PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

Phil. Trans. R. Soc. B (2005) **360**, 1259–1269
doi:10.1098/rstb.2005.1660
Published online 28 June 2005

The Medawar Lecture 2004
The truth about science

Peter Lipton*

Department of History and Philosophy of Science, University of Cambridge, Free School Lane, Cambridge CB2 3RH, UK

The attitudes of scientists towards the philosophy of science is mixed and includes considerable indifference and some hostility. This may be due in part to unrealistic expectation and to misunderstanding. Philosophy is unlikely directly to improve scientific practices, but scientists may find the attempt to explain how science works and what it achieves of considerable interest nevertheless. The present state of the philosophy of science is illustrated by recent work on the 'truth hypothesis', according to which, science is generating increasingly accurate representations of a mind-independent and largely unobservable world. According to Karl Popper, although truth is the aim of science, it is impossible to justify the truth hypothesis. According to Thomas Kuhn, the truth hypothesis is false, because scientists can only describe a world that is partially constituted by their own theories and hence not mind-independent. The failure of past scientific theories has been used to argue against the truth hypothesis; the success of the best current theories has been used to argue for it. Neither argument is sound.

Keywords: philosophy of science; history of science; truth; methodology; Karl Popper; Thomas Kuhn

1. WHAT IS THE PHILOSOPHY OF SCIENCE?

Astronomers study the stars; philosophers of science study the astronomers. That is, philosophers of science—along with historians and sociologists of science—are in the business of trying to account for how science works and what it achieves. This characterization of the philosophy of science as a principled description of science is only a first approximation, but it is a good place to start.

Is trying to describe how science works a worthwhile activity? Science is one of the most complex and impressive communal activities our species performs, so it is natural enough to wish to understand how we do it. Worthwhile, but relatively easy, one might think. For unlike stars, astronomers can talk. To find out how science works, it might seem that all that is required is to find a few cooperative scientists to describe what they do. Of course things are not that easy, because there is such a large gap between what people can do and what they can describe. It is one thing to be able to ride a bicycle; it is something quite different to be good at describing the physics and the physiology behind that ability. The same contrast between doing and describing applies to more cognitive activities. It is one thing to be expert at distinguishing grammatical from ungrammatical strings of words in one's native tongue; it is something quite different to be able to specify the principles by which this discrimination is made. The same applies to science: it is one thing to be a good scientist; it is something quite different to be good at giving a general description of what scientists do. Scientists are not good at the descriptive task. This is no criticism, since their job is to do the science, not to talk about it. Philosophers of science are not very good at describing science either, and this is more embarrassing, since this is their job. It turns out to be remarkably difficult to give even an approximately correct general description of what are apparently the simplest aspects of scientific research. For example, philosophers of science have worked long and hard to provide a general account of the apparently simple tripartite distinction between data that would support a given hypothesis, data that would tell against it, and data that would be irrelevant; but at least most of these attempts turn out to have the highly implausible consequence that the set of irrelevant data is empty. Rather than criticize the descriptive project of the philosophy of science on the grounds that it is too easy to be interesting, one might with considerably more justice complain that it is far too difficult. However, we do the best we can.

The descriptive project—the attempt to tell the truth about science—is thus more than enough to be getting on with, but this first approximation to a characterization of the philosophy of science is incomplete. There are at least two other significant dimensions to the discipline. In the first of these, philosophers of science come close to going native in their involvement with the detail of scientific research: one sees this, for example, in some of the best technical work in the foundations of physics. Often this involvement with the substantive science focuses on various anomalies found in the science itself. Highly successful scientific theories may nevertheless make gross predictive errors, lack mathematical rigour, generate infinities or singularities, fail to fit in the right way with other theories, or use inappropriate concepts, and these and other internal

* (pl112@cam.ac.uk).

Lecture delivered 15 June 2004
Typescript received 20 September 2004

© 2005 The Royal Society

difficulties often attract the philosopher's detailed attention (Sklar 2000).

The second additional dimension to the philosophy of science is normative. Philosophers, it is said with some justice, are concerned not just with what is, but with what ought to be. The ought may be ethical, and in the philosophy of science there is a range of pressing ethical questions about the proper conduct of research, the just use of scientific resources and the legitimate application of science. There are also central questions about norms that are cognitive rather than ethical. At issue is not just how science works, but also what it achieves. What cognitive attitudes scientists actually adopt towards the claims they consider is a descriptive question, what attitude they are entitled to adopt is a normative question in which philosophers of science have an obsessive interest. The question is not just of the truth about science, but whether science tells the truth. Are we entitled to believe that our best scientific theories are at least approximately true? Is this even the proper aim of science? The question whether science is in the truth business is a central concern of this paper.

Alongside the truth questions, however, this paper considers questions of another sort, concerning the attitudes of scientists towards the philosophy of science. Here too one can distinguish descriptive from normative questions. On the descriptive front, of course attitudes vary. Some scientists have an intense interest in the philosophy of science. Peter Medawar, the great immunologist in whose honour the lecture that is the basis of this paper was given, is a good example of this species. Medawar wrote extensively, elegantly and insightfully on the nature and methods of science (e.g. Medawar 1982). Although very much influenced by the work of Karl Popper (to whom we return below), Medawar developed his own positions and arguments in ways that go beyond Popper's own contribution. However, Medawar is not a typical case. Many scientists are indifferent to the philosophy of science and some are downright hostile. What are the sources of this hostility? That is a descriptive question we now consider. Later we will also touch on the normative question: is the hostility justified?

2. HOSTILITY TO THE PHILOSOPHY OF SCIENCE

There are at least three different reasons why a scientist might disdain the philosophy of science: it gets science wrong, it is useless and it is pernicious. There is considerable justice in the first accusation. The best philosophical accounts of scientific practices are thin and primitive as compared with the richness and subtlety of their subject. The accounts are often wrong, seriously misdescribing what is actually going on in the scientific community. This should not be surprising, given the enormous complexity and opacity of scientific practice. However, while the difficulty of the task and the modesty of the results so far are good reasons not to accept any philosopher's dicta as gospel, they are not good reasons to dismiss the project. Although every available philosophical account of science has serious weaknesses, our work, to date, has provided partial, but genuine, illumination of certain aspects of science. Even if this is too optimistic, it of course does not follow that illumination is not possible. Indeed, even if one believed that science is for some reason indescribable in principle, there might yet be some value in the practice of attempting (and failing) to describe it. For example, one might learn something important about science by discovering the reasons it resists description.

