unobservable distinction do not in fact vindicate the realist belief-worthiness of claims about *unobservables*; rather, they attack realism from the bottom up, and undermine the belief-worthiness of claims about *observables*, by showing that even observational claims depend on fallible theoretical assumptions. Obviously, if there is no observable–unobservable distinction, then all scientific claims are in the same boat. But on reflection it seems that the boat they all end up in is the instrumentalist boat of sceptical disbelief, not the realist one of general faith in science.

A number of influential recent philosophers of science, most prominently T. S. Kuhn (1962) and Paul Feyerabend (1976), have embraced this conclusion wholeheartedly, and maintained that no judgements made within science, not even observational judgements, can claim the authority of established truth. Rather, they argue, once scientists have embraced a theory about the essential nature of their subject matter, such as geocentrism, or Newtonian dynamics, or the wave theory of light, they will interpret all observational judgements in the light of that theory, and so will never be forced to

recognize the kind of negative observational evidence that might show them that their theory is mistaken. Kuhn and Feyerabend independently lit on the term *incommensurable* to express the view that there is no common yardstick, in the form of theory-independent observation judgements, which can be used to decide objectively on the worth of scientific theories. Instead, they argue, decisions on scientific theories are never due to objective observational evidence, but are always relative to the presuppositions, interests and social milieux of the scientists involved.

Kuhn's and Feyerabend's blanket relativism has provoked much discussion among philosophers of science, but won few whole-hearted converts. Much of the discussion has focused on the status of observations. Most philosophers of science are prepared to accept that all observational judgements in some sense presuppose some element of theory. But many balk at the conclusion that observations therefore never have any independent authority to decide scientific questions. After all, they point out, most simple observations, such as that a pointer is adjacent to a mark on a dial, presuppose at most a minimal amount of theory, about rigid bodies, say, and about basic local geometry. Since such minimal theories are themselves rarely at issue in serious scientific debates, this minimal amount of theory-dependence provides no reason why observations of pointer readings should not be used to settle scientific disputes. If a scientific theory about the behaviour of gases, say, predicts that a pointer will be at a certain place on a dial, and it is observed not to be, then this decides against the theory about gases. It is not to the point to respond that, in taking the pointer reading at face value, we are making assumptions about rigid bodies and local geometry. For nothing in the debate about gases provides any reason to doubt these assumptions. And this of course is why scientists take such pains to work out what their theories imply about things like pointer readings – since observations of pointer readings do not depend on anything contentious, they will weigh with all sides in the scientific debate.

So, despite the arguments of Kuhn and Feyerabend, nearly all philosophers are realists about pointer readings and similar observable phenomena. But this still leaves us with the original disagreement between realism and instrumentalism about less directly observable entities. For, even if claims about pointer readings are uncontroversially belief-worthy, instrumentalists can still argue that theories about viruses, atoms and gravitational waves are nothing more than useful fictions for making calculations.

The realist response, as I said, is that the observable facts provide good indirect evidence for these theoretical entities, even if we cannot observe them directly. However, there are two strong lines of argument that instrumentalists can use to cast doubt on this suggestion. In the next two sections I shall discuss 'the underdetermination of theory by evidence' and 'the pessimistic meta-induction from past falsity'.

1.7   The underdetermination of theory by evidence

The argument from underdetermination asserts that, given any theory about unobservables that fits the observable facts, there will be other incompatible theories that fit the same facts. And so, the argument concludes, we are never in a position to know that any one of these theories is the truth.

Why should we accept that there is always more than one theory that fits any set of observable facts? One popular argument for this conclusion stems from the

'Duhem–Quine thesis'. According to this thesis, any particular scientific theory can always be defended in the face of contrary observations by adjusting auxiliary hypotheses. For example, when the Newtonian theory of gravitation was threatened by observations of anomalous movements by the planet Mercury, it could always be defended by postulating a hitherto unobserved planet, say, or an inhomogeneous mass distribution in the Sun. This general strategy for defending theories against contrary evidence seems to imply that the adherents of competing theories will always be able to maintain their respective positions in the face of any actual observational data.

Another argument for underdetermination starts, not with competing theories, but with some given theory. Suppose that all the predictions of some particular theory are accurate. We can construct a 'de-Ockhamized' version of this theory (reversing William of Ockham's 'razor' which prescribes that 'entities are not to be multiplied beyond necessity'), by postulating some unnecessarily complicated unobservable mechanism which nevertheless yields a new theory with precisely the same observational consequences as the original one.

Both of these lines of reasoning can be used to argue that more than one theory about unobservables will always fit any given set of observational data. Does this make realism about unobservables untenable? Many philosophers conclude that it does. But this is too quick. For we should recognize that there is nothing in the arguments for alternative underdetermined theories to show that these alternative theories will always be *equally well-supported by the data*. What the arguments show is that different theories will always be *consistent* with the data. But they do not rule out the possibility that, among these alternative theories, one is vastly more plausible than the others, and for that reason should be believed to be true. After all, 'flat earthers' can make their view consistent with the evidence from geography, astronomy and satellite photographs, by constructing far-fetched stories about conspiracies to hide the truth, the effects of empty space on cameras, and so on. But this does not show that we need take their flat-earthism seriously. Similarly, even though Newtonian gravitational theory can in principle be made consistent with all the contrary evidence, this is no reason not to believe general relativity. Nor is our ability to 'cook up' a de-Ockhamized version of general relativity a reason to stop believing the standard version unencumbered with unnecessary entities.

Nevertheless, as I said, many contemporary philosophers of science do move directly from the premise that different theories are consistent with the observational evidence to the conclusion that none of them can be regarded as the truth. This is because many of them address this issue from an essentially Popperian perspective. For if you follow Popper in rejecting induction, then you will not believe that evidence ever provides positive support for any theory, except in the back-handed sense that the evidence can fail to falsify it. Accordingly you will think that all theories that have not been falsified are on a par, and in particular that any two theories that are both consistent with the evidence are equally well-supported by it.

So the arguments for underdetermination do present a problem to Popperians, since Popperians have no obvious basis for discriminating among different theories consistent with the data. But, as I pointed out above, these need not worry those of us who diverge from Popper in thinking unfalsified theories can be better or worse supported by evidence, for we can simply respond to the underdetermination arguments by

observing that some underdetermined theories are better supported by the evidence than others.

