PART I

Scientific realism today

CHAPTER I

Realism and antirealism; metaphysics and empiricism

I.I THE TROUBLE WITH COMMON SENSE

Hanging in my office is a framed photograph of an armillary sphere, which resides in the Whipple Museum of the History of Science in Cambridge, England. An armillary sphere is a celestial globe. It is made up of a spherical model of the planet Earth (the sort we all played with as children), but the model is surrounded by an intricate skeleton of graduated rings, representing the most important celestial circles. Armillary spheres were devised in ancient Greece and developed as instruments for teaching and astronomical calculation. During the same period, heavenly bodies were widely conceived as fixed to the surfaces of concentrically arranged crystalline spheres, which rotate around the Earth at their centre.

This particular armillary sphere has, I expect, many fascinating historical stories to tell, but there is a specific reason I framed the picture. Once upon a time, astronomers speculated about the causes and mechanisms of the motions of the planets and stars, and their ontology of crystalline spheres was a central feature of astronomical theory for hundreds of years. But crystalline spheres are not the sorts of things one can observe, at least not with the naked eye from the surface of the Earth. Even if it had turned out that they exist, it is doubtful one would have been able to devise an instrument to detect them before the days of satellites and space shuttles. Much of the energy of the sciences is consumed in the attempt to work out and describe things that are inaccessible to the unaided senses, whether in practice or in principle. My armillary sphere, with its glorious and complicated mess of interwoven circles, is a reminder of past testaments to that obsession.

In describing the notion of a crystalline sphere, I have already made some distinctions. There are things that one can, under favourable circumstances, perceive with one's unaided senses. Let us call them "observables", though this is to privilege vision over the other senses for the sake of terminological convenience. Unobservables, then, are things one cannot perceive with one's unaided senses, and this category divides into two subcategories. Some unobservables are nonetheless detectable through the use of instruments with which one hopes to "extend" one's senses, and others are simply undetectable. These distinctions are important, because major controversies about how to interpret the claims of the sciences revolve around them. In this chapter, I will briefly outline the most important positions engaged in these controversies, and consider how the tension between speculative metaphysics and empiricism has kept them alive.

There are occasional disputes about what counts as science – concerning how best to exclude astrology but include astronomy, about what to say to creationists unhappy with the teaching of evolutionary biology in schools, etc. I leave these disputes to one side here, and begin simply with what are commonly regarded as sciences today. It is widely held that the sciences are not merely knowledge-producing endeavours, but *the* means of knowledge production *par excellence*. Scientific inquiry is our best hope for gaining knowledge of the world, the things that compose it, its structure, its laws, and so on. And the more one investigates, the better it gets. Scientific knowledge is progressive; it renders the natural world with increasing accuracy.

Scientific realism, to a rough, first approximation, is the view that scientific theories correctly describe the nature of a mind-independent world. Outside of philosophy, realism is usually regarded as common sense, but philosophers enjoy subjecting commonplace views to thorough scrutiny, and this one certainly requires it. The main consideration in favour of realism is ancient, but more recently referred to as the 'miracle argument' (or 'no-miracles argument') after the memorable slogan coined by Hilary Putnam (1975, p. 73) that realism 'is the only philosophy that doesn't make the success of science a miracle'. Scientific theories are amazingly successful in that they allow us to predict, manipulate, and participate in worldly phenomena, and the most straightforward explanation of this is that they correctly describe the nature of the world, or something close by. In the absence of this explanation the success afforded by the sciences might well seem miraculous, and, given the choice, one should always choose common sense over miracles.

Some have questioned the need for an explanation of the success of science at all. Bas van Fraassen (1980, pp. 23–5, 34–40), for example, suggests that successful scientific theories are analogous to well-adapted organisms. There is no need to explain the success of organisms, he says.

Only well-adapted organisms survive, just as only well-adapted theories survive, where 'well-adapted' in the latter case means adequate to the tasks to which one puts theories. These tasks are generally thought to include predictions and retrodictions (predictions concerning past phenomena), and perhaps most impressively novel predictions (ones about classes of things or phenomena one has yet to observe). A well-adapted theory is one whose predictions, retrodictions, and novel predictions, if any, are borne out in the course of observation and experimentation. But saying that successful theories are ones that are well-adapted may be tantamount to the tautology that successful theories are successful, which is not saying much. Whatever the merits of the Darwinian analogy for theories generally, one might still wonder why any *given* theory (organism) survives for the time it does, and this may require a more specific consideration of the properties of the theory (organism) in virtue of which it is well adapted. I will return to the contentious issue of the demand for explanations later in this chapter.