As regards the first objection to the philosophy of science then, I conclude that while some, and perhaps even all, philosophical accounts of science give only poor descriptions, this is not a good reason to reject the project.

The second objection is that philosophy of science is useless. Here too I have considerable sympathy, at least if the criterion for utility is that the philosophy leads directly to improved scientific practice. Scientists who turn to the philosophy of science for this reason are generally disappointed: they do not often find that it helps them with their science. Why not? The simplest reason would be that the philosophy is badly wrong, reducing the objection from uselessness to the first objection, from misdescription. It remains possible that an account of science that seriously misdescribes its subject might yet yield prescriptions that are powerful and productive, but this seems unlikely (although we will consider a possible example below).

In any event, the falsity of a description of a cognitive activity is not the only possible explanation for the failure of that description to support useful prescription. Even if an account of some aspect of scientific practice were entirely correct, it might not be a useful tool for improving that practice. There are several possible reasons for this; the most important harks back to the distinction between doing and describing. Even if there is a good general description of some practice, it does not follow that attempting to follow that description is a good way to acquire or to improve that practice. Juggling is not learned by studying physics and physiology: it is learned by imitation. As Thomas Kuhn and others have stressed, scientists are trained into their practice by begin taught exemplary problem solutions in the field, not by being given general rules of research (Kuhn 1970). Kuhn gives a plausible account of various aspects of scientific practice, and it may help to explain why philosophers who give a rule-account of scientific practice will not be very helpful to scientists at the forefront. Of course, Kuhn's is itself a general account, but even if correct, it is not at all clear that learning it will improve scientific practice. Scientists will learn their craft by learning from exemplars in their own speciality, not by learning the Kuhnian story. Thus there are at least two good reasons for expecting the philosophy of science to be broadly useless: it is wrong, and even if it were right, it would not be in the right form to help the practicing scientist.

How should a philosopher of science react to this? Both an uncompromising and a compromising response are available. The compromising response is to say that there are some aspects of the philosophy of science that are in fact both correct and salutary for a practicing scientist to know. More generally, spending some time worrying about the problems in the philosophy of science may benefit the scientist. I will return to how this may happen towards the end of this paper. The uncompromising response is that the philosophy of science is not supposed to be useful. Astronomy can be justified without claiming it will improve galactic behaviour; similarly, one can justify the philosophy of science without supposing that it will improve scientific behaviour. Science is such an important part of our lives that it seems worthwhile to try to understand a little better how it works, for the sake of understanding itself rather than for any more utilitarian end. Bacon (1620) wrote of both the fruit and the light that science produces. The light is the insight, the fruit is the technology, the increased control over nature. Many areas of science do indeed have both these benefits, but not all of them do, and the chance of illumination alone may be sufficient to justify a scientific activity. The same can be said about the philosophy of science. If it can illuminate, however partially, some aspects of scientific practice, then it is probably worth the very modest investment we make in it. As in other areas of philosophy, the justification for the philosophy of science cannot plausibly be supposed to lie primarily in its technological benefits.

This leaves the third objection, which is that the philosophy of science is worse than useless: it is pernicious. One reason one might believe this follows on from the point just made about philosophical descriptions being in the wrong form to be useful. One might go so far as to worry that a scientist sufficiently misguided as to attempt to do research by following some philosophical principle would actually risk degrading their practice. This is cousin to the thought that someone may juggle competently until thinking too much about it, at which point all the balls are dropped. The correct response here, however, is not to reject the philosophy of science, but only its misapplication, if such there be.

Scientists who think science studies pernicious usually have a quite different concern. The accusation is that many of us who study scientists are putting science down, by maintaining that it is all just politics, 'social construction', and so forth. Not only does this traduce the glory of science, but if it were widely believed, it would threaten society's support. Surely, there is no need to invest such enormous resources in something that scientists in the end just make up. This is a delicate matter and one that arouses the passions of the 'science wars', but in my view the accusation that science studies diminish science is largely an understandable misperception. Many of the issues here revolve around the truth question that will shortly occupy our attention, the question of whether we should see science as in the business of revealing the truth about a largely unobservable reality. The view of science as an extraordinarily successful truth-generating enterprise is a glorious image, and it is fair to say that this is not an image that all of science studies promotes; but to fail to promote this image is not to denigrate science. Heedless of the dangers of compressing a very complicated situation into a couple of paragraphs, here follows the situation as I see it.

Compare the enormous scope of scientific theories with the extremely limited causal interaction of the scientific community and its instrumentation with the world it is out to describe. We sense directly only a tiny portion of the world, past and present, and on this slender basis, we affirm an elaborate and extensive story about the causes of what we sense, and more generally, about the nature of the world, past, present and future. That scientific story can never be unique: there must always be alternative causes that would have produced the same sensations. The same point can be made in terms of the relation between theory and data. Theories go far beyond the data that support them; indeed, the theories would be of little interest if this were not so. However, this means that a scientific theory is always 'underdetermined' by the available data. Even if there were no anomaly in sight, so the theory is compatible with all known data, it is impossible to deduce or prove the theory from them. (Here we ignore the worry that some of the data may themselves be incorrect.) This means that there always exists, in principle, incompatible theories that also would have fit those data. This logical point—the underdetermination of theory by data—which has been developed with particular force by Duhem (1914) and Quine (1951), puts severe strain on the idea that science reveals the truth. Among historians and sociologists of science, the reaction has been not so much to deny that science is in the truth business as to leave the question to one side. For underdetermination makes vivid the impossibility fully to explain why a scientific community ends up with a particular theory by appeal only to the data that are mustered in its support, since those data will not explain why scientists ended up with that theory rather than any other one that would also have fit them. That is the kind of contrastive question—why this theory rather than another one that would have fit the data equally well—that interests many historians and sociologists of science, and it is the kind of question that it seems can only be answered by appeal to social and cognitive factors, not by appeal to the physical facts. For this contrastive question, the way the world is, is irrelevant, since it is something in common to both sides of the contrast and so will not explain it. The situation among philosophers of science is quite different, but the deep concern with underdetermination remains. Unlike historians and sociologists, philosophers are obsessed with the truth view of science and are often eager to defend it, but as we will see, the problem of underdetermination makes this very challenging. It is the same story as in other areas of the theory of knowledge: the sceptical arguments—the arguments that purport to show that we are not entitled to our truth claims—are always the arguments that seem the most powerful, even if the least desired.