Now that we have returned to Popper, it is worth noting that the Duhem–Quine argument also raises a more specific problem for Popperians. Recall that Popper's overall philosophy raised the 'problem of demarcation', the problem of how to distinguish science from other kinds of conjecture. Popper's answer was that science, unlike astrology, or Marxism and psychoanalytic theory, is falsifiable. But the Duhem–Quine argument shows that even such eminently scientific theories as Newtonian physics are not falsifiable in any straightforward sense, since they can always save themselves in the face of failed predictions by adjusting auxiliary hypotheses.

Not only does this cast doubt on Popper's dismissal of Marxism and psychoanalysis as unscientific, but it seems to undermine his whole solution to the demarcation problem. If such paradigmatic scientific theories as Newtonian physics are not falsifiable, then it can scarcely be falsifiability that distinguishes science from non-science. Still, this is Popper's problem, not ours (see Harding 1975). If we do not reject induction, then we do not have a problem of demarcation. For we can simply say that what distinguishes successful scientific theories from non-science is that the observational evidence gives us inductive reason to regard scientific theories as true.

The arguments in the latter part of this section have presupposed that a certain kind of inductive argument is legitimate. The kind of inductive argument relevant to underdetermination is not simple 'enumerative' induction, from observed *A*s being *B*s to 'All *A*s are *B*s', but rather inferences from any collection of observational data to the most plausible theory about unobservables that is consistent with that data. But these are species of the same genus; indeed, enumerative inductions can themselves be interpreted as treating 'All *A*s are *B*s' as the most plausible extrapolation consistent with the observed *A*s being *B*s. My attitude to this more general category of inductive inferences remains the same as my attitude to enumerative induction, which I outlined earlier. We do not yet have an explanation of why inductive inferences are legitimate, and to that extent we still face a problem of induction. But it is silly to try to solve that problem by denying that inductive inferences are ever legitimate. And that is why the underdetermination of theory by data does not constitute a good argument for instrumentalism. For to assume that we are never entitled to believe a theory, if there are others consistent with the same data, is simply to assume the illegitimacy of induction.

The Underdetermination of Theory by Observational Data (UTD)

There are two arguments for the UTD. The first is based on the Duhem–Quine thesis, originally formulated by the French philosopher and historian Pierre Duhem (1861–1916) and later revived by the American logician W. V. O. Quine (b. 1908). Duhem (1951) and Quine (1951) point out that a scientific theory T does not normally imply predictions P on its own, but only in conjunction with auxiliary hypotheses H.

$$T \& H \Rightarrow P$$

So when P is falsified by observation, this does not refute T, but only the conjunction of T & H.

not-P $\Rightarrow$ not-(T & H)

So T can be retained, and indeed still explain P, provided we replace H by some alternative, H', such that

T & H' $\Rightarrow$ not-P.

This yields the Duhem–Quine thesis: any theoretical claim T can consistently be retained in the face of contrary evidence, by making adjustments elsewhere in our system of beliefs. The UTD follows quickly. Imagine two competing theories T_1 and T_2. Whatever evidence accumulates, versions of T_1 and T_2, conjoined with greatly revised auxiliary hypotheses if necessary, will both survive, consistent with that evidence, but incompatible with each other.

The other argument, first put forward by physicists like Henri Poincaré (1854–1912) and Ernst Mach (1838–1916) at the turn of the twentieth century, has a different starting-point. Imagine that T_1 is the complete truth about physical reality, and that it implies observational facts O. Then we can always construct some 'de-Ockhamized' T_2 which postulates more complicated unobservable mechanisms but makes just the same observational predictions O. (Glymour 1980: ch. 5.)

For example, suppose we start with standard assumptions about the location of bodies in space-time and about the forces acting on them. A de-Ockhamized theory might then postulate that all bodies, including all measuring instruments, are accelerating by 1ft/sec.^{2} in a given direction, and then add just the extra forces required to explain this. This theory would clearly have exactly the same observational consequences as the original one, even though it contradicted it at the unobservable level.

To bring out the difference between the two arguments for UTD, note that the Duhem–Quine argument does not specify exactly which overall theories we will end up with, since it leaves open how T_1's and T_2's auxiliary hypotheses may need to be revised; the de-Ockhamization argument, by contrast, actually specifies T_1 and T_2 in full detail, including auxiliary hypotheses. In compensation, the Duhem–Quine argument promises us alternative theories whatever observational evidence may turn up in the future; whereas the de-Ockhamization argument assumes that all future observations are as T_1 predicts.

1.8 The pessimistic meta-induction

I turn now to the other argument against realism. This argument takes as its premise the fact that past scientific theories have generally turned out to be false, and then moves inductively to the pessimistic conclusion that our current theories are no doubt false too. (This is called a 'meta-induction' because its subject matter is not the natural world, but scientific theories about the natural world.)

There are plenty of familiar examples to support this argument. Newton's theory of space and time, the phlogiston theory of combustion, and the theory that atoms are indivisible were all at one time widely accepted scientific theories, but have since been recognized to be false. So does it not seem likely, the pessimistic induction concludes, that all our current theories are false, and that we should therefore take an instrumentalist rather than a realist attitude to them? (See Laudan 1981.)

This is an important and powerful argument, but it would be too quick to conclude that it discredits realism completely. It is important that the tendency to falsity is much more common in some areas of science than others. Thus it is relatively normal for theories to be overturned in cosmology, say, or fundamental particle physics, or the study of primate evolution. By contrast, theories of the molecular composition of different chemical compounds (such as that water is made of hydrogen and oxygen), or the causes of infectious diseases (chickenpox is due to a herpes virus), or the nature of everyday physical phenomena (heat is molecular motion), are characteristically retained once they are accepted.

Nor need we regard this differential success rate of different kinds of theories as some kind of accident. Rather, it is the result of the necessary evidence being more easily available in some areas of science than others. Paleoanthropologists want to know how many hominid species were present on earth 3 million years ago. But their evidence consists of a few pieces of teeth and bone. So it is scarcely surprising that discoveries of new fossil sites will often lead them to change their views. The same point applies on a larger scale in cosmology and particle physics. Scientists in these areas want to answer very general questions about the very small and the very distant. But their evidence derives from the limited range of technological instruments they have devised to probe these realms. So, once more, it is scarcely surprising that their theories should remain at the level of tentative hypotheses. By contrast, in those areas where adequate evidence is available, such as chemistry and medicine, there is no corresponding barrier to science moving beyond tentative hypotheses to firm conclusions.

The moral is that realism is more defensible for some areas of science than others. In some scientific subjects firm evidence is available, and entitles us to view certain theories, like the theory that water is composed of H_2O molecules, as the literal truth about reality. In other areas the evidence is fragmentary and inconclusive, and then we do better to regard the best-supported theories, such as the theory that quarks and leptons are the ultimate building blocks of matter, as useful instruments which accommodate the existing data, make interesting predictions, and suggest further lines for research.

