Chapter 10

Philosophy of Science

(Alexander Bird, University of Edinburgh)

Introduction

The central concerns of the philosophy of science as a whole divide into two areas. One of these areas, perhaps the traditionally more dominant side, encompasses such questions as: Is scientific knowledge possible? If so, how? What methods of reasoning do scientists use? What is the role of observation and experiment in scientific reasoning? Are theories justified that concern entities and processes that cannot be observed? This area is where philosophy of science overlaps with epistemology, the theory of knowledge.

The second area is situated where philosophy of science overlaps with metaphysics. Here we are concerned not with whether and why our theories amount to knowledge, but instead we focus on what our theories are about—in the most general terms. For example, the world contains *laws of nature* and is divided up into different *kinds* of things. Almost all scientific theories concern one or other of these two general features of the world—or at least their subject matter will reflect these features. Associated with features is the fact that most theories attempt to *explain* some phenomenon or other. These are metaphysical topics.

In this chapter we shall look first at the second area, metaphysical philosophy of science, before returning to the first, epistemological philosophy of science.

The metaphysics of science

I mentioned three key components in the metaphysical part of the philosophy of science: laws of nature, natural kinds, and explanation. We can see these all operating together in John Dalton's atomic hypothesis. The starting point is the obvious fact that chemists deal with different substances. The stuff of chemistry divides naturally in to different chemical kinds, which for the most part remain or can be made to remain as they are and distinct from one another. The task of the chemist is to discover by what processes these kinds will react to produce other kinds—and to enquire into why they do so. In the process of investigating the reactions that they undergo, it was discovered that substances would react together only in fixed proportions; and in the case of gases the proportions, when measured by volume, were always ratios of small whole numbers. In reacting hydrogen with oxygen we find that they will combine without residue to form water if we use exactly twice as much hydrogen as oxygen. But if we deviate from this proportion then some hydrogen or some oxygen will be left over unreacted. So we have an experimentally discovered law: the law of constant proportions. Dalton's explanation for this was that the basic constituents of chemical matter are atoms, a different kind of atom for each chemical element. When substances react they do so because atoms combine in small clusters—molecules. The molecules are the basic form of the compound substance. So that substance will contain the elements in exactly the proportion that the elements occur in the compound molecule. Hence a *law of nature* (the law of constant proportions) governs *kinds* at one level (laboratory substances), is *explained* by the existence of kinds at another level (atoms and molecules).

In the next three sections we shall look at these components in detail. What exactly does it mean to explain a phenomenon? What is the relationship between explanations and causes and laws of nature? What for that matter are laws of nature? Where does our sense that the world divides into different kinds of thing come from? Is it a reflection of the division of the world into natural kinds, or does it reflect more human interests and categories?

Explanation

The explanation of phenomena such as the law of constant proportions is a characteristic and indeed central task of science. The nature of scientific explanation raises a range of philosophical problems and controversies. A very basic one asks whether science really should be engaged in explanation at all. One view, associated with *positivist* philosophy of science, is that to seek explanations in terms of hidden processes is just so much mystery-mongering and that science should aim solely to describe carefully the various regularities to be found in the observable phenomena. Dalton's atomic hypothesis was attacked in those grounds and the related ground that even if the phenomena are explained by hidden processes we could never know what those processes really are (because, for example, we could never have good reason to choose between different possible underlying explanations). We shall return to positivist philosophy of science later. Another problem raised by the notion of explanation concerns the variety of explanation. We are all familiar with *causal* explanation, a form of explanation found in the physical sciences. But in biology we also find explanation through *natural selection*. We can explain why giraffes have long necks by citing the adaptive advantage a long neck gives an animal, enabling it to reach a greater range of leafy food. This seems different from causal explanation, since an adaptive advantage does not cause a giraffe to have a long neck; what causes a giraffe to have a long neck is its genes. We might ask whether behind this difference in kind of explanation is an unity in the underlying mechanism. For example, it is plausible to suggest that whenever one can give a correct natural selectionist explanation that is only because a more complex causal explanation is also available: mutations cause certain giraffes to have longer necks which cause them to eat more than other giraffes which in turn causes those giraffes to have more offspring, and so on.

Whether this constitutes a reduction of selectionist explanation to causal explanation is a question for philosophers of biology; similar questions arise in the philosophy of psychology and in other areas of science. A form of explanation that is perhaps most common in physics is *nomic* explanation—explanation in terms of the laws of nature.   We will see an attempted unification with causal explanation here too. In the 16th century Johannes Kepler discovered that the planets rotate about the Sun in ellipses, rejecting centuries of conviction that planetary motion was essentially circular. A century later Isaac Newton was able to explain Kepler's discovery. All objects, the planets included, obey an universal law of gravitation and basic laws of motion. Newton deduced from his laws the fact that a small object rotating under gravity around a much more massive one will travel in an ellipse. In a different case physicists can now give a full explanation of why an aeroplane wing generates lift. We take certain general laws concerning the motion of the air (or any fluid). These laws, plus a description of the particular shape of the aerofoil allow us to deduce that there will be a difference in pressure between the lower and upper sides of the wing, which in turn means that there will be an upwards force on the wing. These two examples suggest a general pattern for explanations involving laws of nature: from a statement of certain relevant laws plus a statement concerning the particular conditions surrounding the phenomenon in question we *deduce* the facts we want to explain. Diagrammatically this can be represented thus:

from	laws; particular conditions	e.g.	laws of gravitation and motion; the Sun is much more massive than the planets
we deduce	————————		————————
	the facts to be explained		the planets travel in ellipses with the Sun at one focus

The idea that we can explain a phenomenon by showing how it is in instance of a law or several laws is known as the *covering law* model of explanation. The particular version we have been looking at is the *deductive-nomological* version, where the phenomenon is deduced from the laws (plus other relevant conditions).

The originator the of the deductive-nomological version of the covering law model, Carl Hempel (Hempel 1965, 1966), extended this idea to cases where explanations do not permit us to deduce that the phenomenon in question will occur but nonetheless make it *likely*. Subatomic events, such as the decay of an atomic nucleus, are inherently probabilistic. This means that we can never deduce that such an event will occur, and so no deductive-nomological explanation is available. It might be thought that we may never the less explain why some atomic explosion occurred by citing the laws and circumstances that make it likely. This is the *probabilistic-statistical* (or *inductive-probabilistic*) version of the covering law model. The laws still make the phenomenon expected, but only with a high degree of probability rather than certainty. We can see a connection between the covering law model of explanation and prediction. If we know the explanation of an event, then we could have predicted it in advance. If we had known the laws of gravitation and motion before we had ever observed the planets, then we would have been able to expect them to travel in ellipses. If we knew that a compound consists of molecules made up of a few atoms, then we would expect the reaction producing this compound to require precise proportions of the reagents.   Conversely, according to Hempel, if we can predict an event then we are in a position to explain it once it has occurred.

One immediate problem with this view is that it seems we can explain things with information that is insufficient to permit prediction. An unfortunate baggage handler at London's Heathrow airport once contracted malaria from a mosquito which has travelled from overseas attached to some luggage. We can explain why he got malaria—he was bitten by a mosquito. But this is not enough to allow us to predict in advance that he would. We cannot deduce that he gets malaria, for not all mosquitoes carry malaria, and not all bites from infected mosquitoes lead to malaria. The fact that it was unlikely that he would get malaria is consistent with the fact that we can explain why he did after the event. So explanation does not seem, after all, to require that we deduce the fact to be explained. To a great extent this observation can be reconciled with the covering law model of explanation by distinguishing complete from incomplete (or so-called 'elliptical') explanations. A complete explanation is one that gives all the laws and conditions required to deduce the phenomenon to be explained (or to make the phenomenon highly likely). An incomplete explanation supplies just part of a complete explanation. A complete explanation of the baggage handler's malaria would include the fact that the mosquito was infected, the fact that the bite was in a certain place, the fact that the baggage handler had a physiology susceptible to malarial infection, and other such facts that altogether, along with the relevant biological and biochemical laws, would allow us to deduce that he would contract malaria. The explanations we give in most circumstances are incomplete, leaving out facts that are not of especial interest or facts about which we are ignorant. In fixing upon the mosquito bite we are focussing on only a part of the complete explanation, but an especially salient part. As this case suggests, the idea of an incomplete explanation allows us to link nomic to causal explanation. A causal explanation is an incomplete explanation where we have chosen to cite one of the salient conditions, such as the mosquito bite, rather than any of the laws.