Although this cannot be the whole story, the sensitivity to the problem of underdetermination goes a long way towards explaining the problematic role of truth in science studies, and so also, the perception by some that science is somehow being diminished. Perhaps there is some diminishment here: it depends on the image with which one begins. Certainly, if the starting point were a picture of science according to which all of our best current theories were absolutely correct and proven, then considerable diminishment is only to be expected. However, for all the disputes within science studies, there is a consensus that there is no better way than science to find out about the physical world and a consensus over the enormous power and control over that world that science and engineering create. Indeed, among many of us there is considerable 'science envy' and the fond hope that science studies might come just a little closer to the level of profound achievement displayed by the history of the sciences. To give up the view that science is revelatory of the unvarnished truth is by no means to affirm that science is a put-up job, or to suppose that saying things are a certain way would make them so. That is why I turn now directly to consider different views about what it is that science achieves. I myself remain deeply attracted to the truth view; but there are alternatives which should be taken seriously and which would leave science glory enough.

3. WHAT DOES TRUTH HAVE TO DO WITH IT?

Philosophy abhors a vacuum, so there are many different views about what science should be taken to achieve: here, I consider just three broad options. The first of these, known in the philosophical literature as scientific realism, is the view that science is indeed in the truth business, that its aim and achievement are to provide increasingly accurate and comprehensive descriptions of a largely unobservable reality whose nature is independent of the theories used to describe it. Science gives us maps of reality. It is no part of this map view to claim that all, or indeed any of our best current theories are entirely correct, or even that science will ever generate such theories. Nevertheless, truth is the name of the game according to the map view, and science is good at it. In particular, according to this view, later scientific theories tend to be more comprehensive and closer to the truth than those they replace. Realism is probably the view most popular among scientists themselves; it is also a view directly threatened by the problem of underdetermination.

The second position is instrumentalism, although it might also be called the computer view. Scientific theories are not to be understood as attempts to describe or map an invisible realm, but rather are calculation devices meant to provide increasingly accurate predictions about observable states of affairs. Realism and instrumentalism coincide so far as the observable claims of a scientific theory go—in both cases these aim to be true—but the two positions diverge in the case of the claims that appear to describe unobservable entities, properties and processes. According to realism, those claims too are meant to be true; but according to instrumentalism that is to misunderstand their purpose: theories are just models, not designed to be true to reality, but to 'save the appearances'—to give the right answers about what we observe. On some versions of instrumentalism, the sentences of high theory are not even claims at all, not the sort of thing that could be true or false, any more than an integrated circuit in a computer could be true or false; rather, they are reliable or unreliable in supporting calculation. On another and currently more philosophically popular version of instrumentalism, developed by van Fraassen (1980), scientific theories are indeed true or false, we just do not care which. Van Fraassen's position brings out the relationship between realism and instrumentalism with particular clarity. The two views agree about what scientific theories say—they agree that the theories are making claims about a largely unobservable world. However, where the proponent of realism suggests that we believe that the best such theories are at least approximately true, van Fraassen advises that we only 'accept' such theories, where to accept in his sense is only to believe the observable claims of the theory to be true, not the theories as a whole. As for the balance of the theory's content, we are advised to be agnostic: we cannot know whether any of those claims about what cannot be observed are correct, and there is no scientific need to know. What matters is not that our theories be true, but that they be 'empirically adequate'. By severely restricting the range of our belief, we reduce the problem of underdetermination.

The third and final view to sketch here, which I will call projectivism, is the most difficult to explain. Like realism and unlike instrumentalism, projectivism has it that scientific theories are attempts to describe the world accurately in both its observable and its unobservable regimes. However, unlike realism, projectivism denies that the world being described is entirely independent of the process of investigating and describing it. To see what this denial might come to, it is helpful to have the example of a distinction developed by philosophers and scientists (or rather 'natural philosophers') of the seventeenth century. This is the distinction between primary and secondary qualities. Galileo, Robert Boyle and John Locke all maintained that while certain 'primary' properties that we attribute to objects, including size, shape and mass, are indeed entirely independent of our experience of them, there are also 'secondary' properties, such as colour, where this is not so. According to Locke (1689), colours are partially constituted by the experience we have of them: they are dispositions to produce certain kinds of experience. (Galileo held the simpler view that colours are simply the experiences; Boyle vacillated between Galileo and Locke's view.) Of course, the disposition of grass to produce in me the experience of green has a physical basis in the properties of the microstructure of the surface of the blades; but those properties are not themselves colours. By contrast, if I perceive a primary quality of an object, say the roundness of a ball, although the ball presumably has the disposition to produce that experience of roundness, here the basis of that disposition is the actual roundness of the ball itself. Colours, on Locke's account, provide a good example for projectivism. When I see the grass to be green, I am not hallucinating; nor can I see whatever colour I want to see. The colour I see depends in part on what is going on quite independently of me. Nevertheless, on the Lockean view, colour is defined in part in terms of human response: for an object to be green is for it to be disposed to produce a certain kind of experience. Thus although we see colours as being 'out there' in the objects, there is a sense in which that perception is in part a projection of the inner experience.

In the seventeenth century, this distinction between primary and secondary qualities was typically employed in aid of realism: the fact that secondary qualities such as colours have a peculiar anthropocentrism is taken to be the reason that science should not appeal to them. Scientific theories should appeal only to primary qualities, since these are the objective and mind-independent features of reality that science is out to describe. This was a central part of the 'corpuscularian philosophy', however, towards the end of the eighteenth century, Immanuel Kant (1783) turns this on its head. According to him, we can neither experience nor represent properties of things in the world as they are 'in themselves', but only as they manifest themselves as secondary properties. Kant held that all the properties that science might describe—including size, shape and mass—are secondary in approximately Locke's sense.

This is a version of projectivism. When a bridge is constructed, it cannot be built in any way one likes: the world constrains what materials and designs are possible. At the same time, a bridge is dependent on human activity: it is a human construct, if anything is. In that sense, a bridge is a joint produce of the human world and the world quite apart from human activity. Similarly, although that analogy will only go so far, on Kant's version of projectivism, the properties that science attributes to the world are real, but are joint products of the things in themselves and the organizing, cognitive, descriptive activities of scientists. This view of science is not easy to articulate or indeed to contemplate, but it is an important alternative to both realism and instrumentalism, and one to which we will return below, when we discuss the views of Thomas Kuhn.

4. KARL POPPER

The two distinguished twentieth-century philosophers of science whose names are most familiar to scientists are Karl Popper and Thomas Kuhn. Popper is a realist, Kuhn a projectivist, and a brief discussion of their views will help to clarify those positions. (Those who wish to learn more about instrumentalism might look at ch. 2 of van Fraassen 1980.) Neither Popper nor Kuhn are typical proponents of realism and projectivism, respectively, but their views have been influential, and they also provide particularly illuminating cases for considering scientists' attitudes towards the philosophy of science.