At first sight this might look like a victory for instrumentalism over realism. For did not instrumentalists always accept that we should be realists about *observable* things, and only urge instrumentalism for uncertain theories about unobservable phenomena? But our current position draws the line in a different place. Instrumentalism, as originally defined, takes it for granted that everything *unobservable* is inaccessible, and that all theories about unobservables are therefore uncertain. By contrast, the position we have arrived at places no special weight on the distinction between what is observable and what is not. In particular, it argues that the pessimistic meta-induction fails to show that falsity is the natural fate of all theories about unobservables, but only that there is a line *within* the category of theories about unobservables, between those theories that can be expected to turn out false and those whose claims to truth are secure. So our current position is not a dogmatic instrumentalism about all unobservables, but merely the uncontentious view that we should be instrumentalists about that sub-class of theories which are not supported by adequate evidence.

1.9 Naturalized epistemology of science

In the last decade or so a number of philosophers of science have turned to a *naturalized* approach to scientific knowledge (Kitcher 1992). In place of traditional attempts to establish criteria for scientific theory-choice by *a priori* philosophical investigation, the naturalized approach regards science itself as a subject for *a posteriori* empirical investigation. Accordingly, naturalized epistemologists look to the history, sociology and psychology of science, rather than to first principles, to identify criteria for the acceptability of scientific theories.

One apparent difficulty facing this kind of naturalized epistemology of science is that it is unclear how empirical investigation can ever yield anything more than *descriptive* information about how scientists actually operate. Yet any epistemology of science worth its name ought also to have a *normative* content – it ought to *prescribe* how scientists should reason, as well as describe how they do reason. David HUME (chapter 31) first pointed out that there is a logical gap between 'is' and 'ought'. A naturalized epistemology based on the empirical study of science seems fated to remain on the wrong side of this gap.

However, there is room for naturalized epistemologists to reply to this charge. They can agree that the empirical study of science cannot by itself yield prescriptions about how science ought to be done. But empirical study can still be *relevant* to such prescriptions. Suppose it is agreed that *technological fertility*, in the sense of generating technological advances, is a virtue in a scientific theory. Then the history, sociology and psychology of science might be able to show us that certain kinds of research strategies are effective at developing technologically fertile theories. More generally, given any agreed theoretical end Y, empirical study can show that research strategy X is an effective *means* to that end. The empirical study of science can thus yield the *hypothetical prescription* that, if you want Y, then you *ought* to adopt means X. It is this kind of hypothetical prescription that naturalized philosophers of science seek to establish: they look to the history, sociology and psychology of science to show us that scientists who *choose* theories on grounds X will in general *achieve* theories with characteristic Y.

Can the naturalized study of science tell us which research strategies are an effective means to theoretical *truth*? Different naturalized philosophers of science give different answers to this question. Many are suspicious of the idea of theoretical truth, and instead prefer to stick to the study of how to achieve more practical ends like technological fertility, simplicity and predictive accuracy. However, there seems no good reason for this restriction. There is nothing obviously incoherent in the idea of looking to the empirical study of science to tell us which research strategies have proved a good way of developing true theories. Indeed, the discussion of the 'pessimistic meta-induction' in the previous section amounted to the sketch of just such an investigation, in that it appealed to the history of science to decide whether or not the standard procedures of scientific theory-choice succeed in identifying true theories. It is not difficult to imagine more detailed and specific studies of this kind of issue.

Let me now return briefly to the issue with which I began, namely, the problem of induction. It is possible that the naturalized study of how to get at the scientific truth will enable us to make headway with this problem. For an empirical investigation into

science might be able to show us that a certain kind of *inductive* inference is in general a reliable guide to scientific truth. And this would then provide a kind of vindication of that inductive method (see Papineau 1993: ch. 5).

It is true that this kind of defence of induction will inevitably involve an element of circularity. For when we infer that certain kinds of induction are *in general* a reliable guide to truth, on the basis of evidence from the history of science, this will itself be an inductive inference. It is a matter of some delicacy, however, whether this circularity is vicious.

Defenders of this naturalized defence of induction will point out that, from their point of view, a legitimate criterion of theory-choice need not be an *a priori* guide to truth, but only an empirically certifiable one. Given this, the original argument against induction, that it is not logically valid, will not worry naturalized philosophers of science. Induction may not provide any *a priori* guarantee for its conclusions; but from the naturalized point of view, this does not show that induction is in any way illegitimate, since it leaves it open that induction may be an empirically reliable guide to the truth. And if there is nothing to show that induction is illegitimate, naturalized philosophers of science can then argue, why should we not use it to investigate the worth of inductive inferences? Maybe this is less satisfying a defence of induction than we might originally have hoped for. But perhaps it is defence enough.

2   The Metaphysics of Science

2.1   Causation

Many issues in the metaphysics of science hinge on the notion of *causation*. This notion is as important in science as it is in everyday thinking, and much scientific theorizing is concerned specifically to identify the *causes* of various phenomena. However, there is little philosophical agreement on what it means to say that one event is the cause of another.

Modern discussion of causation starts with David Hume, who argued that causation is simply a matter of CONSTANT CONJUNCTION (p. 720). According to Hume (1978), one event causes another if and only if events of the type to which the first event belongs regularly occur in conjunction with events of the type to which the second event belongs. This formulation, however, leaves a number of questions open. Firstly, there is the problem of distinguishing genuine *causal laws* from *accidental regularities*. Not all regularities are sufficiently lawlike to underpin causal relationships. Being a screw in my desk could well be constantly conjoined with being made of copper, without its being true that these screws are made of copper because they are in my desk. Secondly, the idea of constant conjunction does not give a *direction* to causation. Causes need to be distinguished from effects. But knowing that A-type events are constantly conjoined with B-type events does not tell us which of A and B is the cause and which the effect, since constant conjunction is itself a symmetric relation. Thirdly, there is a problem about *probabilistic causation*. When we say that causes and effects are constantly conjoined, do we mean that the effects are always found with the causes, or is it enough that the causes make the effects probable?