A number of further problems beset the covering law model and its various forms. One such problem concerns the probabilistic-statistical version. Some philosophers argue that a complete explanation that shows only that some event is likely (rather than certain) does not really explain it; after all, that does not explain why it happened rather than did not happen. On this view inherently probabilistic events cannot be explained. At best we can explain why they were likely to occur, not why they actually occurred.

Different problems face both versions of the covering law model. Peter Achinstein gives this example, concerning an unfortunate Mr Jones who ate a pound of arsenic and died shortly thereafter (Achinstein 1983):

.

from the law	everyone who eats a pound of arsenic dies within 24 hours;
and condition	Jones ate a pound of arsenic
we deduce	————————————
	Jones dies within 24 hours

We would certainly have been able to predict Mr Jones' death from the law and condition given. However, this does not, in this case, explain his death. For just after eating the arsenic and while still in good shape, Mr Jones left his house and was run down by a bus. Not every correct prediction translates into a correct explanation. This suggests that there is more to explanation that just being able to deduce the phenomenon from laws and conditions.

The preceding problem can be called the problem of pre-emption. Another problem concerns the asymmetry of explanation. If A explains B then B cannot explain A. But the deductive pattern of deductive-nomological explanation does not show this asymmetry. It seems that we may explain why the pressure in a syringe increased by 50% when we pushed in the plunger by one third using a deductive-nomological explanation:

from law	under conditions of constant temperature, PV=constant;
and conditions	the temperature remained constant,
	the volume decreased by one third
we deduce	———————————
	the pressure increased by 50%

But note that the following also fits the pattern:

from law	under conditions of constant temperature, PV=constant;
and conditions	the temperature remained constant,
	the pressure increased by 50%
we deduce	———————————
	the volume decreased by one third

It looks as if we can also explain the decrease in volume by citing the change in pressure—but that is clearly the wrong way round.

One way out of both the pre-emption and the asymmetry problems would be to suggest that an explanation should mention something known to be a cause of the phenomenon to be explained. But that seems too easy an escape. First, this would prevent us from using this model of explanation to tell us what causal explanation is—for it would now assume we understood causal explanation already. Secondly, one might expect similar problems to arise when trying to give an independent account of what causes are (and indeed similar problems do arise). So this 'solution' just moves the problem elsewhere. As regards the problem of asymmetry one might want to add the proviso that the explaining conditions should precede the phenomenon to be explained in time. This would work but again it might be too easy a solution. For what determines the direction of the arrow of time? One plausible answer is that it is the direction of causation. So we could not assume we know what the direction of time is without circularity.

Another approach is to look at the bigger picture. We knew that the poison did not explain Jones' death because we knew that the bus did. If we had looked at Jones' physiology we would have seen that the chain of events that normally lead from ingesting poison to death had not been completed. Similarly, we knew that the change in the gas pressure did not cause the change in volume in the syringe, because we knew that it was that the experimenter's pushing on the plunger that caused the volume to change. So perhaps the way to discern explanations is to see how the proposed explanation fits into a larger pattern of explanations. We shall see one way how this might work in the next section of this chapter. Hempel's covering law model might still be part of the picture, since all the explanations will have to conform to it. But merely fitting that model won't be enough—integration with other explanations will also be required. However, even this approach is unlikely to be enough, because at the most basic level some of these problems will refuse to go away. One basic idea in Hempel's model might still be right—that to explain some fact fully is to show how it is the *instantiation* of some law or combination of laws of nature. That is to say, that the fact is an instance of that law in action, as when a falling stone is the law of gravity in action. However, what our examples have shown is that being deducible from a law does not show that a fact is indeed an instantiation of the law. In fact it may not be possible to give a perfect account of when laws are instantiated, or, correspondingly, of when one event causes another.

Laws of nature

Even if we cannot say exactly what it is for a fact to be an instance of a law, for a fact to show a law in action, we may at least be able to make some progress by investigating the concept of law itself. That is what this section is about. At the beginning of this chapter I mentioned the law of constant proportions. This was discovered as a result of careful laboratory measurements taken by chemists such as Jeremias Richter. The observed fact that substances reacting together in laboratory experiments do so in fixed proportions was taken to be a sign of an universal fact, that *all* chemical substances always react together in fixed proportions. So perhaps this is just what a law of nature is, an u*niversal regularity*. This is the simplest and perhaps initially most attractive view of what a law of nature is. Some laws may not seem to be fully universal. Kepler's laws of planetary motion were formulated to describe the motions of planets in the Solar System, not any other planetary system. However, even Kepler's laws are general to the extent that they tell us how the planets move at any time, not, for example, just when they are being observed. Furthermore, thanks to Newton, we know that any planetary system will obey Kepler's laws, so long as the gravitational effect of the planets on each other is negligible compared to that of their sun. Newton showed that Kepler's laws hold if and only if the Sun and each planet exert a force on each other that is attractive, that acts along the line joining them, whose strength is proportional to the product of the two masses and is inversely proportional to the square of their separation. If this is generalised from the Sun and planets to all masses whatsoever we get Newton's law of universal gravitation. As its name suggests, this law is truly universal: currently we believe that it holds everywhere and for all times. Some physicists suggest that the law is not truly universal, claiming that the so-called gravitational constant is not constant after all—the strength of the force of gravity is changing over time. But even if this is true, we would expect there to be some underlying law that shows why the gravitational constant should change. That underlying law would be genuinely universal.

It is wise to note some ambiguity in our language. Strictly speaking Kepler's laws do not hold of our Solar System, since inter-planetary forces mean that there is some deviation from those laws. Does that mean that, strictly speaking Kepler's laws are not laws at all? Or does it mean that although still laws, those laws are strictly false? Similarly, if we discovered the gravitational 'constant' to be changing, would that mean that Newton's law of gravitation is not a law after all, or that it is a law that is false? Both ways of speaking are correct. One way means by 'law' some general proposition or statement, something which can be asserted by people, be believed, and found, in due course, to be wrong (or right). The other way takes ‘law’ to be referring to the objective features of the world we are trying to uncover. On the latter view a law is just *there* making the world behave in the way that it does, independently of whether we know or believe it to be there. Since we are at this point doing metaphysics, it is more useful to reserve the term ‘law’ for this objective notion, and use ‘law-statement’ for the other way of speaking. Then we can say that law statements are attempts to describe the laws that there actually are.

So far we have formulated the view that a law (an objective law) is just an universal regularity is the way nature behaves. But some reflection suggests that this cannot be all there is to being law. A coincidence is two events happening together or in close succession. A *mere* coincidence is such a co-occurrence that is purely accidental—nothing connects the two events. Too many coincidences and we begin to suspect that they are not mere coincidences and that the kinds of event are in fact related. Conspiracy theorists tend to think that there are very few repeated mere coincidences—they see sinister connections where there is just the co-occurrence of independent happenings. A rational person must accept that mere coincidences can be repeated. If the coincidence is repeated again and again, it may be reasonable to suppose that there is some connection. But one should still admit that a repeated mere coincidence can happen, even if it is unlikely. (And we know that unlikely events do happen—which is why people hope to win lotteries). Now imagine that a chain of mere coincidences constitutes a fully universal regularity—it has no exceptions. To make this clearer, imagine that on some busy London bus at some particular moment all its passengers are people who that evening have Italian food for dinner. That could happen by sheer fluke. Now imagine that something akin to that happens on a comic scale. Would that cosmic coincidence be a law? You would think not—it would be a mere coincidence, a fluke, an accident. And the occurrences of laws are just not that sort of thing—occurrences of laws are not accidental; they have to happen the way they do.