The ancient saying is that the fox has many ideas, but the hedgehog has one big idea. Popper (1959, 1962, 1972) is a philosophical hedgehog, and his big idea is the power of negative evidence. The fundamental point is simple: no number of, for example, white swans, can ever prove that all swans are white (since there might always be a black one hiding in the rushes); but a single non-white swan disproves the hypothesis. Even if every other swan, past, present and future, were white, the statement that *all* swans are white must still be false. Popper gives this logical asymmetry between positive and negative evidence hyperbolic application, maintaining that positive evidence has no probative value whatever and that negative evidence should often be treated as tantamount to disproof. Moreover, a claim is scientific only if there are possible data that would contradict it, and scientific research consists in generating bold general hypotheses and then attempting to refute them.

Popper's refusal to give any weight to positive evidence seems at first incompatible with realism, for if there is no such thing as positive evidence, then there can be no reason to claim that a particular hypothesis is likely to be true or approximately true. Yet Popper finds a way staunchly to defend his somewhat heterodox version of realism. He insists the goal of science is truth, and argues that scientific research can be shown to be a rational activity with respect to that goal, if it is conducted along Popperian lines. The trick is to shift perspective from the question of justifying a particular theory to justifying a choice between competing theories. We can never have any reason to say a theory is true, according to Popper, but if we have two competing theories where one of them has been refuted, then we have a telling reason for preferring the other. It may too, of course, be false, but at the moment what we know is that the first is false while the second may be true, so we are eminently rational to prefer the second, given that truth is our goal. We are also rational then to turn to the preferred theory and to try to refute it too, replacing it with yet another theory should the attempt at refutation be successful. Scientists will never be out of work on this view, since the process of conjecture and refutation never yields an end state where there is reason to believe that the latest theory is indeed correct.

In my experience, if a practicing scientist has time for any philosopher of science, that philosopher is most likely to be Popper. What explains this (relative) popularity? It helps that Popper's writing is unusually clear and engaging, and that he displays some acquaintance with real science. However, the main reason for Popper's popularity is that he is strongly prescriptive. As I suggested above, what scientists would like from the philosopher is useful advice on how to improve their practice, and Popper is one of the very few philosophers in the history of the subject who seems to do this. His central prescription is very simple. Do not look for evidence in favour of your latest hypothesis; look rather for evidence that would refute it.

Philosophers have, on the whole, reacted to Popper very differently from scientists. Although he has attracted a few vigorous supporters, most philosophers of science find his basic position unacceptable. They find two main difficulties. The first is that the simple logical asymmetry Popper exploits does not adequately capture the actual situation the scientist faces. The simple logical point is that if a hypothesis deductively entails a false prediction, then the hypothesis must be false as well. When scientists actually derive predictions, however, they can almost never do so from the hypothesis under test on its own: they almost always need diverse additional 'auxiliary' premises, which appeal to other theories, the correct functioning of the instrumentation, the absence of disturbing forces, and so on. When a prediction fails in such a circumstance, all that logic tells us is that at least one of the premises must be false, not which one. The point is familiar: when an experiment does not work out as predicted, there is always more than one possible explanation. Positive evidence is never conclusive; but neither is negative evidence, nor would it be a good idea to pretend that it was.

The second main difficulty that philosophers find in Popper's scheme is that while it is a version of realism, it is also entirely sceptical. Popper's attempt entirely to do without positive evidence is noble, particularly in light of the remarkably resilient argument against positive evidence, known as 'the problem of induction', constructed by the great eighteenth-century philosopher David Hume (1748). However, almost all philosophers find the price Popper pays prohibitive. That price is being forced to accept that there can never be any reason whatsoever to believe that any scientific claim is even approximately correct.

It is remarkable that a philosopher who holds such an extreme sceptical view of science is, by some margin, the leading candidate for scientists' philosopher of science. Part of the explanation is probably rhetorical. Popper's tone is so resolutely pro-science, that I conjecture that most scientists who look at his work simple miss or elide the extreme sceptical consequences. Another part of the explanation is simply that Popper's advice does seem useful. Indeed, perhaps it is so, and an example of how a philosophy of science that is itself fundamentally flawed may nevertheless have practical value. Popper can serve as a corrective to an irrational but natural 'confirmation bias' that leads almost all of us to look disproportionally for evidence that confirms our own ideas (Kahneman *et al.* 1982).

5. THOMAS KUHN

Thomas Kuhn is also a hedgehog: his big idea is the importance of 'paradigms'. Unfortunately, that word has been used in so many senses, not least in Kuhn's (1970) classic discussion itself, that its meaning has been almost blurred out of existence. A central and relatively clear sense of the world, however, is the one that Kuhn referred to with the term 'exemplar': a concrete problem solution in a given scientific speciality that guides research in that field. Having turned from doing physics to the history of science, Kuhn was struck by the prevalence of periods of what he came to call 'normal science' in which there is striking research consensus. What struck him was not just that the scientists in such a period tend to agree about theory and data, but that they tended to go on in the same way: the agreement extended to choice of new problems, choice of techniques to solve them and standards of solution. It is as if all the scientists in the group had the same secret rulebook for doing good science in their speciality.

The nonexistence of the rulebook gave Kuhn his question: how does one explain the rule-like behaviour of a scientific community in the absence of rules? Kuhn's answer: by exemplars. In content, rules and exemplars are quite different: rules are by nature general, while exemplars are specific problem solutions in specific specialities. However, in function, argues Kuhn, exemplars are similar to rules, because they create 'perceived similarity relations' that give them a general import. Shared exemplars explains shared research trajectory, because practitioners will choose new problems that seem similar to the exemplary problems; they will try to solve them using techniques that seem similar to those in the exemplar solutions; and they will judge the adequacy of any proposed solution by tacit appeal to the standards that the exemplars exemplify. Thus exemplars function like the rules that are not there.