Many philosophers of science during the past century have preferred to talk about *explanation* rather than causation. According to the covering-law model of explanation, something is explained if it can be deduced from premises which include one or more laws. As applied to the explanation of particular events, this implies that one particular event can be explained if it is linked by a law to some other particular event. However, while they are often treated as separate theories, the covering-law account of explanation is at bottom little more than a variant of Hume's constant conjunction account of causation. This affinity shows up in the fact that the covering-law account faces essentially the same difficulties as Hume: (1) in appealing to deductions from 'laws', it needs to explain the difference between genuine laws and accidentally true regularities; (2) it omits the requisite directionality, in that it does not tell us why we should not 'explain' causes by effects, as well as effects by causes; after all, it is as easy to deduce the height of a flagpole from the length of its shadow and the laws of optics, as to deduce the length of the shadow from the height of the pole and the same laws; (3) are the laws invoked in explanation required to be exceptionless and deterministic, or is it acceptable, say, to appeal to the merely probabilistic fact that smoking makes cancer more likely, in explaining why some particular person developed cancer?

In what follows I shall discuss these three problems in order (treating them as problems that arise equally both for the analysis of causation and the analysis of explanation). After that I shall consider some further issues in the metaphysics of science.

The Covering-Law Model of Explanation

According to this model (originally proposed by Hempel and Oppenheim (1948) and further elaborated in Hempel (1965)) one statement (the *explanandum*) is explained by other statements (the *explanans*) if and only if the explanans contains one or more *laws*, and the explanandum can be *deduced* from the explanans. In the simplest case, where the explanandum is some particular statement to the effect that some individual *a* has property E, we might therefore have:

a has C
For all x, if x has C, then x has E
a has E

For example, we might deduce that a piece of litmus paper turned red, from the law that all litmus paper placed in acid turns red, together with the prior condition that this piece of litmus paper was in fact placed in acid. The model can accommodate more complicated explanations of particular events, and can also allow explanations of laws themselves, as when we deduce Kepler's law that all planets move in ellipses, say, from Newton's law of universal gravitation and his laws of motion.

As applied to the explanation of particular events, the covering-law model implies a symmetry between *explanation* and *prediction*. For the information that, according to the model, suffices for the explanation of some known event should also enable us to predict that event if we did not yet know of it. Many critics have fastened on this implication of the model, however, and pointed out that we can often predict when we do not have enough information to explain (as when we predict the height of the flagpole from

its shadow) and can often explain when we could not have predicted (as when we explain X's cancer on the basis of X's smoking).

These examples suggest that genuine explanations of particular events need to cite genuine causes, and that the reason the covering-law model runs into counter-examples is that it adds nothing to the inadequate constant conjunction analysis of causation, except that it substitutes the term 'law' for 'constant conjunction'. To get a satisfactory account of explanation we need, firstly, to recognize that explanations of particular events must mention causes, and, secondly, to improve on the constant conjunction analysis of causation.

There is a variant of the covering-law model which allows non-deterministic explanation as well as deterministic ones. This is termed the 'inductive–statistical (I–S)' model, by contrast with the original 'deductive–nomological (D–N)' model. An example would be:

> *a* drinks 10 units of alcohol per diem
> For *p* per cent of *x*s, if *x* drinks 10 units of alcohol per diem, *x* has a damaged liver
> *a* has a damaged liver

Here the explanandum cannot be *deduced* from the explanans, but only follows with an *inductive* probability of *p*; and the inference appeals to a *statistical* regularity, rather than an exceptionless *nomological* generalization. In Hempel's original version of this model, it was required that the probability of the explanandum be *high*. A better requirement, however, as explained in the section on probabilistic causation below, is that the particular facts in the explanans need only make the probability of the explanandum *higher* than it would otherwise have been.

2.2 Laws and accidents

There are two general strategies for distinguishing laws from accidentally true generalizations. The first stands by Hume's idea that causal connections are mere constant conjunctions, and then seeks to explain why some constant conjunctions are better than others. That is, this first strategy accepts the principle that causation involves nothing more than certain events always happening together with certain others, and then seeks to explain why some such patterns – the 'laws' – matter more than others – the 'accidents'. The second strategy, by contrast, rejects the Humean presupposition that causation involves nothing more than happenstantial co-occurrence, and instead postulates a relationship of 'necessitation', a kind of 'cement', which links events that are connected by law, but not those events (like being a screw in my desk and being made of copper) that are only accidentally conjoined.

There are a number of versions of the first Humean strategy. The most successful, originally proposed by F. R. Ramsey (1903–30), and later revived by David Lewis (1973), holds that laws are those true generalizations that can be fitted into an ideal system of knowledge. The thought here is that the laws are those patterns that are somehow explicable in terms of basic science, either as fundamental principles themselves, or as consequences of those principles, while accidents, although true, have no such explanation. Thus, 'All water at standard pressure boils at 100°C' is a consequence

of the laws governing molecular bonding; but the fact that 'All the screws in my desk are copper' is not part of the deductive structure of any satisfactory science. Ramsey neatly encapsulated this idea by saying that laws are 'consequences of those propositions which we should take as axioms if we knew everything and organized it as simply as possible in a deductive system' (Ramsey 1978: 130).

Advocates of the alternative non-Humean strategy object that the difference between laws and accidents is not a *linguistic* matter of deductive systematization, but rather a *metaphysical* contrast between the kind of links they report. They argue that there is a link in nature between *being at 100°C* and *boiling*, but not between *being in my desk* and *being made of copper*, and that this is nothing to do with how the description of this link may fit into theories. According to D. M. Armstrong (1983), the most prominent defender of this view, the real difference between laws and accidents is simply that laws report relationships of natural *necessitation*, while accidents only report that two types of events *happen* to occur together.

Armstrong's view may seem intuitively plausible, but it is arguable that the notion of necessitation simply restates the problem, rather than solving it. Armstrong says that necessitation involves something more than constant conjunction: if two events are related by necessitation, then it follows that they are constantly conjoined; but two events can be constantly conjoined without being related by necessitation, as when the constant conjunction is just a matter of accident. So necessitation is a stronger relationship than constant conjunction. However, Armstrong and other defenders of this view say very little about what this extra strength amounts to, except that it distinguishes laws from accidents. Armstrong's critics argue that a satisfactory account of laws ought to cast more light than this on the nature of laws.

2.3   *The direction of causation*

Hume said that the earlier of two causally related events is always the cause, and the *later* the effect. However, there are a number of objections to using the earlier–later 'arrow of time' to analyse the directional 'arrow of causation'. For a start, it seems in principle possible that some causes and effects could be simultaneous. More seriously, the idea that time is directed from 'earlier' to 'later' itself stands in need of philosophical explanation – and one of the most popular explanations is that the idea of 'movement' from earlier to later depends on the fact that cause–effect pairs always have a given orientation in time. However, if we adopt such a 'causal theory of the arrow of time', and explain 'earlier' as the direction in which causes lie, and 'later' as the direction of effects, then we will clearly need to find some account of the direction of causation which does not itself assume the direction of time.