There could be an universal regularity that is merely a cosmic coincidence, not a law. So a law must be more than just any universal regularity. It must be a non-accidental universal regularity. But what do we mean by ‘non-accidental’? This is a tricky question, since we cannot simply reply that a regularity is accidental if it does not come about as the result of a law—since the notion of law is precisely the one that we are ultimately trying to explain. Our explanation would be circular. There are two standard responses to this problem. I shall look at one solution in detail. It notes that we could say without immediate circularity that an accidental regularity is one that does not come about as a result of an *underlying* or *basic* law. Kepler’s laws are non-accidental, since they can be explained by Newton’s law of gravitation, the law that underlies them. But to avoid circularity, we must be able to give a satisfactory explanation of what an underlying or basic law is. What features do we think the basic laws should have? For a start, a basic law (such as Newton’s) explains other regularities that are non-basic laws (like Kepler’s laws) as well as individual events (such as the behaviour of a comet or space vehicle); but the basic laws does not itself get explained. How much should be explained by the basic laws? If the world is deterministic then the basic laws should between them explain every individual event. In fact the world is not deterministic; individual events such as the nuclear decay of a radioactive atom do not happen as the inevitable result of the existence of a law. Even so, the probability of that atom's decaying is subject to a law of nuclear physics. So at least the probabilities of events should be captured by the laws. In general as much information as possible about the occurrence of individual events should be captured by our basic laws. Another feature of the way basic laws operate is illustrated by the explanation of Kepler's laws or of the motion of a comet. To give a full explanation we need not only Newton's law of gravitation, we need Newton's laws of motion as well. The laws operate together to explain events and regularities. And if we are talking about the motion of charged particles, we will have to introduce further basic laws. So we expect out basic laws to form an integrated system. Lastly, we note that the history of science shows that the laws we uncover as we get closer to our basic laws are fewer in number, more general in scope, and simpler in form. For example, before Kepler and Newton, the behaviour of objects in the universe were thought to be governed by different laws for heavenly objects (the Sun, planets etc.) from the laws for earthly objects. Furthermore, the mathematically quantified laws for the motions of the planets were very complicated (both in Ptolemy's Earth-centred system and in Copernicus' Sun-centred system). Kepler then showed how the laws governing the planets could be reduced to three simple laws, while Newton further simplified matters by introducing the single even simpler law of gravitation, which with the laws of motion explain the behaviour not just of planets but (in part at least) of all objects whatsoever. The various sciences have a multiplicity of non-basic laws, but physicists have shown how to reduce the underlying laws of physics to just three laws (gravity, the electro-weak force and the strong nuclear force), and hope to unify even these. So our conception of law seems to include the thought that the basic system of laws should not only seek to encompass as many phenomena as possible, but should also be as simple as possible and should integrate into a system. A god-like mind should be able to deduce all the non-basic laws from these basic ones. And so the *systematic* account of what a law is says that a law is either basic, in which case it is one of the regularities of nature that is part of an integrated system of regularities, which is optimal in the sense of capturing as many facts, actual and potential, as possible while also being as simple as possible (Lewis 1973: 72-77). Or the law is non-basic, in which case it can be derived from the basic ones. We can now see why cosmic accidental regularities don't get to be laws. This is because we would not expect them to help explain many things, and so we would not want them as part of our basic system. At the same time we cannot derive the accidental regularity from the basic laws. We can explain the individual events making up the accidental regularity but we cannot give a unified explanation for all of them together. Return to the case of the bus passengers; we know that this is just an accident because we can give an explanation of why each passenger ate Italian food, and the explanations will be independent of one another and independent of the bus journey.

The systematic account of laws has its critics. For sure, they say, it looks as if *our* universe is governed by a system of well-integrated simple laws of great scope. But things did not have to be that way. There is no contradiction in supposing that an universe could exist with many complicated laws that do not integrate well but act largely independently. Furthermore, one could imagine that by fluke there occurred an accidental regularity in that world which would very much improve the simplicity and power of the system of supposed basic laws for that world, if it were added to that system. But that would not make the accidental regularity a law. A really amazing coincidence could still be a mere coincidence without being a law (it only looks like a law). A different sort of problem with the systematic account, and indeed with any account that thinks of laws as some kind of regularity, is that it doesn't do justice to the thought that laws explain things. A regularity, even a basic one, is just a collection of resembling co-occurrences. In what sense does the collection of such co-occurrences explain why any particular co-occurrence happened? The worry is that the natural answer is: none at all. It may be a botanical regularity that all toadstools are poisonous. But that does not explain why some particular one poisoned me when I ate it. For an explanation we want something behind the regularity, something that can itself explain the regularity, and therefore something which is not the regularity itself. These critics suggest that we should not focus on individual events or facts, nor on collections of them (i.e. not on regularities). What we should focus upon are *properties.* These critics have a particular conception of what a property is. Properties are what they call *universals*, entities that exist, in some sense, in their own right, more or less independently of their belonging to specific objects. We can then think of laws as relations among properties. So we can think of the property *black* as being an universal—all black things share one and the same property. And it can relate to other properties in a law of nature: For example, being black makes something heat up more quickly when irradiated than being white does.

Thinking of laws as relations among properties has the advantage of making it plausible that laws explain individual facts and events as well as regular co-occurrences of events. That blackness is related to heating explains why all black things heat up quickly. The law is not the regularity but stands behind the regularity. The relation between properties is often called *nomic necessitation*. We can think of nomic necessitation as something like causation (after all, it is reasonable to say that being read causes rapid eating). However that is not very helpful at explaining what 'nomic necessitation' means, especially if we aim, as well we might, to explain causation in terms of laws/nomic necessitation. We can often explain what property or relation we are thinking of by explaining what it does: *black* is the property that makes things appear black to normal observers, or is a property that makes things heat up rapidly; *being heavier than* is the relation that explains why some thing take more force to lift or accelerate than others. What would explain the relation of *nomic necessitation*? To reply 'it is the relation that relates properties in laws' is clearly going to be circular if we want to use the concept of nomic necessitation in explaining what laws are.

I'll mention one last view of laws that differs importantly from both the systematic regularity and the nomic necessitation views. This view at least shares this with the nomic necessitation view, that it takes properties to be of fundamental importance. But the nomic necessitation view thinks that properties exist independently of one another and might be conjoined in any combination. It might be that in some other possible universe, white things heat up more rapidly than black, or that colour and heating are not connected at all. The *dispositionalist* view of properties and laws notes something important from the previous paragraph, that properties can be specified by their causal or functional role. Perhaps these roles are *essential*. That is, to say what a property is, is to say what causal role it fulfils. It might, for example, be the role of *being negatively charged* that things with this property repel one another and attract things with the property of being positively charged. So, necessarily, negatively charged objects repel one another. If some objects with property *P* failed to repel one another, *P* would not be negative charge. If so, then the law that negatively charged objects repel just follows from the existence of the property of negative charge. We do not need a separate account of laws (see Mumford 1998). One intriguing aspect of this view is that it makes the laws of nature hold in all possible worlds (where the property exists). This appears to conflict with an intuition that the laws of nature could have been different from what they actually are. However, it is unclear just what weight we should put on this intuition.