Kuhn's exemplar mechanism has considerable plausibility. As every teacher and student of science ought to know, the function of problem sets is not to test but to instil understanding. And it seems undeniable that past achievements in a scientific discipline profoundly influence researchers' choice of new problems and of techniques deployed to solve them. Kuhn makes the most out of the exemplar mechanism, parlaying it into an ambitious account of the sociology, semantics and metaphysics of science. Sociologically, he uses the exemplar mechanism to distinguish different periods of scientific development. A period during which the community of specialists share a set of exemplars that are performing their functions successfully is normal science. Communal scientific behaviour looks very different in the 'pre-paradigm' period before shared and effective exemplars are found, in the crisis period when the exemplars continue to throw up new problems but cease to support adequate solutions, and during a period of scientific revolutions when, according to Kuhn, one set of exemplars is jettisoned in favour of another. Semantically, he argues that exemplars combined with theory structure are what provide the meaning of the novel terms introduced by a scientific theory, so that after a scientific revolution, with the change in both theory and exemplars, the change of the meaning of terms is so extensive that the competing theories are 'incommensurable': the claims of the one theory cannot even be precisely expressed in the terms of the other, and the comparison between competing theories becomes a much more complicated and messy matter than philosophers of science have traditionally supposed.

Metaphysically, Kuhn's notorious claim is that, after a scientific revolution, 'the world changes'. What can this mean? It cannot mean merely that after a scientific revolution people's *beliefs* about the world change: that is too obvious. On the other hand, it had better not mean that the physical world, *as it is entirely independently of human thought or intervention*, changes. That is not obvious: it is crazy. We need here a third hand. What Kuhn means depends on the projectivism he endorses, the view that the subject matter of science is not the world entirely independent of us, but only the world as it is structured by us. Kuhn understands that structuring primarily in terms of systems of classification into kinds and properties, a task in which exemplars and the similarity relations they induce play a central role. The structured world is thus a kind of joint product of the things in themselves and the structuring activity of scientific description. When Kuhn says that the world changes, his claim is that it is this structured world—not just belief and not the world in itself—that changes. This is not a crazy view: since the structured world is partially constituted by scientific belief, a radical change in belief could change such a world. However, it is a radical claim, since Kuhn is asserting that this changing world is the only one that science can describe. Kuhn is thus like Kant in supposing that our knowledge can only extend to the description of a world partially constituted by our knowledge; but whereas Kant thought that there was only one form that the human contribution to this structured world could take, Kuhn maintains that the contribution changes across scientific revolutions. Kuhn is Kant on wheels.

The response of scientists to Kuhn has been mixed. On the one hand, some social scientists have been impressed and have seen in Kuhn a passport to the higher status they see physical scientists enjoy. All they need to do, they think, is to settle on a paradigm. Kuhn himself rejected this appropriation of his work. For something to serve as an exemplar it is not enough that scientists agree to endorse it: it must support the solution of new problems, and that power is not something that can be voted into existence. Kuhn held that the social sciences are still in the pre-paradigm period, and only time will tell if that will change. A more common reaction to Kuhn, however, at least among physical scientists, is hostile: they see him as one of those who would demean science, making the wild and damaging claim that it is all just politics. But this hostility is also based on misunderstanding. Kuhn certainly does not give everything that a proponent of realism would like to have, but the idea that Kuhnian science is power politics is a parody of his views. For one thing, unlike some other figures in science studies, Kuhn's account is strikingly 'internalist', explaining the development of science entirely in terms of features internal to the scientific community. Indeed, with some justice he apologized for this in the preface of *The Structure of Scientific Revolutions*: '…except in occasional brief asides, I have said nothing about the role of technological advance or of external social, economic and intellectual conditions in the development of the sciences' (Kuhn 1970, ix). In my view it is no accident that Kuhn's account is fundamentally internalist. In order for the exemplar mechanism to work effectively, scientists must be free to choose their own problems, guided by their exemplars, rather than having those problems imposed from without. Kuhn himself recognized this elsewhere, pointing out that the structure of research must be different in engineering, where problems do generally have an external source (Kuhn 1977).

Nor did Kuhn maintain that 'might makes right' within the scientific community. Scientific activity is of course sharply constrained by the world: the data are not whatever you might wish them to be. Indeed, for Kuhn, the central activity of normal science is the attempt to deal with the anomalies that face any interesting scientific theory, and the clearest example of an anomaly is precisely the clash between theoretical expectation and empirical result. Nor would it be just to accuse Kuhn of denying that science is in the business of describing the world. That is science's job, according to Kuhn, although he insists that the only world that is describable is one that is partially structured by the scientists themselves and that those structures have changed across the history of science. This projectivism is not a common-sense position, nor, as we have seen, is it anything like the only option on the truth question, but neither is it silly or demeaning.

It is not just scientists who have given Kuhn a hostile reception: many philosophers of science have also disapproved. In many cases, this was based on the same misunderstandings that provoked scientists, although that is much less common now than it was in the first 20 years after the publication of *Structure*. Even now, however, it is common enough to hear philosophers attributing to Kuhn the view that since competing theories are incommensurable they are incomparable, and hence that scientific change is an irrational activity. This is so even though Kuhn explicitly maintained that incommensurable does not mean incomparable, but rather that comparison, although essential, is complicated and sometimes inconclusive. He also maintained that science is the model of cognitive rationality, although here again the nature of rationality turns out to be complicated. Fortunately, however, not all philosophical criticism of Kuhn is based on misapprehension. For example, one may accept that much of what Kuhn says about the ways in which exemplars organize a research tradition is correct and important, while objecting that he has not made out the case for the enormous semantic and metaphysical power he goes on to attribute to those model problem solutions. In particular, one might argue that an exemplar-based methodology is compatible with realism's answer to the truth question, compatible that is, with the view that science can describe aspects of the world as it is 'in itself'.

6. THE PESSIMISTIC INDUCTION

Having considered some of the options available for answering the truth question and having seen the choices made by two of the most important figures in the philosophy of science, it is time to turn to a direct appraisal of realism, the view that science has the goal of truth. I will consider two of the most discussed arguments in the philosophical literature on this question, one against realism, the other for it. Both arguments are somewhat unusual by general philosophical standards, because they both treat the truth question as a broadly empirical hypothesis. If scientific communities are considered machines that generate theories, what is the empirical evidence for and against the hypothesis that the output of these machines will tend over time to be theories that are true?

The argument against the truth hypothesis is the pessimistic induction, and the evidence it appeals to is from the history of science (Laudan 1984). That evidence strongly suggests that scientific theories have a sell-by date. The history of science is a graveyard of theories that were empirically successful for a time, but are now known to be false, and of theoretical entities—the crystalline spheres, phlogiston, caloric, the ether and their ilk—that we now know do not exist. Science does not have a good track record for truth, and this provides the basis for a simple empirical generalization. Put crudely, all past theories have turned out to be false, therefore it is probable that all present and future theories will be false as well. That is the pessimistic induction. It is no argument for the impossibility of truth in science, but it is an argument that the relevant evidence we have goes against the truth hypothesis, against the hypothesis that the machinery of science is such as to generate output that will tend over time to be true. The best theories of the moment may seem very impressive for a time, but if we stand back from the present we see that they are unlikely to avoid the fate of their predecessors.