A number of such accounts have been proposed. David Lewis (1979) has argued that the asymmetry of causation derives from an 'asymmetry of overdetermination'. The overdetermination of present events by past events – consider a person who dies after simultaneously being shot and struck by lightning – is a very rare occurrence. By contrast, the multiple 'overdetermination' of present events by future events is absolutely normal. This is because the future, unlike the past, will always contain multiple traces of any present event. To use Lewis's example, when the president presses the red button in the White House, the future effects do not only include the dispatch

of nuclear missiles, but also his fingerprint on the button, his trembling, the further depletion of his gin bottle, the recording of the button's click on tape, the emission of light waves bearing the image of his action through the window, the warming of the wire from the passage of the signal current, and so on, and on, and on.

Lewis relates this asymmetry of overdetermination to the asymmetry of causation as follows. If we suppose the cause of a given effect to have been absent, then this implies the effect would have been absent too, since (apart from freaks like the lightning–shooting case) there will not be any other causes left to 'fix' the effect. By contrast, if we suppose a given effect of some cause to have been absent, this does not imply the cause would have been absent, for there are still all the other traces left to 'fix' the cause. Lewis argues that these counterfactual considerations suffice to show why causes are different from effects.

Other philosophers appeal to a probabilistic variant of Lewis's asymmetry. Following Reichenbach (1956), they note that the different causes of any given type of effect are normally probabilistically independent of each other; by contrast, the different effects of any given type of cause are normally probabilistically correlated. For example, both obesity and high excitement can cause heart attacks, but this does not imply that fat people are more likely to get excited than thin ones; on the other hand, the fact that both lung cancer and nicotine-stained fingers can result from smoking does imply that lung cancer is more likely among people with nicotine-stained fingers. So this account distinguishes effects from causes by the fact that the former, but not the latter, are probabilistically dependent on each other.

2.4   Probabilistic causation

The just-mentioned probabilistic account of the *direction* of causation is normally formulated as part of a more general *theory of probabilistic causation*. Until relatively recently philosophers assumed that the world fundamentally conforms to deterministic laws, and that probabilistic dependencies between types of events, such as that between smoking and lung cancer, merely reflected our ignorance of the full causes. The rise of quantum mechanics, however, has persuaded most philosophers that determinism is false, and that some events, like the decay of a radium atom, happen purely as a matter of chance. A particular radium atom may decay, but on another occasion an identical atom in identical circumstances might well not decay.

Accordingly, a number of philosophers of science have put forward models of causation which require only that causes probabilify, rather than determine, their effects. The earliest such model was the 'inductive–statistical' version of the covering-law model of explanation (Hempel 1965). Unlike deterministic 'deductive–nomological' explanations, such inductive–statistical explanations required only that prior conditions and laws imply a *high* probability for the event to be explained, not that this event will certainly happen. However, even this seems too strong a requirement for probabilistic causation. After all, smoking unequivocally causes lung cancer, but even heavy smokers do not have a *high* probability of lung cancer, in the sense of a probability close to 1. Rather, their smoking increases their probability of lung cancer, not to a high figure, but merely from a low to a less low figure, but still well below 50 per cent. So more recent models of probabilistic causation simply require that causes

should *increase* the probability of their effects, not that they should give them a high probability (Salmon 1971).

This kind of model needs to guard against the possibility that the probabilistic association between putative cause and putative effect may be *spurious*, like the probabilistic association between barometers falling and subsequent rain. Such associations are not due to a causal connection between barometer movements and rain, but rather to both of these being joint effects of a *common cause*, namely, in our example, falls in atmospheric pressure. The obvious response to this difficulty is to say that we have a cause–effect relationship between A and B if and only if A increases the probability of B, and this association is not due to some common cause C. However, this is obviously incomplete as an analysis of causation, since it uses the notion of (common) *cause* in explaining causation.

It would solve this problem if we could analyse the notion of common cause in probabilistic terms. It seems to be a mark of common causes that they probabilistically '*screen off*' the associations between their joint effects, in the sense that, if we consider cases where the common cause is present and where it is absent separately, then the probabilistic association between the joint effects will disappear. For example, if it is given that the atmospheric pressure has fallen, then a falling barometer does not make it any more likely that it will rain; and similarly, if the atmospheric pressure *has not* fallen, a faulty falling reading on a barometer is no probable indicator of impending rain. (Numerically, if C is a common cause, and A and B its joint effects, we will find that A and B are associated – Prob(B/A) < Prob(B) – but that C and its absence render A irrelevant to B – Prob(B/A & C) = Prob(B/C) and Prob(B/A & not-C) = Prob(B/not-C).) It remains a matter of some debate, however, whether this characteristic probabilistic structure of common causes is enough to allow a complete explanation of causation in probabilistic terms, or whether further non-probabilistic considerations need to be introduced.

2.5   Probability

Philosophical interest in probabilistic causation has led to a resurgence of interest in the philosophy of probability itself. Probability raises philosophical puzzles in its own right, quite apart from its connection with causation. What exactly is the 'probability' of a given event? The only part of the answer that is uncontroversial is that probabilities are quantities that satisfy the axioms of the probability calculus I specified earlier when discussing Bayesianism. But this leaves plenty of room for alternative philosophical views, for there are a number of different ways of interpreting these axioms.

One interpretation is the *subjective* theory of probability, which equates probabilities with subjective degrees of belief. This is the interpretation assumed by Bayesian confirmation theory. Most philosophers are happy to agree that subjective degrees of belief exist, and that the Dutch Book Argument (see the above box on Bayesianism) shows why they ought to conform to the axioms of probability. But many, if not all, philosophers argue that we need a theory of objective probability in addition to this subjective account.

One possible objective interpretation is the *frequency* theory, originally put forward by Richard von Mises (1957). According to this theory, the probability of a given kind

of result is the number of times this result occurs, divided by the total number of occasions on which it might have occurred. So, for example, the probability of heads on a coin toss is the proportion of heads in some wider class of coin tosses.

This theory, however, faces a number of difficulties. For a start, it has problems in dealing with 'single-case probabilities'. Consider a particular coin toss. We can consider it as a member of the class of all coin tosses, or of all tosses of coins with that particular shape, or of all tosses made in just that way, or so on. However, these different 'reference classes' may well display different frequencies of heads. Yet intuitively it seems that there ought to be a unique value for the probability of heads on a particular toss of a particular coin. Perhaps this difficulty can be dealt with by specifying that the single-case probability should equal the relative frequency in the reference class of all tosses that are *similar in all relevant respects* to the particular toss in question. But there remain difficulties about which respects should count as 'relevant' in this sense.