Natural kinds

The things and stuffs we find in this world divide into different kinds. Some differences in kind are due to the fact that the things in question are made by us for different purposes, knives are different from forks, tables from chair, buses from trains, because these items have been designed with different functions in mind. Much of science proceeds upon the assumption that there are also natural divisions of things into kinds—animal species, chemical elements, kinds of subatomic particles, and so forth. The central philosophical problem surrounding natural kinds asks what it is that divides things into natural kinds. The parallel with kinds of artefacts is unlikely to be helpful, since that would suggest that natural kinds are divided by virtue of being designed for a purpose. Although that answer might be congenial to some religious views, it is inadequate if we seek a natural rather than a supernatural explanation of the existence of natural kinds. Another outlook that sees a close analogy to artefactual kinds regards the division of things into kinds as reflecting human interests, even if natural objects are not designed by us. For example, the greengrocer's division of produce into fruit and 'vegetables' reflects a culinary perspective, and hence tomatoes and cucumbers are not regarded as fruit, even though they perform the same seed-bearing role as apples and strawberries. We think that there is a natural botanical division as well as the culinary one, but can we be sure that this is not just a matter of a different anthropocentric perspective?

One way we might ensure that we achieve a natural division with our concepts is to provide careful definitions of our kind concepts in terms of purely scientific properties: the definition of 'fruit' as 'the seed bearing part of a plant' seems to be scientifically respectable in a way that 'sweet-tasting part of a plant' does not. This however just postpones the problem. For the definitions will succeed in picking out natural divisions only if the terms employed in the definitions themselves reflect genuinely natural kinds, in this case 'seed' and 'plant'. Any definition must refer to properties that things may or may not have and these properties must be natural properties if the definitions are to define natural kinds. But the question, what makes a property a natural property is no advance on the question, what makes a kind a natural kind.

There are two ways of tackling this problem, both of which make progress towards an answer. The first approach points out that we would expect humans to have some innate ability to perceive difference in kinds. This is because such an ability would have been required by our ancestors in coping with the environment in which they found themselves. An individual who cannot spot the differences between a lion and an antelope or between a poisonous plant and nutritious one will not survive for long in the African savannah. Looking at the other end of the evolutionary chain, even single celled creatures respond differently to different kinds of chemicals in their environments. So natural selection will confer on us at least some ability to detect differences in natural kinds.

Building on this initial innate ability we can develop means of detecting yet further kinds. For example, natural selection has given us the ability to see different colours. This may not be enough to tell the difference between two different chemical substances which look exactly the same; but those substances may burn to produce flames with different colours, and so this test will allow us to know of another natural division.

The second approach denies that our kind concepts work by having precise verbal definitions of the sort envisaged. Take as a simple example the concept of 'water'. Today we might employ a scientific definition: "compound with formula H_2O". But that water is H_2O was a discovery made in the nineteenth century about stuff that people we already calling 'water'. So before the discovery, the term 'water' must have been referring to water even without the benefit of such a definition. Saul Kripke and Hilary Putnam have argued that it is plausible that no definition was available then that could have successfully picked out precisely water (Kripke 1980; Putnam 1975). This is because in a limited state of chemical knowledge the best definition of water would have been something like 'a colourless, tasteless liquid that boils at 100°C and freezes at 0°C'. Now imagine that on some other planet superficially like Earth there is a colourless, tasteless liquid that boils at 100°C and freezes at 0°C, and which is found in the places where water is found on Earth. However this stuff is not H_2O but has some other formula, say XYZ. On the accepted assumptions that what 'water' refers to hasn't changed since the eighteenth century and that water is precisely H_2O, it follows that XYZ is not water and never has been. However, had people in the eighteenth century employed the supposed definition of water ('a colourless, tasteless liquid that boils at 100°C and freezes at 0°C') then XYZ would have passed their definition and would have been water. Hence it follows that in the eighteenth century neither this nor anything else like it was the *definition* of water. (This argument from Putnam and others from Kripke are discussed in more detail by Alexander Miller in his chapter on the Philosophy of Language.)

So it seems that in the eighteenth century people were able to refer to water without having a definition of water. That itself should be no great surprise, for not every word can have an informative, non-circular verbal definition. Indeed it is plausible that most do not. But if a verbal definition is not what gives the concept of water its content, what does? The proper names of people provide an useful analogy here. Proper names don't have verbal definitions. People get their names in naming ceremonies. Other people may get to know the name by being introduced to the person in question, or by having them pointed out. But we can use a name to refer to someone we have never met nor seen. In such a case, what makes that name refer to them is the fact that our use of the name is linked, via the use of the name by other people, all the way back to the person who bears that name. So we can explain that what ties the use of a name to the person is an initial naming ceremony followed a series of causal connections between one person and another. Often such connections will be reinforced by 'repetitions' of the naming ceremony, as when someone is introduced by name.

Something similar may be regarded as happening with common kind names, such as 'water'. At some time in the past something like a naming ceremony took place when our ancestors agreed to use 'water' (or some predecessor of that word) to refer to the stuff they found around them, which they could see and point to. That links our use of ‘water’ to the stuff water. That link is reinforced (rather more strongly that the original link) by the fact that as children we are introduced to the word ‘water’ in connection with samples of water. Since we are not linked in this way to any samples of XYZ, our word does not refer to XYZ. A similar story may be told about many other natural kind terms. According to this view of how we refer to natural kinds, we can do so without being able to describe them definitively. Perhaps we need a conception of ‘kind of stuff’ so that we can pick out water/H_2O (but not XYZ) when we say ‘water is this kind of stuff’. The pointing (‘this’) and the notion of stuff will be enough to permit reference to a single natural kind. Of course not *all* natural kinds can be identified this way—subatomic particles are not easily pointed to. But we can nonetheless forge a similar sort of connection with them via their effects. We can see the light emitted by a gas in the cathode ray tube or the traces in a cloud chamber. This may not count as seeing electrons, but we can say ‘electrons are the things that explain these effects’. In yet more sophisticated cases we might have to articulate some theory in which a kind of thing or stuff is postulated to explain some phenomenon—as neutrinos were postulated to explain the ‘missing’ mass in beta decay.

It is reasonable to suppose that, speaking generally, there are natural kinds. The structure there is in the universe, which requires both difference and sameness, would not be possible without natural kinds. The laws of nature govern the behaviour of entities according to the kinds of thing that they are. Natural kinds are those kinds that play a part in genuine explanations of natural phenomena. But not all the kind terms that occur in language are natural kind terms. So what, we asked, distinguishes genuine natural kind terms from other linguistic expressions? What reason do we have to suppose that they really do reflect naturally occurring divisions among things? The first part of the answer was that the meanings of some words are attached to our innate perceptual discriminatory abilities. We would expect for evolutionary reasons that these should pick out real rather than imagined differences between things. But not all differences in kinds will be perceptual. The second part of the answer suggests that our terms refer to natural kinds because the way our language works. On the (reasonable) assumption that there are indeed natural kinds, the function of certain terms is to latch onto the natural kinds that there are. This latching on occurs independently of our discriminatory abilities. In some cases we are in direct contact with a kind. If we inaugurate the use of the word ‘water’ by saying ‘water is this stuff’ pointing at what happens to be H_2O, then water is H_2O and not XYZ, even if we cannot tell H_2O and XYZ apart. In other cases we are not in direct contact with the kind. But we may know about the phenomena that a kind explains. In such cases a kind term my be introduced by an explanatory theory. If we say ‘neutrinos are the particles whose emission explains the missing mass of beta decay’ then ‘neutrino’ refers to that kind of particle (if there is such a kind) even when we had not worked out any way of detecting neutrinos.

In the last case, either the word ‘neutrino’ will pick out the natural kind of neutrinos—if the theory is true—or it will pick out nothing at all—if the theory is false. Knowing whether ‘neutrino’ successfully refers will depend on knowing whether the theory of neutrinos is true or false. Knowledge of the truth of theories is the subject of the next part of theis chapter.