Defenders of the truth hypothesis have not taken the pessimistic induction lying down; but most of their responses have been concessive. For example, one might concede that it is probable that no interesting and detailed scientific theory is going to be perfectly correct, but emphasize that some falsehoods are closer to the truth than others. That is, as we have already noted, a proponent of realism may appeal to the notion of approximate truth and characterize scientific progress as a steady approach towards the strict truth, even if that approach is asymptotic. This reply concedes both the premises and the conclusion of the pessimistic induction. That is, it concedes that all scientific theories—past, present and future—may well be false, while holding on to a somewhat weakened version of realism. This may be a sensible strategy, but it faces a number of challenges. First, the notion of approximate truth turns out to be surprisingly obscure: philosophers have not so far generated a general account of what it means to say that one falsehood is closer to the truth than another that is itself even approximately true. Second, it has been argued that even if we had a workable account of approximate truth, it would not apply to many of the most important theories in the history of science, since those theories, as the pessimist has noted, often contain central terms that refer to nothing, and so it seems those theories could not be even approximately true. Third, the truth-as-asymptote version of realism appears to demand a fairly smooth trajectory of scientific development, but Kuhn and others have maintained that this condition is not met by the actual history of science. Thus, while Einstein's physics is often presented as a refinement of Newton's, making it easy to imagine Newton as having given just a slightly less good approximation to the truth than did Einstein, this masks the radical conceptual transformation, a transformation too extreme to place the two theories on a common scale. (This is another sense of Kuhn's term 'incommensurable'.)

Another family of responses to the pessimistic induction is even more concessive, giving up on realism but looking for the nearest defensible alternative. According to these 'semi-realist' positions, we are not to affirm the truth (or even the approximate truth) of the full content of a scientific theory, but only of selected parts of it. Scientists, of course, often have different degrees of confidence in different components of the theories they consider, but the principle of selection that the semi-realists deploy in response to the pessimistic induction is different. They, in effect, look for a least common denominator in the history of science, by looking for aspects of theories that tend to be conserved across scientific changes, and advise us to place our truth bets there. Thus, it has been argued that although we may never be entitled to believe in the full truth of a scientific theory, we are entitled to believe in some of its structural claims, because these claims tend to be preserved as the rest of the theory changes (Worrall 1996). Here again, the soundness of the pessimistic induction has effectively been conceded: past theories have been false, so it is likely that future theories, taken in their entirety, will be false as well.

In my view, these approaches concede too much to the pessimistic induction. All of them as much as admit that the fact that old theories have turned out to be false makes it likely that all present and future theories will turn out to be false as well. These approaches then fall back on a weakened form of realism, retreating either to approximate truth or searching for components of past theories that have not been subsequently rejected. However, the conservative policy of a retreat to believing only the least common denominator across the history of science seems perverse, since in many cases our current theories seem strongest precisely where they diverge from their predecessors. Some retreat from unattenuated realism may be required for other reasons, but not because of the pessimistic induction, since that argument is in fact unsound.

Note first of all that the pessimistic induction is an exercise in what might be called 'judo epistemology'. The principle of judo is to use one's opponents' strength against them. (At least that is what I was told as a child: in practice, it always seemed rather that my opponents were using their strength all too effectively against me.) Similarly, in the pessimistic induction, we begin with current science and then judge past science to be false because it is incompatible with what we now believe. The negative judgement about past science is then parlayed by the pessimistic induction into the conclusion that current science is false as well. The strength of current science—its improvement over past science—is thus used against itself. This is suspicious. Similarly, it bears noting that the only empirical fact upon which the pessimistic induction ultimately depends is that the series of theories in a given area has been a series of contraries, a series of incompatible claims about the world, rather than a simple process of cumulative accretion. The incompatibility of past and present science is a plausible claim about much of the history of science, but is the fact that later theories have tended to correct earlier ones in itself really enough to provide a substantial reason for saying that science will never find the truth?

The pessimistic induction has the form of a simple inductive generalization: all observed As have been B, so all As are B. All past theories have been false, so all theories are false. An inference of this form is only warranted when there is no reason to suppose that the sample is biased, and in the case of the pessimistic induction there are several such reasons. For one, different theories are produced and tested in different data environments, and we may have much more relevant data for current theories than we did for past theories. The series of theories is not a set with independent members from which we have taken a random sample, for the series is the output of a process designed precisely to convert false theories into true ones. In this context, the fact that past theories have been false does not provide a reason to believe future ones will be false as well. Indeed, one might even argue that the history of science to which the pessimistic induction appeals points in the opposite direction. One does not need to accept Popper's extreme rejection of induction in order to agree with him that science is in the business of learning from its mistakes, and finding out what does not work may be an indispensable guide to finding out what does work. So one might go so far as to argue that we have more reason to believe some of our current theories in light of our extensive knowledge about how various earlier alternatives to them have failed, than we would if, by some miracle, the current theory had been the first one in its domain, a theory without a history.

7. THE MIRACLE ARGUMENT

Although it would be an exaggeration to claim that the history of science provides a powerful argument for realism, the pessimistic induction certainly does not refute the truth hypothesis. There is, however, a much-discussed argument directly for the truth hypothesis—known as the miracle argument for realism. It rests on a compelling intuition. Suppose that you are given a map of some region of rainforest, but initially have no way of knowing whether the map is accurate or not. Of course you cannot check every claim that the map makes, the rainforest being such a dense and difficult environment. However, you do what you can, and for every place in the rainforest that you do manage to check, things are exactly as the map says. What should you conclude from this? One possibility is that the map is wrong everywhere you have not checked, but that would require an evil miracle. What is far more likely is that the reason the map was accurate everywhere you checked is because the map is accurate generally. The same train of thought applies to science. Given a highly successful scientific theory with many and precise predictions that have been found to be correct (and none found to be incorrect), it remains logically possible that the rest of the theory is false. However, the only view that does not make this success a miracle is realism, the view that the theory is generally at least approximately correct, and not just where we happened to have checked (Putnam 1978). In other words, we are entitled to infer from predictive success to truth.