In addition, there is the problem that many of these more specific reference classes will only be finite in extent. Coins with a certain distinctive shape may only be tossed in some given way ten times in the whole history of the universe. Yet the probability of heads on these tosses is unlikely to be equal to the relative frequency in the ten tosses, for luck may well yield a disproportionately high, or low, number of heads in ten tosses. Because of this, frequency theorists standardly appeal, not to actual reference classes, but to hypothetical infinite sequences, and equate the probability with the *limit* that the relative frequency *would* tend to *if* the relevant kind of trial were repeated an infinite number of times. Critics of the frequency theory object that this reliance on hypothetical infinite reference sequences makes probabilities inadmissibly abstract.

Because of these difficulties, many contemporary philosophers of probability have adopted the '*propensity*' theory of probability in place of the frequency theory. The earliest version of this theory, proposed by Popper (1959b), simply modified the frequency theory by specifying that only those relative frequencies generated by repeated trials on a given 'experimental set-up' should count as genuine probabilities. This arguably deals with the problem of single-case probabilities, but it still leaves us with hypothetical reference classes. To avoid this, later versions of the propensity theory do not define probabilities in terms of frequencies at all, but simply take probabilities to be primitive propensities of particular situations to produce given results.

This kind of propensity theory does not seek to define objective probabilities in terms of frequencies, but in effect simply takes single-case probabilities as primitive (see Mellor 1971). But it can still recognize a connection between probabilities and frequencies. For, as long as propensities are assumed to obey the axioms of the probability calculus (though this assumption itself merits some debate), it will follow that, in a sufficiently long sequence of independent trials in each of which the propensity to produce B is p, the overall *propensity* for the observed frequency of B to differ by more than a given amount from p can be made arbitrarily small, in accord with the Law of Large Numbers. (Note how the italicized second use of 'propensity' in this claim prevents it serving as a definition of propensities in terms of frequencies.)

Both the frequency theory and the propensity theory have their strengths. The frequency theory has the virtue of offering an explicit definition of probability, where the propensity theory takes probabilities as primitive. On the other hand, the propensity

theory has no need to assume hypothetical reference sequences, whereas these are essential to the frequency theory.

It may seem that the frequency theory, because it offers an explicit definition, is better able than the propensity theory to explain how we find out about probabilities. But this is an illusion. The trouble is that the frequency theory's explicit definition is in terms of frequencies in INFINITE SEQUENCES (p. 355). But our evidence is always in the form of frequencies in *finite* samples. So the problem of explaining how we can move from frequencies in finite samples to knowledge of probabilities is as much a problem for the frequency theory as for the propensity theory. (There are various suggestions about how to solve this 'problem of statistical inference', none of them universally agreed. My present point is merely that this problem of statistical inference arises in just the same way for both frequency and propensity theorists.)

In the face of continued debate about the interpretation of objective probability, some philosophers have turned to physics, and in particular to the notion of probability used in quantum mechanics, to resolve the issue. Unfortunately, quantum mechanics is no less philosophically controversial than the notion of probability. There are different philosophical interpretations of the formal theory of quantum mechanics, each of which involves different understandings of probability. Because of this it seems likely that philosophical disputes about probability will continue until there is an agreed interpretation of quantum mechanics.

The Interpretation of Quantum Mechanics

Modern quantum mechanics says that the state of any given system of microscopic particles is fully characterized by its 'wave function'. However, instead of specifying the exact positions and velocities of the particles, as is done in classical mechanics, this 'wave function' only specifies *probabilities* of the particles displaying certain values of position, and velocity, *if* appropriate measurements are made. *Schrödinger's equation* then specifies how this wave function evolves smoothly and deterministically over time, analogously to the way that Newton's laws of motion specify how the positions and velocities of macroscopic objects evolve over time – except that Schrödinger's equation again only describes changes in probabilities, not exact values.

On the orthodox interpretation of quantum mechanics, quantum probabilities change into actualities only when 'measurements' are made. If you measure the position of a particle, say, then its position assumes a definite value, even though nothing before the measurement determined exactly what this value would be.

There is something puzzling about this, however, since any overall system of measured particle and measuring instrument is itself just another system of microscopic particles, which might therefore be expected to evolve smoothly according to Schrödinger's equation, rather than to jump suddenly to some definite value for position. To account for this, the orthodox interpretation says that in addition to the normal Schrödinger evolution, there is a special kind of change which occurs in 'measurements', when the wave function suddenly 'collapses' to yield a definite value for the measured quantity.

The 'measurement problem' is the problem of explaining exactly when, and why, these collapses occur. The story of 'Schrödinger's cat' makes the difficulty graphic. Suppose that some unfortunate cat is sitting next to a poison dispenser which is wired

up to emit cyanide gas if an electron emitted from some source turns up on the right half of some position-registering plate, but not if the electron turns up on the left half. The basic quantum mechanical description of this situation says that it is both possible that the electron will turn up on the right half of the plate and that it will turn up on the left half, and therefore both possible that the poison is emitted and that it is not, and therefore both possible that the cat is alive and that it is dead. One of these possibilities only becomes actual when the wave function of the whole system collapses. But when does that happen? When the electron is emitted? When it reaches the plate? When the cat dies or not? Or only when a human being looks at the cat to see how it is faring?

There seems no principled way to decide between these answers. Because of this, many philosophers reject the orthodox view that physical systems are completely characterized by their wave functions, and conjecture that, in addition to the variables quantum mechanics recognizes, there are various 'hidden variables' which always specify exact positions and velocities for all physical particles. It is difficult, however, for such *hidden variable theories* to reproduce the surprising phenomena predicted by quantum mechanics, without postulating mysterious mechanisms that seem inconsistent with other parts of physics.

A more radical response to the measurement problem is to deny that the wave function ever does collapse, and somehow to make sense of the idea that reality contains both a live cat and a dead cat. This 'many-worlds' interpretation of quantum mechanics flies in the face of common sense, but its theoretical attractions are leading an increasing number of philosophers to take it seriously.

2.6   Teleology

We normally explain some particular fact by citing its *cause*: for example, we explain why some water freezes by noting that its temperature fell below 0°C. There are cases, however, where we seem to explain items by citing their *effects* instead. In particular, this kind of explanation is common in biology. We often explain some biological trait by showing how it is useful to the organism in question: for instance, the explanation of the polar bear's white fur is that it camouflages it; the explanation of human sweating is that it lowers body temperature, and so on. Similar explanations are also sometimes offered in anthropology and sociology.