The epistemology of science

Scientists carry out experiments and make observations. They do so in order to generate and gather evidence for or against theories or hypotheses. A *hypothesis* is a claim whose truth cannot be known directly through observation but only indirectly via inference from evidence. A *theory* is a set of related hypotheses, which will typically have some degree of generality and will help explain a range of phenomena. So, for instance, Newton's theory of gravitation explains, among other things, the orbits of the planets around the Sun. When it was observed that the orbit of Uranus did not precisely fit the theory, it was hypothesised that an unseen planet was affecting Uranus's orbit (which was how the planet Neptune was discovered).

A hypothesis for which there is supporting evidence is said to be *confirmed* by that evidence. The better confirmed a hypothesis is, the more justified a scientist is in believing it. An important question therefore is, when does evidence confirm a hypothesis? A related but distinct question is, under what circumstances does a well-confirmed hypothesis count as scientific *knowledge*? Both these questions have generated a considerable amount of philosophical discussion, and continue to do so.

Confirmation

The first and simplest account of confirmation is *enumerative induction* (sometimes called Humean induction). According to naive enumerative inductivism evidence takes the form of an observation of a regular conjunction of properties. Your evidence might be that fact that all ripe bananas you have eaten have tasted sweet. Or, at a more complicated level, your evidence might be that having tested a variety of pendulums of different lengths and measured the period of their swings, the period (in seconds) is always 2.01 x the square root of the length (in meters).

The inductivist account of confirmation is simple: observational evidence confirms those hypotheses that say that what has not been observed will be like what has been observed. "More of the same' is the motto of inductivism. The weakest hypothesis to be confirmed by the evidence will say that the next observed thing will be like the previously observed thing: the next ripe banana will taste sweet, the period of the next pendulum will also be 2.01 x the square root of its length. A more ambitious hypothesis confirmed by the same evidence says that *all* relevant things, whether past, present or future, whether observed or unobserved will be like the observed—all ripe bananas are sweet, the periods of all pendulums are equal to 2.01 x the square root of their lengths.

Enumerative induction gives the impression that the way to do science is to go about observing things and then to infer that what is unobserved and what is in the future will be just like what has been observed in the past. It is also a feature of enumerative induction that it cannot tell us about the nature of things which cannot ever be observed, such as the interiors of atoms or stars. Since we cannot observe these at all, we cannot use enumerative induction to learn about them. A more natural way to look at matters is to suggest that a scientist first constructs a hypothesis. She deduces observable consequences from it, and then checks whether these observable consequences are true. If they are, then those observations confirm the hypothesis. This is the hypothetico-deductive model of confirmation, devised by Carl Hempel. An example of this model in operation is the confirmation of Einstein's general theory of relativity, which says that light should be bent by massive objects. Sir Arthur Eddington deduced how the light from a star should be bent by the Sun. This allowed him to say where the star should be observed to be during a solar eclipse. In 1919 Eddington made the relevant observations which were in close agreement with the deductions from the theory (but not with those from Newton's theory). Thus Einstein's theory was spectacularly confirmed. When presented schematically,

from	hypothesis;	e.g.	the general theory of relativity;
	particular known conditions		facts about the position of stars, Earth, and Sun
we deduce	————————		————————————
	observation		the star is to be observed in such-and-such a place during the eclipse
and thus confirm the hypothesis			

we can see the parallel between Hempel's deductive-nomological model of explanation and his hypothetico-deductive model of confirmation. Note that the hypothetico-deductive model can incorporate enumerative induction in that according to the former 'all ripe bananas are sweet' is confirmed by 'this ripe banana is sweet', as enumerative induction maintains, because the particular case is deducible from the hypothesized general case.

Although hypothetico-deductivism is an improvement on enumerative induction, it still has its problems. One of these is known as the paradox of confirmation, or as the raven paradox. Let our hypothesis be 'all ravens are black'. If we come across a raven, we deduce from the hypothesis that it should be black. And so if it is black, the hypothesis is confirmed. So far so good. We next come across some white object in a tree. Notice now that the hypothesis allows us to deduce something about this white object—that it is not a raven. So we investigate this white thing and find that it is a plastic bag, and so certainly not a raven. The observable deduction from the hypothesis was found to be true. According to hypothetico-deductivism this too means that the hypothesis is confirmed. But is seems strange that the hypothesis 'all ravens are black' should be confirmed by a white bag, and indeed by any non-black thing which is not a raven. (Another way to see why this is a consequence of hypothetico-deductivism is to see that the hypothesis 'all ravens are black' is equivalent to 'all non-black things are non-ravens'.)

There are two response to the raven paradox. The first takes the conclusion—that a white bag can confirm a hypothesis about ravens—to be an absurdity. Therefore the hypothetico-deductive model should be abandoned and we should look for a better account of confirmation. The second response is to accept that the white bag does confirm the hypothesis, albeit to a very small degree. After all, it might be pointed out, on observing the white object it is still an open possibility that it will turn out to be a raven, thereby proving our hypothesis false. So it is some small success for the theory that on closer inspection it turns out to be a bag not a raven.

A rather different picture of scientific reasoning is provided by *abductivism*, which was first articulated by Charles Sanders Peirce. Abductivism is also called *Inference to the Best Explanation* (I use the terms interchangeably), which spells out the basic idea. Scientists look for possible explanations of the phenomena they observe. When sufficient evidence has been gathered they may be in a position to infer that the best of the possible explanations is the actual, true explanation. For example, palaeontologists have long asked, why did the dinosaurs become extinct? Answers to this question come in the form of different possible explanations: (i) they lost out to mammals in a Darwinian struggle for existence, (ii) dust from an enormous volcanic eruption blotted out the sun, killing off their food sources, (iii) a huge meteor collided with the Earth, creating a similar dust cloud. It is thought to be in favour of the meteor hypothesis that it explains the unusual amounts of the element iridium found in geological strata associated with dinosaur extinctions. Iridium is rare in the Earth but more common in meteors. That sort of evidence may show one explanation to be much better than the others. According to abduction, that is a reason to believe it to be true.

Abductivism can also incorporate enumerative induction. If all diamonds have been found to be harder than quartz, then perhaps the best explanation of this observed fact is that it is a law of nature that diamond is harder than quartz. If that is true, all future and unobserved diamonds are harder than quartz. So enumerative induction is a special case of abduction, where the best explanation is some natural law. While abduction can include enumerative induction, it also goes further. Like hypothetico-deductivism, abductivism allows us to make inferences concerning things we have not ever seen, nor could ever see, like subatomic particles, or the Big Bang, or the movements of continents, or the extinction of the dinosaurs. Although we may have never seen a certain thing or event it may still be the best explanation of what we do see. Dalton, we saw, regarded the existence of unseen atoms as the best explanation of the observed facts summarised in the law of constant proportions.

Abductivism provides us with an answer to the raven paradox. Hypothetico-deductivism told us that the white shoe confirms the hypothesis that all ravens are black. Abduction tells us that evidence is relevant to a hypothesis if it bears on the explanatory power of that hypothesis.. Since the white shoe in the tree is irrelevant to the explanatory power of the hypothesis that ravens are black, it provides no confirmation. The hypothetico-deductive model of confirmation is mistaken. Not everything deducible from a hypothesis confirms that hypothesis. This ties in with what we discovered about explanation. According to the deductive-nomological conception of explanation we can explain an event by deducing its occurrence from a law. But we saw that this was wrong, thanks to the example of Mr Jones, whose death was deducible from the law that a pound of arsenic is fatal but not explained by that law (being explained by the bus instead). Correspondingly, Mr Jones' death does not confirm the hypothesis that a pound of arsenic is fatal.