The miracle argument exerts a strong intuitive pull towards realism, especially in the context of highly successful mathematized theories that makes predictions of quite extraordinary precision and accuracy. Nevertheless, the argument has been beset with objections. One is that the miracle argument begs the question against anti-realists. The empirical success of our best scientific theories is admitted on all sides, but the opponents question what this shows. Put differently, it is admitted on all sides that scientists themselves cite evidence in favour of their preferred theories; what the philosopher does is obnoxiously ask what such evidence really shows. Answering that question would seem to require some independent evidence or argument, yet the miracle argument's appeal to the empirical success of a theory seems merely to reiterate scientific evidence, since that evidence seems nothing more or less than the theories' empirical successes.

Another line of objection to the miracle argument is to question just how big a miracle a false but successful theory would involve. Success, after all, is a temporal thing, and while the history of science may not support the destructive aspirations of the pessimistic induction, it certainly seems to provide instances of theories that were highly successful for a time, yet were eventually found to be false. Moreover, the Popperian process of filtering out theories as they fail makes it seem rather less surprising that the theories we now have, have so far been successful; had they not been, they would have been weeded out earlier (van Fraassen 1980). This selection mechanism appears to explain why our current theories are successful whether or not they are true. On the other hand, it is a peculiarity of selection mechanisms that they may explain why all members of a population have a trait without explaining why each member does. Thus the fact that this is a meeting of the red-haired league explains why everyone in the room has red hair; but it does not explain why Tim, who is at the meeting, has red hair. Similarly, while the selection mechanism may explain why all accepted theories are successful (if they had not been successful, they would not have been accepted), it does not explain why any particular theory has been successful. Theories are not successful because they have been selected: instead it is the other way round. However, the truth of a theory does explain its success.

Unfortunately, there is a third line of objection to the miracle argument that is even harder to shake, and this is that the argument ignores adverse base rates (Howson 2000). The miracle argument trades on the intuition that most false theories would have been unsuccessful. (Here one is considering the set of possible theories, not just those that have actually been formulated.) The intuition is correct, but the inference to truth is fallacious. Of course if all false theories were unsuccessful, then all successful theories would be true. However, from the fact that most false theories are unsuccessful, it just does not follow that most successful theories are true. One way to see this is to start not with falsehood, but with success. Given the constraint of success, we know that a true theory is one possibility; but given the underdetermination of theory by data, we know that there are also many false theories which would have enjoyed that same success. Most false theories would not meet the constraint, but many would, so alas it looks as though most successful theories are false! Success may be a good test in the sense that it has a low false positive rate, since most false theories are unsuccessful. However, that is not enough to show that most successful theories are true.

This tendency to infer from most false theories are unsuccessful to most successful theories are true is an instance of a general cognitive mistake that we almost all make: we tend to ignore base rates. Suppose a test for a disease has a false negative rate of nil: nobody who has the disease will test negative. The false positive rate is 5%: 5 out of 100 of the people who do not have the disease will nevertheless test positive, and 1 in 1000 people have the disease. You give your patient the test and the result is positive. What is the probability that your patient has the disease? When staff and students at Harvard Medical School were asked, most said 0.95. The correct answer is just under 0.02. If you give the test to 1000 people, approximately 50 of them will test positive even though they do not have the disease, as against the one poor guy who has the disease. The probability that your patient has the disease is much higher after the test—from 1 in 1000 to approximately 1 in 50—but that is still fairly unlikely. The medics who thought that the probability is 0.95 were wildly out because they ignored the base rate, which should have told them that results for healthy people are swamping results for the ill. Beware of even a very low false positive rate, if the condition is rare.

The miracle argument for realism works the same way, with realists playing the role of the probabilistically challenged medics. On this analogy, being true is the disease and making lots of true and precise predictions is testing positive for it. The false negative rate is nil: no true theory makes false predictions. The false positive rate is low, since relatively very few false theories are so empirically successful. Hence we are inclined to infer that highly successful theories are likely to be true. What we ignore is the base rate that, to put it crudely, the vast majority of theories are false, so even a very small probability that a false theory should make such successful predictions means that the great majority of successful theories are false. Most false theories are unsuccessful, but alas what counts, is that most successful theories are false. Beware even the low false positives rate of the success test, because truth is a rare condition.

This is an interesting and disturbing analysis of our situation, presenting a powerful objection to the miracle argument. Why are so many of us prone to commit a base-rate fallacy here? We are apparently inclined to compare the relative number of successful false theories with unsuccessful false theories, but not with successful true theories, although it is the latter comparison that matters. That philosophers of science should suffer from this cognitive blindspot is particularly surprising, given how familiar we are with the problems of underdetermination. Medics act as if the test makes the prior probability irrelevant. Something similar may explain our tendency to be impressed by the miracle argument. What seems to be going on is that the miracle argument encourages us to assess the reliability of empirical success as a test for truth by estimating its false-positive rate (the chance that a false theory is successful), a rate we rightly judge to be very low. However, we ignore how incredibly unlikely it would be that, prior to testing, a given theory should be true. This has the effect of hiding from our view all those other theories that would be just as successful even though they are false. Our perception, however, is easily altered. Underdetermination arguments force us to face up to the existence of innumerable false, but successful, possible theories, and our intuitions flip.

8. THE CONSOLATIONS OF PHILOSOPHY

The status of the truth hypothesis, and so of realism, thus remains unsettled: it is neither undermined by the pessimistic induction nor confirmed by the miracle argument. Nor do I know of any other arguments that even come close to closing the question. This suggests that the rational attitude towards a scientific theory should never rise above the level of agnosticism. This view is in fact very close to what Popper recommends, because of his specific rejection of the notion of supporting evidence and because of his generalized hostility to the notion of belief. In my view, however, this is not realistic. As Hume said of the sceptical arguments in philosophy, 'they admit of no answer and produce no conviction' (1748). Philosophers may give brilliant lectures establishing that we have no more reason to think that we are safer leaving the lecture theatre by the door than by the window, but they may be sure that everyone will still leave by the door. What Hume called our 'natural instincts' are stronger than any philosophical argument. In my view this applies to science as well. Scientists are spontaneous realists and in many cases cannot but immerse themselves in the world of theory, believing a great deal of what they say. To be sure, they are capable of a limited agnosticism, treating a particular hypothesis as only a conjecture or as a model, but a comprehensive refusal to believe any scientific claim not entailed by the data seems nearly a psychological impossibility.