Until fairly recently most philosophers of science took such functional or teleological explanations at face value, as an alternative to causal explanation, in which items are explained, not by their causal antecedents, but by showing how they contribute to the well-being of some larger system. Carl Hempel's covering-law model of explanation embodied an influential version of this attitude. According to Hempel, causal explanations and functional explanations are simply two different ways of exemplifying the covering-law model: the only difference is that in causal explanations the *explaining fact* (lower temperature) temporally precedes the *explained* fact (freezing), whereas in functional explanations it is explained fact (white fur) that comes temporally before the consequence (camouflage) which explains it.

Most contemporary philosophers of science, however, take a different view, and argue that all explanations of particular facts are really causal, and that functional explanations, despite appearances, are really a *subspecies* of causal explanations. On

this view, the reference to future facts in functional explanations is merely apparent, and such explanations really refer to past causes. In the biological case, these past causes will be the evolutionary histories that led to the natural selection of the biological trait in question. Thus the functional explanation of the polar bear's colour should be understood as referring us to the fact that their *past* camouflaging led to the natural selection of their whiteness, and not to the fact that they may be camouflaged in the *future*. Similarly, any acceptable functional explanations in anthropology or sociology should be understood as referring us *back* in time to the conscious intentions or unconscious selection processes which caused the facts to be explained (see Wright 1973; Neander 1991). (There remains the terminological matter of whether functional explanations understood in this way ought still to be called 'teleological'. Traditional usage reserves the term 'teleology' for distinctively non-causal explanations in terms of future results. But most contemporary philosophers are happy to describe disguised causal explanations that make implicit reference to selection mechanisms as 'teleological'.)

2.7   Theoretical reduction

Another philosophical question about subject matters like biology is whether they can be reduced to lower-level (in the sense of ontologically more basic) sciences like chemistry and physics. Obviously, this is an issue that arises not just for biology, but also for such other 'special' natural sciences as geology and meteorology, and also for such human sciences as psychology, sociology and anthropology.

One science is said to 'reduce' to another if its categories can be defined in terms of the categories of the latter, and its laws explained by the laws of the latter. *Reductionists* argue that all sciences form a hierarchy in which the higher can always be reduced to the lower. Thus, for example, biology might be reduced to physiology, physiology to chemistry, and eventually chemistry to physics.

Reductionism can be viewed either historically or metaphysically. The historical question is whether science characteristically progresses by later theories reducing earlier ones. The metaphysical question is whether the different areas of science describe different realities, or just the one physical reality described at different levels of detail. Though often run together, these are different questions.

Taken as a general thesis, historical reductionism is false. Recall the earlier discussion of the 'pessimistic meta-induction from past falsity'. This involved the claim that new theories characteristically show their predecessors to be false. To the extent that this claim is true, historical reductionism is false: for a new theory can scarcely explain why an earlier theory was true, if it shows it is false.

In the earlier discussion I argued that there are some areas of science, like molecular biology and medical science, to which the pessimistic meta-induction does not apply. If this is right, then we can expect that in these areas new theories will indeed normally reduce old ones. But I did not dispute that there are other areas of science, like cosmology and fundamental particle physics, in which the normal fate of old theories is to be thrown out. It follows that we must reject historical reductionism, understood as the thesis that all science proceeds by new theories reducing old ones.

This does not mean, however, that metaphysical reductionism is false. Even if science proceeds towards the overall truth by fits and starts, there may be general reasons for expecting that this overall truth, when eventually reached, will reduce to physical truth.

One possible such argument stems from the *causal interaction* between the phenomena discussed in the special sciences and physical phenomena. Biological, geological and meteorological events all unquestionably have physical effects. It is difficult to see how they could do this unless they are made of physical components.

It is doubtful, however, whether this suffices to establish full-scale reductionism, as opposed to the weaker thesis (sometimes called 'token-identity') according to which each *particular* higher-level event is identical with some *particular* physical event. Thus, for example, it might be true that one animal's aggressive behaviour can be equated with a given sequence of physical movements, and another animal's aggressive behaviour can be equated with another sequence of physical movements, without there being any uniform way of defining 'aggressive behaviour', for all animals, in terms of physical movements. The case-by-case token-identity will explain how each instance of aggressive behaviour can have physical effects, like causing intruding animals to move away. But without any uniform definition of 'aggressive behaviour' in terms of physical movements there is no question of reducing ethology (the science of behaviour) to physics, and so no question of explaining ethological laws by physical laws. Instead, the laws of ethology and other special sciences will be *sui generis*, identifying patterns whose instances vary in their physical make-up, and which therefore cannot possibly be explained in terms of physical laws alone (see Fodor 1974).

Acknowledgements

Further Reading

For an introduction to the problem of induction, and Popper's solution, see Popper (1959a), especially chapter 1. The problem and Popper's solution are further discussed in O'Hear (1989). There are two excellent introductions to Bayesian philosophy of science: Horwich (1982) and Howson and Urbach (1989).

The best modern defence of instrumentalism is Van Fraassen (1980). Churchland and Hooker (1985) offers a good collection of essays on the realism–instrumentalism debate. Kitcher (1993) contains a strong defence of realism.

The classic works on the theory-dependence of observation and the incommensurability of theories are Hanson (1958), Kuhn (1962) and Feyerabend (1976). A good collection of essays on these issues is Hacking (1981). The best sources for the underdetermination of theories by evidence and the pessimistic meta-induction are respectively Quine (1951) and Laudan (1981). For a survey of recent work on naturalized epistemology, see Kitcher (1992).

Most modern discussions of explanation begin with the title essay in Hempel (1965). Explanation and its relation to causation are further explored by the essays in Ruben (1993). Armstrong (1983) provides an excellent account of the general problem of distinguishing laws from accidents, as well as his own solution. Chapter 7 of O'Hear (1989) contains a good introduction to both probability and probabilistic causation. The best contemporary non-specialist discussion of the problems of quantum mechanics is to be found in chapters 11–13 of Lockwood (1989).

DAVID PAPINEAU

The view that events discussed in the special sciences are token-identical but not reducible to physical events is defended in Fodor (1974).

References

Armstrong, D. 1983: *What is a Law of Nature?* Cambridge: Cambridge University Press.

Ayer, A. J. 1956: The Problem of Knowledge. London: Macmillan.

Boyd, R. 1980: Scientific Realism and Naturalistic Epistemology. In P. Asquith and R. Giere (eds) PSA 1980 vol. 2, 613–62. East Lansing, MI: Philosophy of Science Association.

Cartwright, N. 1983: *How the Laws of Physics Lie*. Oxford: Oxford University Press.