Scepticism and science

David Hume believed that people do actually infer, on the whole, in accordance with naive enumerative induction (Hume 1739). This he regarded as a manifestation of the tendency of the human mind to habituate itself to repeated occurrences of events. He argued, however, that natural though enumerative induction might be, it nonetheless cannot be the product of reason. His successors have taken this to mean that induction cannot lead to a rationally justified belief in a hypothesis. If he is right then evidence cannot ever inductively confirm that hypothesis.

For, said Hume, in order for the rule of enumerative induction to yield knowledge of the truth of a hypothesis, we must know that the rule is reliable. After all, there are other rules one could employ, for instance 'counter-induction' which says: if all observed Fs have been G, then expect the next F *not* to be G. Some gamblers argue like this when they think that a run of heads in tossing a coin makes tails more likely on the next toss, in which case the rule will lead one astray. Sometimes the rule of counter-induction will be more reliable, for example when one is predicting the colour of the ball to be taken from a bag with red and green balls in it, from which many red but few green balls have already been taken. Let us say that we are on some occasion using the rule of enumerative induction to confirm a hypothesis. Note that however much evidence we have, it does not logically guarantee that the next F will be G. However many ravens we have seen, all of which have been black, it is still conceivable that the next one will be white. So it cannot be logic which tells us to use the rule of induction. If not logic, then perhaps experience tells us to use the rule. After all, as a matter of fact we do use induction much of the time, and on the whole we get by pretty well—experience tells us that the rule of induction has worked for us in the past. So we should expect it to work well for us now. But here is Hume's catch. The argument from experience says that we should expect the rule to work in the future because it has worked in the past. That argument is itself inductive—it justifies our use of induction by using an inductive argument. It therefore assumes what it sets out to justify. The argument from experience seems to be a circular argument, and so Hume concluded that there is no rational justification for inductivism. Our use of it is not a matter of reason but of habit.

Nelson Goodman discovered another puzzle which also seems to undermine enumerative induction (Goodman 1954). We can invent a new adjective *grue* which is defined as follows:

X is grue = either X is green and observed before midnight on 31 December 2050
or X is blue and not observed before midnight on 31 December 2050

So an emerald observed today is grue, while a blue sapphire dug up for the first time in 2051 will also be grue. It follows from this definition that anything seen until today which is green can be counted as grue. So, in particular, all observed emeralds are grue. The rule of enumerative induction says: the fact that all observed Fs are Gs confirms the hypothesis that all Fs are Gs (for any F and G). If we put 'emerald' for 'F' and 'grue' for 'G' we have: the fact that all observed emeralds are grue confirms the hypothesis that all emeralds are grue. We have seen that it is true that all observed emeralds are grue. So now let's look at the hypothesis which is confirmed: all emeralds are grue. This covers *all* emeralds, whether observed or not. Which means that an emerald which is first mined and seen in 2051 is grue. Now look at the definition of 'grue'. It says that when something is grue, if it has not been observed before midnight on 31 December 2050, then it is blue. So this emerald, first seen in 2051, should be blue. That is, enumerative induction allows us, using the term 'grue', to infer that emeralds first seen after 2050 will be blue. Not only is this a strange conclusion, it is contradicted by employing enumerative induction using our normal word 'green', which tells us that the same emeralds are green. So it looks as if, depending on which vocabulary we use, we can use any evidence to confirm any hypothesis about the future, however bizarre. At the very least, in order to use enumerative induction, we must restrict our vocabulary to words like 'green' which seem well-behaved, and exclude ones like 'grue'. But how do we decide what to allow and what to exclude? It looks as if the ones we want to include are natural kind terms; the problem is that we discover which the natural kinds are by using scientific reasoning. If Goodman's problem casts doubt on the latter, then we are prevented from using this route to a solution. Furthermore, related problems arise even with well-behaved vocabulary. Let us say you are an ornithologist. You have observed many ducks, and they have all been brown. So using enumerative induction you conclude that all ducks are brown. You have also made a study of arctic birds in general and observed that they are all white. So you conclude that all arctic birds are white. However, you have never seen an arctic duck. Should you conclude that arctic ducks are brown or that they are white?

Some philosophers have believed that Hume's problem shows that induction cannot possibly be the basis of scientific knowledge and rational belief. So either one must say that science is irrational. Or one must show that something other than induction is the foundation of scientific reasoning. Sir Karl Popper was one such philosopher and believing that science is rational he sought to develop a non-inductive account of science (Popper 1959). Since he rejected induction, Popper held that observing a black raven could never confirm the hypothesis that all ravens are black. But consider the person observing the white object in the tree. Imagine this time that the white object does turn out to be a raven. Now we know for sure that the hypothesis is false. So although we can never know that hypotheses are correct, we can know that they are mistaken.

This leads to Popper's characterization of the route by which science progresses, called *falsificationism*. This can be summed up as *conjecture and refutation*. First the scientist constructs a bold hypothesis—a conjecture. Then he deduces observable consequences from it. So far this is similar to hypothetico-deductivism. But instead of seeking agreement between the predictions and what is observed, the falsificationist looks for sources of possible *disagreement*. He tries his hardest to refute the theory. If he succeeds in finding a refutation—a disagreement between what the conjecture predicts and what is observed—then progress of some sort has been made. We know the conjecture is false, and now we construct a new one. However long this process goes on, we do not get to know that a general theory is true. Even if a theory stubbornly resists every attempt to falsify it we do not know it is true. Such a theory is said to be *well-corroborated*. But, insists Popper, corroboration should not be confused with confirmation. It is never a reason to believe a general theory or hypothesis.

Many scientists have found Popper's picture of science attractive. It describes the imaginative construction of conjectures and their rigorous testing. It tells us never to rule out the possibility that a successful theory might yet be falsified. This strikes a chord with physicists who regard Newtonian physics as refuted by Einstein, after centuries of remarkable corroboration. But other scientists will not be so happy with the idea that we can never know any theory to be true. Biologists and geneticists regard Darwinian evolution as supplemented by Mendelian genetics as sufficiently well-corroborated as to be regarded, in the main, as a body of knowledge. Some would say that even if we do not know our theories to be true, we have reason to think that they are getting closer to the truth. The idea of 'nearness to the truth' is called *verisimilitude*. Without induction, increasing corroboration does not show even increasing verisimilitude. This degree of scepticism seems incompatible with depicting science as rational. How could science rationally aim at truth or verisimilitude if there is no way of knowing whether we have achieved those things?

Realism and anti-realism

What is the aim of science? We have just seen one answer, that science aims at true theories—or at least at theories with increasing verisimilitude. This is the answer given by the realist. But the anti-realist argues that this cannot or should not be the aim of science. Instead the purpose of a scientific theory is to be able to generate true predictions. A theory can yield accurate predictions even if it is not true. Since our only way of evaluating a theory is through testing its predictions, the most we can expect of a successful theory is that it should go on producing true predictions, not that it is itself true.

Debates along these lines have surfaced at several points in the history of science. Between the 2nd century and the 15th century, astronomy in Europe and in Arab countries was dominated by the theories of the Greek-Egyptian astronomer Ptolemy. The main component of his picture was that the Earth is at the centre of the Universe, with the Sun, planets and stars circulating around the Earth. Other details of his theory enabled him to achieve a high degree of accuracy. The fit between his theory and observation was impressively close, enabling detailed and successful predictions to be made. In 1543 the Polish cleric and astronomer Copernicus published a rival theory, which has the Earth being among the other planets rotating about the Sun which is at the centre of the Universe. Copernicus's system, which has many of the complexities of Ptolemy's theory, was able to achieve an even greater degree of predictive accuracy, which made it popular with many astronomers. On the other hand the theory conflicted with a number of religious beliefs, as well as contemporary opinions in physics. Therefore in a preface to Copernicus's book *On the Revolutions of the Spheres* the theologian Osiander proposed that the new theory should not be seen as aiming to state truths about the planets, but rather as seeking to provide an improved basis for computing their positions in the night sky as observed from the Earth. Osiander was proposing an anti-realist interpretation of Copernicus's work. The view, called *instrumentalism*, holds that theories do not aim at presenting a true picture of the world; instead the purpose of a theory is only to produce accurate predictions.