Scientists' natural realism is neither surprising nor pernicious. Science is a practice that has assertion making at its heart, and to assert a claim is to say the claim is true. Moreover, scientific inferences are often driven by asking how good an explanation a given hypothesis would provide of the phenomena if it *were* true, and then, if it would be good enough, inferring that *is* true (Lipton 2004). Indeed, the activity of explanation itself supposes truth. Science aims to explain why, and an actual explanation is not just a good story but a good story that is also *true*. Thus, in so far as scientists take themselves to be providing actual explanations of the phenomena, they are taking themselves to be saying things that are true.

Perhaps all this assertive practice arises from systematic overconfidence, with scientists systematically believing and claiming more than the evidence could warrant. Even if this is the case, it may not be as bad as it sounds. Overconfidence may be the price for almost any cognitive activity, and it may encourage the enthusiastic defence of ideas that are worthwhile but otherwise would have died a premature death. Of course, it may also lead to dogmatism, but probably not if the scientists in the next laboratory are equally confident in an incompatible hypothesis. The dispute between two confident but opposed groups may promote greater progress than a uniform agnosticism. Finally, even dogmatism may have an essential role in scientific development. Kuhn (1963) has argued that it is only through the unquestioning acceptance of a great deal of theory (if only temporarily) that the kind of esoteric and articulated research which characterizes mature science is possible at all. The philosopher has many alternatives to realism, but the practicing scientist may not.

Where does this leave the relationship between the scientists and the philosophers of science? I have suggested that the problem is not that philosophers put science down; nor is it that they fail to give direct, practical advice. Perhaps the real problems are that philosophers never seem to settle their own arguments and, in so far as their argument yields any result, it is one of generalized agnosticism that is not a live option for scientists embedded in their practices. Indeed, one might go so far as to worry that if philosophy did have any impact on scientists, it would be pernicious, depriving them of the kinds of commitment and confidence upon which their practice depends. Perhaps there are still some areas of the philosophy of science that science ought to consider—such as the ethical work on science and the close conceptual engagement with particular current scientific theories—but why should scientists pay any attention to the sort of philosophical work I have discussed in this essay?

There is no obligation for scientists to consider this sort of philosophy. Just as people can use their eyes to acquire knowledge without understanding how vision works, so scientists can generate knowledge about the world without understanding how science works. What is required for knowledge is the de facto reliability of the methods we deploy, not our understanding of how those methods work or some demonstration of their reliability. If scientific practices are taking us towards the truth, then they are generating knowledge of a mind-independent world that is also independent of any philosophical account or justification; and if those practices are not taking us to the truth, that situation will not be reversed by philosophy. On the other hand, there is little risk that the study of philosophy will handicap scientists by turning them into destructive sceptics about their own practices. Scientists' natural instincts are too strong for this.

However, it is not just that the philosophy of science is safe for scientists. A little of it may even do you good. Like spending time in another culture, the pursuit of the philosophy of science, and of science studies generally, helps to reveal contingencies in scientific practices that may look like necessities from within the practices themselves. Kuhn may well be right that various forms of limited-term dogmatism are important to scientific development, but he (and here Popper would agree) is also right that many certainties have a limited shelf life. Without going as far as the American president who said that economic growth is inevitable, but that could change, there may an advantage to having a limited form of cognitive schizophrenia in science, an advantage to being able to immerse fully in the world of the scientific commitments of the moment while also being able to stand back and see that there may well arise circumstances where those deep commitments ought to be abandoned. 'Often wrong but never in doubt' is a strategy with certain cognitive advantages, but there may also be advantages to being self-aware. Kuhn argues that there is an 'essential tension' between conservatism and innovation in scientific practice, because 'the scientist requires a thoroughgoing commitment to the tradition with which, if he is fully successful, he will break' (Kuhn 1977, p. 235). My suggestion is that a dose of philosophy may help scientists to maintain that tension in a healthy state.

The virtues of self-awareness also provide a more general justification for an interest in the philosophy of science. Even if scientific self-understanding helped not at all in improving scientific performance, it is still worthwhile. It addresses a healthy and natural curiosity. Scientists are people who want to understand how things work, and it is natural for them to wish to turn this curiosity back on themselves, and so to wish to understand how their own enterprise works. The achievements of the philosophy of science in illuminating the nature of science should not be oversold, but my hope is that this essay goes some way towards showing why the project of trying to explain science is worthwhile.

REFERENCES

Bacon, F. 1620 Novum Organum. [Reprinted in *The new organon and related writings* (ed. F. H. Anderson). Bobbs-Merrill (Indianapolis, IN) 1960.]

Duhem, P. 1914 *The aim and structure of physical theory.* [Reprinted by Princeton University Press 1954.]

Howson, C. 2000 *Hume's problem*. Oxford University Press.

Hume, D. 1748 *An enquiry concerning human understanding*. [Reprinted by Oxford University Press 1999.]

Kahneman, D., Slovic, P. & Tversky, A. 1982 *Judgment under uncertainty: heuristics and biases*. Cambridge University Press.

Kant, I. 1783 *Prolegomena to any future metaphysics*. [Reprinted by Hackett (Indianapolis, IN) 1977.]

Kuhn, T. 1963 The function of dogma in scientific research. In *Scientific change* (ed. A. Crombie), pp. 347–369. Oxford University Press.

Kuhn, T. 1970 *The structure of scientific revolutions*, 2nd edn. University of Chicago Press.

Kuhn, T. 1977 *The essential tension: tradition and innovation in scientific research*. *The essential tension*, pp. 225–239. University of Chicago Press.

Laudan, L. 1984 A confutation of convergent realism. In *Scientific realism* (ed. J. Leplin), pp. 218–249. Berkeley: University of California Press.

Lipton, P. 2004 *Inference to the best explanation*, 2nd edn. London: Routledge.

Locke, J. 1689 *An essay concerning human understanding*. [Reprinted by Oxford University Press 1975.]

Medawar, P. 1982 *Pluto's republic*. Oxford University Press.

Popper, K. 1959 *The logic of scientific discovery.* London: Hutchinson.

Popper, K. 1962 *Conjectures and refutations*. London: Routledge.

Popper, K. 1972 *Objective knowledge*. Oxford University Press.

Putnam, H. 1978 *Meaning and the moral sciences*. London: Hutchinson.

Quine, W. V. O. 1951 Two dogmas of empiricism. *Phil. Rev.* **60**, 20–43.

Sklar, L. 2000 Interpreting theories: the case of statistical mechanics. *Brit. J. Phil. Sci.* **51**, 729–742.

van Fraassen, B. C. 1980 *The scientific image*. Oxford University Press.

Worrall, J. 1996 Structural realism: the best of both worlds? In *Philosophy of science* (ed. D. Papineau), pp. 139–165. Oxford University Press.