Churchland, P. and Hooker, C. (eds) 1985: *Images of Science*. Chicago: University of Chicago Press.

Duhem, P. 1951 [1906]: *The Aim and Structure of Physical Theory* (translated by P. Wiener). Princeton, NJ: Princeton University Press.

Feyerabend, P. 1976: *Against Method*. London: New Left Books.

Fodor, J. 1974: Special Sciences. *Synthèse*, 28, 97–115.

Friedman, M. 1984: *Foundations of Spacetime Theories*. Princeton, NJ: Princeton University Press.

Glymour, C. 1980: Theory and Evidence. Princeton, NJ: Princeton University Press.

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

——(ed.) 1981: *Scientific Revolutions*. Oxford: Oxford University Press.

Hanson, N. R. 1958: *Patterns of Discovery*. Cambridge: Cambridge University Press.

——1963: *The Concept of the Positron*. Cambridge: Cambridge University Press.

Harding, S. (ed.) 1975: *Can Theories be Refuted?* Dordrecht: Reidel.

Hempel, C. 1965: *Aspects of Scientific Explanation*. New York: Free Press.

Hempel, C. and Oppenheim, P. 1948: Studies in the Logic of Explanation. *Philosophy of Science*, 15, 135–75.

Horwich, P. 1982: *Probability and Evidence*. Cambridge: Cambridge University Press.

Howson, C. and Urbach, P. 1989: *Scientific Reasoning*. La Salle, IN: Open Court.

Hume, D. 1978 [1739]: *A Treatise of Human Nature* (edited by P. H. Nidditch). Oxford: Clarendon Press.

Kitcher, P. 1992: The Naturalists Return. *Philosophical Review*, 101, 53–114.

——1993: *The Advancement of Science*. New York: Oxford University Press.

Kuhn, T. S. 1962: *The Structure of Scientific Revolutions*. Chicago: Chicago University Press.

Laudan, L. 1981: A Confutation of Convergent Realism. *Philosophy of Science*, 48.

Lewis, D. 1973: *Counterfactuals*. Oxford: Blackwell.

——1979: Counterfactual Dependence and Time's Arrow. *Noûs*, 13, 455–76.

Lockwood, M. 1989: *Mind, Brain and the Quantum*. Oxford: Blackwell.

Mellor, D. 1971: *The Matter of Chance*. Cambridge: Cambridge University Press.

Mises, R. von 1957: *Probability, Statistics and Truth*. London: Allen and Unwin.

Neander, K. 1991: The Teleological Notion of Function. *Australasian Journal of Philosophy*, 69, 454–68.

O'Hear, A. 1989: *An Introduction to the Philosophy of Science*. Oxford: Clarendon Press.

Papineau, D. 1993: *Philosophical Naturalism*. Oxford: Blackwell.

Popper, K. 1959a: *The Logic of Scientific Discovery*. London: Hutchinson.

——1959b: The Propensity Interpretation of Probability. *British Journal for the Philosophy of Science*, 10, 25–42.

——1963: *Conjectures and Refutations*. London: Routledge and Kegan Paul.

——1972: *Objective Knowledge*. Oxford: Clarendon Press.

Quine, W. V. O. 1951: Two Dogmas of Empiricism. In *From a Logical Point of View*. New York: Harper.

314

Ramsey, F. 1978 [1929]: General Propositions and Causality. In D. H. Mellor (ed.) *Foundations*. London: Routledge and Kegan Paul.

Reichenbach, H. 1956: *The Direction of Time*. Berkeley: University of California Press.

Ruben, D.-H. (ed.) 1993: *Explanation*. Oxford: Oxford University Press.

Salmon, W. 1971: *Statistical Explanation and Statistical Relevance*. Pittsburgh, PA: University of Pittsburgh Press.

Smart, J. 1963: *Philosophy and Scientific Realism*. London: Routledge and Kegan Paul.

Teller, P. 1973: Conditionalization and Observation. *Synthèse*, 28, 218–58.

Van Fraassen, B. 1980: *The Scientific Image*. Oxford: Clarendon Press.

Worrall, J. 1989: Why Both Popper and Watkins Fail to Solve the Problem of Induction. In F. D'Agostino and I. Jarvie (eds) *Freedom and Rationality*. Dordrecht: Kluwer.

Wright, L. 1973: Functions. *Philosophical Review*, 82, 139–68.

Discussion Questions

1   Is it any less rational to accept induction than it is to accept deduction?

2   Is it always a mistake to save a theory when it has been falsified?

3   How can we show that it is more rational to believe some hypotheses than to believe others?

4   How do scientific theories move beyond the stage of conjecture?

5   Is belief a matter of degree?

6   Does Bayes's theorem help us to deal rationally with evidence?

7   Do scientific claims about unobservable entities differ in status from scientific claims about observable entities?

8   Is instrumentalism mistaken if it cannot account for some features of scientific practice? How can we determine what features are really a part of scientific practice?

9   Are general theories or piecemeal procedures more important to our basic characterization of science?

10   Does all observation depend on theoretical assumptions? What are the implications of your answer for an account of unobservables?

11   'Given any theory about unobservables which fits the observed facts, there will always be other incompatible theories which fit the same facts.' Does this make realism about unobservables untenable?

12   Can we be realists in some areas of science and instrumentalists in others?

13   Should we look to the history, sociology, and psychology of science, rather than to first principles, to identify criteria for the acceptability of scientific theories?

14   Can a naturalized study of science tell us which research strategies are an effective means to theoretical truth?

15   Is there any good reason for a philosopher to prefer to talk about explanation rather than causation?

16   How can we distinguish laws from accidentally true generalizations?

17   Must we posit an ideal system of knowledge in order to understand the notion of 'law'? What if there cannot be such a system?

18   Do scientific laws involve necessity? In what sense?

19   How can we explain the direction of causation?

20   Can we accept a model of causation according to which causes probabilify, rather than determine, their effects?

21   What problem poses greater difficulties for frequency theories of probability: 'single-case probabilities' or a reliance on hypothetical infinite reference sequences?

22   If 'propensities' are primitive, can a propensity theory give us any insight into the nature of probability?

23   How are interpretations of quantum mechanics relevant to philosophical disputes about probability?

24   Can we give up common sense in favour of a 'many-worlds' reality containing both Schrödinger's cat alive and Schrödinger's cat dead?

25   Are teleological explanations an alternative to causal explanations or a kind of causal explanation?

26   How can we determine whether all sciences form a hierarchy in which the higher can always be reduced to the lower?

27   Do different areas of science describe different realities, or just one physical reality at different levels of detail?