Earlier I explained that the chemist John Dalton had proposed invisible atoms and unobservable modes of their combination as the best explanation of the experimental law of constant proportions. Since the hypothesis could not be confirmed—or refuted—by observation, many of Dalton's contemporaries and successors rejected it is as pointless speculation. Chemists such as Sir Benjamin Brodie considered that chemical theories should confine themselves to articulating laws concerning quantities and operations which could be observed or performed in the laboratory. However, Dalton's atomic hypothesis became more widely accepted as a result of its being able to explain a new phenomenon, *isomerism*, discovered after his time. Isomers are distinct substances, which have different physical or chemical properties, but which share the same chemical formula and so are compounds of the same proportions of chemical elements. On the atomic hypothesis this is easily explained, since the same set of atoms might combine in different structures to form molecules with different properties. For instance they may form molecules which are mirror images of one another and so will bend light in different directions. The explanatory power of atomism ensured its success.

Brodie, in rejecting Dalton's atoms, was influenced by *positivism*. Positivism has its roots in the empiricism of David Hume and was first expounded by the French philosophers Saint-Simon and Comte in the early part of the nineteenth century. Positivists reject metaphysics and hypotheses concerning the unobservable—for them the point of science is to establish correlations among observable phenomena. A very influential form of positivism—logical positivism—was developed by the Vienna Circle group of philosophers in the 1920s. According to their principle of verification, the meaning of a sentence is the way in which it is verified. So, roughly, to ask what a sentence means is to ask how one might decide whether it is true. Since the logical positivists took the ground of all empirical knowledge to be experience and observation, often understood as the sensations of the observers, it follows that for them the meaning of any sentence is reducible to some claim about what one may observe or sense. So, in particular, any theoretical claim in science is really just about observations and sensations. One version of this positivism about meaning is Bridgman's *operationalism*. According to operationalism, any scientific statement is really a statement about various sorts of operations and measurements—there is no genuine theoretical difference without an observational difference. A statement about electrons is not really about invisible particles but instead is about laboratory items such as vacuum tubes or bubble chambers.

If apparently theoretical statements are really about observable things, then it ought in principle be possible to translate the theoretical statements into purely observational ones. This turned out impossible to do. So some positivists developed an alternative view according to which theoretical statements are strictly speaking not about anything at all. Hence theories cannot properly be said to be true or false. But we can still deduce observational predictions from them. Predictive utility is the measure of a theory, not truth. Either way, the positivists held that there must be a part of scientific language that is purely observational. For example, if theoretical statements are (as on the earlier view) reducible to statements about observations, it must be that there is some language for discussing observations which is itself free from any element of theory. The idea that there is a pure 'observation' language has since been discredited. Ludwig Wittgenstein attacked the idea that there could be a 'private language' of sensation terms defined by their user directly by association with his or her experiences. Various philosophers have argued that all our scientific language is to some extent theoretical (for example, Maxwell 1962). Furthermore there seems no clear dividing line between what is observable and what is not. The term 'gene' was introduced by Wilhelm Johannsen to name the mechanism whereby biological characteristics are passed from generation to generation. Genes were at that time unobservable, theoretical entities. But now with high powered microscopes and the techniques of DNA sequencing, genetic material can be observed and manipulated. The unobservable has become observable.

Recent empiricists, such as Bas van Fraassen, have rejected the positivist idea that theoretical talk is different in kind from observation talk (van Fraassen 1980). And he accepts that there is a spectrum from observable to unobservable. Many theories postulate the existence of highly unobservable entities with unobservable properties, while our evidence is all at the observable end of the spectrum. Van Fraassen's *constructive empiricism* points out that many different theories about unobservables will have the same consequences as regards observables. Our evidence is like seeing the tip of an iceberg; lots of differently shaped icebergs might have similarly shaped tips; hence we cannot know from looking just at the tip of some iceberg what it looks like beneath water. Similarly van Fraassen argues that we cannot know from just our observable evidence which the true theory about unobservables is. The best we can hope for is a theory which is *empirically adequate*—one which generates true predictions about observables. This reminds us of Osiander's preface to Copernicus' planetary theory—stating that his theory was just an instrument for accurately calculating the visible positions of planets. So van Fraassen agrees with the positivists that the aim of a theory is to yield accurate predictions. But they disagree in their views regarding what theories are about. The positivist thinks that a theory is either strictly speaking about our observations and instruments or is really about nothing at all. Van Fraassen thinks that there could be electrons, neutrinos, and other subatomic particles and that our theories do attempt to describe such things. Only, he argues, we simply have no reason to suppose that they are successful in this respect.

Realists reject positivism and empiricism about theories. Against the positivists, they claim that our theories are genuinely about unobservable entities, and against empiricists such as van Fraassen, realists hold that if there is enough evidence we can know those theories to be true. Realists have been successful in defeating positivism, primarily because the theory-observation distinction could not be maintained in the way the positivists required. This is now accepted by other anti-realists, such as van Fraassen. As regards the debate between realism and van Fraassen's constructive empiricism, a key issue centres on the ability of abductive inference to yield knowledge. Realists will claim that we know that electrons exist, since their existence is the best explanation of many phenomena, such as chemical bonding, cathode rays and static electricity. More generally, many realists will claim that without a realist view of science, the success of science would be a miracle. In other words, the best explanation of the predictive and technological success of science is that science has, by and large, given us true theories.

Unsurprisingly, Van Fraassen rejects inference to the best explanation as an admissible mode of inference, at least when hypotheses concerning the unobservable are concerned. His main criticism is that because there are infinitely many possible competing hypotheses that might fit a given set of observational evidence, we have little chance of even thinking of the one true hypothesis. So even if abduction can select the best hypothesis *among those considered*, the chances are slim that this hypothesis is true. Furthermore, why should we think that the best explanation is more likely to be true? In his novel *Candide*, Voltaire makes fun of the philosopher Leibniz in the character of Dr Pangloss. Like Leibniz, Pangloss believes, in the face of disaster, that this the actual world is the best of all possible worlds. Similarly we may ask, why should this world be one in which the best explanations tend to be true? Might not the world be one where poor explanations are very often true and good explanation tend to be false? For example, we think that to be a good explanation, a hypothesis should be simple and powerful, should unify diverse phenomena, and should give us an intellectually satisfying insight or understanding of the nature of things. But might not the world be a complicated place, where diverse phenomena are indeed diverse and not unified, where the true nature of things is not intellectually satisfying? If the world were like that, then an inference to the best explanation would lead us to get things wrong. It might be thought that the success of science shows that abductive inferences tend to lead to the truth. That is to say, the best explanation of the success of science is the fact that the world is after all a place where simple, powerful, unifying, and satisfying explanations tend to be true. That looks to be a case of using abduction to justify abduction, in which case we have another instance of Hume's problem.

Reliabilist epistemology and the philosophy of science

To conclude I shall describe how many philosophers of science now think that we should tackle the most general epistemological problems in the philosophy of science. In this area we should expect the philosophy of science to draw upon advances in general epistemology. There the refutation of scepticism draws upon 'externalist' epistemology—epistemology that says that the world may play a part in enabling us to know things; justified believing does not depend solely on the way we organise our beliefs from an internal perspective. Among philosophers of science a favoured form of externalist epistemology is *reliabilism*. (Externalist epistemology and reliabilism are discussed in more detail in Alan Goldman's chapter on Epistemology.) According to reliabilism a person knows something when they have a true belief brought about by a reliable process: a process that generally delivers true beliefs. What makes this 'externalist' is that what makes a process reliable is whether the world is such a process that brings about true beliefs and not false ones. It does not matter whether or not the person in question knows that the process is reliable. (The process is reliable as a matter of fact in the 'external' world, but may not be *seen* to be reliable from the internal perspective of the thinker's mind.) For example, a perceptual process like seeing can let us have knowledge of our surroundings so long as it is reliable, even if we do not know that it is reliable. The same goes for the sorts of belief-forming process employed in science. Some such processes involve enumerative induction. These can give us knowledge so long as they are reliable. That is, if we live in a world where the laws of nature make nature regular in the relevant respects, then an inductive belief-forming process can be knowledge-yielding. If I hypothesise that all acids will turn my test solution red, and have discovered that a wide variety of acids do indeed turn it red, then my conclusion that the hypothesis is true will be knowledge if nature is indeed uniform in this area. The sceptic claims that we must justify our use of induction in order to be able to use it to get knowledge. This the externalist denies; the latter says that I don't need to have proven first that nature is uniform to get knowledge from induction.

This sort of scientific use of induction is not so very far from the pre-scientific inductive learning from experience that we all employ every day. Our disposition to make inductions of a simple sort is innate, another product of evolution and natural selection. Since creatures, especially humans, have to live in a diversity of environments, not all useful knowledge can be innate. We have to be able to learn. But a disposition to learn that made faulty inductive inferences would be no advantage. So we are born with a disposition to infer inductively, and furthermore this disposition is accompanied by an innate sense of when inductions are likely to be successful, that is we are disposed to make induction in circumstances when nature is relevantly uniform. The lesson of Goodman's new riddle of induction was that we cannot infer inductively with just any old property that comes to mind. The properties in question have to be natural ones. As suggested when discussing natural kinds, another of our innate abilities is the ability to distinguish at least some natural kinds and properties. Our inductive disposition is a disposition to infer with respect to these kinds and properties and not with respect to others.

Our scientific inferential practices go well beyond these basic inductive capacities, which are shared to some degree with many animals. But science nonetheless builds upon these capacities. One respect in which science seems to go well beyond simple induction is in our ability to make sophisticated abductive inferences (inferences to the best explanation). Quite what the nature of inference to the best explanation is, and when we may expect it to deliver knowledge, are questions of current debate. Just as we have an intuitive sense of when inductions are likely to be reliable, it may be that scientists acquire, through long experience, a sense of just when abduction is likely to be reliable. Whatever answers we come up with, externalist epistemology tells us that we do not need to worry about general problems of induction. So long as our abductive dispositions are reliable they can be knowledge yielding. In particular, if they are reliable, then our use of inference to the best explanation to generate beliefs in electrons, neutrinos and other unobservables, in fact generates knowledge of these things. So externalist epistemology applied to abduction allows us to respond to anti-realist scepticism as well.

Conclusion

For much of its history the philosophy of science has been dominated by empiricism. With its emphasis on observation as the basis of knowledge and understanding in science it has been a fruitful doctrine in encouraging an experimental and investigative approach to science. Empiricism does however also limit the ambitions of science. We saw that empiricism in the guises of positivism, instrumentalism, and constructive empiricism rejects the view that science can uncover the hidden natures of things. All we can hope to know concerns the observable features of the world. Empiricism even limits our conception of what there is in the world. Positivism denied that we could even talk about genuinely theoretical, unobservable things. Elsewhere we encountered the regularity conception of laws of nature. The empiricist thinks of laws as just regularities in the way things are, while the anti-empiricist want to think of laws as something deeper—laws are what stand behind regularities and explain why they exist.

It is fair to say that empiricism is no longer the dominant philosophy of science. Most philosophers of science think that there are indeed hidden natures to things and that science has the ability to give us knowledge of them, even if there are debates about quite how it does so. Nonetheless there persist powerful traces of empiricism, such as the conviction that scientific evidence is ultimately observational and that the laws of nature are contingent. It remains a task for philosophy of science to decide whether these are lasting legacies or residues that will ultimately wither.

Questions

1. How does explanation differ from 'prediction after the event'?
2. What is a law of nature? Is it any more than just a regularity found in nature?
3. Are there really *natural* kinds? What (if anything) makes green things form a natural kind but not grue things?
4. When does evidence confirm a hypothesis? What is the relationship between confirmation and explanation?
5. What are the ground for scepticism concerning knowledge of (a) things we could observe but have not observed yet, and (b) things we cannot observe? Do we have a response to such scepticism?

Bibliography of works cited

Achinstein, Peter 1983 *The Nature of Explanation* Oxford: Oxford University Press.
Goodman, Nelson 1954 *Fact, Fiction and Forecast* Atlantic Highlands, NJ: Athlone Press.
Hempel, Carl 1965 *Aspects of Scientific Explanation* New York: Free Press.
Hempel, Carl 1966 *Philosophy of Natural Science* Englewood Cliffs NJ: Prentice-Hall.
Hume, David 1739 *A treatise of Human Nature* London.
Kripke, Saul 1980 *Naming and Necessity* Cambridge, Mass: Harvard University Press.
Lewis, David 1973 *Counterfactuals* Oxford: Blackwell.
Maxwell, Grover 1962 "The Ontological Status of Theoretical Entities" in Feigl, Herbert and Maxwell, Grover *Minnesota Studies in the Philosophy of Science vol. III* eds. Minneapolis: University of Minnesota Press.
Mumford, Stephen 1998 *Dispositions* Oxford: Oxford University Press
Popper, Sir Karl 1959 *The Logic of Scientific Discovery* London: Hutchinson.
Putnam, Hilary 1975 "The Meaning of 'Meaning'" in *Philosophical Papers* vol. ii *Mind, Language and Reality* Cambridge: Cambridge University Press.
van Fraassen, Bas 1980 *The Scientific Image* Oxford: Oxford University Press.

Further reading

Much of this chapter relates to my book *Philosophy of Science* (London: UCL Press/Routledge, 1998). A useful collection of readings that connect well with my book and this chapter is to be found in Martin Curd and Jan Cover *Philosophy of Science* (London: Norton 1998). Good introductions to the philosophy of science include James Ladyman *Understanding Philosophy of Science* (London: Routledge 2002), John Losee *A Historical Introduction to the Philosophy of Science* (Oxford: Oxford University Press, 4th edition 2001), and Carl Hempel *Philosophy of Natural Science* (Englewood Cliffs NJ: Prentice-Hall, 1965). The latter is out-of-date in many respects but is nonetheless a very clear and stimulating introduction to the subject. Hempel's work is still a starting point for most discussions of scientific explanation. David-Hillel Ruben's *Explaining Explanation* shows how matters have progressed, and the collection edited by Ruben *Explanation* (Oxford: Oxford University Press, 1993) contains key papers by Hempel, Railton, and others. On laws of nature I recommend David Armstrong's *What is a Law of Nature?* (Cambridge: Cambridge University Press, 1983). W.V. Quine's "Natural Kinds" in his *Ontological Relativity and Other Essays* (New York: Columbia University Press, 1969) is a centrally important paper on that subject. The problems of inductivism are forcefully articulated by Sir Karl Popper in his *Logic of Scientific Discovery* (London: Hutchinson, 1959) along with his falsificationist alternative. For hypothetico-deductivism we may return to Hempel's *Philosophy of Natural Science*. Goodman's new riddle of induction is introduced in his *Fact, Fiction and Forecast* (Atlantic Highlands, NJ: Athlone Press, 1954). The many issues surrounding abductivism/Inference to the Best Explanation are very clearly addressed in Peter Lipton's *Inference to the Best Explanation* (London: Routledge, 1991). On realism and anti-realism you cannot do better than Stathis Psillos' *Scientific Realism* (London: Routledge, 2000). A good collection of important papers in the philosophy of science that deals mainly with the realism debate is David Papineau (ed) *The Philosophy of Science* (Oxford: Oxford University Press, 1996).