The attempt to satisfy the desire for an explanation of scientific success has produced the bulk of the literature on scientific realism. As arguments go, the miracle argument is surprisingly poor, all things considered, and consequently alternatives to realism have flourished. The poverty of the miracle argument and consequent flourishing of rivals to realism stem from difficulties presented by three general issues, which I will mention only briefly:

1 the use of abductive inference, or inference to the best explanation (IBE)
2 the underdetermination of theory choice by data or evidence (UTD)
3 discontinuities in scientific theories over time, yielding a pessimistic induction (PI)

Abduction is a form of inference famous from the writings of Charles Saunders Pierce, inspiring what is now generally called 'inference to the best explanation' (some use the term synonymously with 'abduction' while others, more strictly, distinguish it from Pierce's version). IBE offers the following advice to inference makers: infer the hypothesis that, if true, would provide the best explanation for whatever it is you hope to explain. Note that the miracle argument itself is an abductive argument. Why are scientific theories so successful at making predictions and accounting for empirical data? One answer is that they are true, and this seems, to the realist at any rate, the best explanation. One might even think it the only conceivable explanation, but as we shall see, in light of UTD and PI, this is highly contestable. First, however, let us turn from the particular case of the miracle argument to the merits of IBE as a form of inference in general. There is little doubt that this sort of inferential practice is fundamental to everyday and scientific reasoning. The decision to adopt one theory as opposed to its rivals, for example, is generally a complex process involving many factors, but IBE will most certainly figure at some stage.

Antirealists are quick to point out that in order for an instance of IBE to yield the truth, two conditions must be met. Firstly, one must rank the rival hypotheses under consideration correctly with respect to the likelihood that they are true. Secondly, the truth must be among the hypotheses one is considering. But can one ensure that these conditions are met? Regarding the first, it is difficult to say what features a truth-likely explanation should have. Beyond the minimum criterion of some impressive measure of agreement with outcomes of observation and experiment, possible indicators of good explanations have been widely discussed. Some hold that theories characterized by features such as simplicity, elegance, and unity (with other theories or domains of inquiry) are preferable. Quite apart from the matter of describing what these virtues are, however, and knowing how to compare and prioritize them, it is not immediately obvious that such virtues have anything to do with truth. There is no *a priori* reason, one might argue, to reject the possibility that natural phenomena are rather complex, inelegant, and disjoint. And regarding the second condition for successful IBE, in most cases it is difficult to see how one could know in advance that the true hypothesis is among those considered.[1]

In practice it is often difficult to produce even one theory that explains the empirical data, let alone rivals. This, however, does not diminish the seriousness of the problem. In fact, it turns out that it may be irrelevant whether one ever has a choice to make between rival theories in practice. For some maintain that rival theories are always possible, whether or not one has thought of them, and this is sufficient to raise concerns about IBE. Confidence in the possibility of rivals stems from the underdetermination thesis, or UTD. Its canonical formulation due to Pierre Duhem, later expressed in rather different terms by W. V. O. Quine (hence also called the 'Duhem-Quine thesis'), goes this way. Theoretical hypotheses rarely if ever yield predictions by themselves. Rather, they must be conjoined with auxiliary hypotheses – background theories, related theories, theories

[1] A case in which one does have this knowledge is where rival hypotheses are contradictories. See Lipton 1993 for a discussion of this and its implications for IBE.

about the measurement of relevant parameters, etc. – in order to yield predictions. If observation and experimentation produce data that are not as one predicts, one has a choice to make concerning which of the prediction-yielding hypotheses is culpable. One can always preserve a favoured hypothesis at the expense of something else. Since there are different ways of choosing how to account for recalcitrant data, different overall theories or conjunctions of hypotheses may be used to account for the empirical evidence. Thus, in general, there is always more than one overall theory consistent with the data.

In more contemporary discussions, UTD is usually explicated differently. Given a theory, T_1, it is always possible to generate an empirically equivalent but different theory, T_2. T_2 is a theory that makes precisely the same claims regarding observable phenomena as T_1, but differs in other respects. T_2 might, for example, exclude all of the unobservable entities and processes of T_1, or replace some or all of these with others, or simply alter them, but in such a way as to produce exactly the same observable predictions. Given that this sort of manoeuvring is always possible, how does one decide between rival theories so constructed? Here again the realist must find a way to infer to a particular theory at the expense of its rivals, with the various difficulties this engenders.

In addition to challenges concerning IBE and UTD, at least one anti-realist argument aspires to the status of an empirical refutation of realism. PI, or as it is often called, the 'pessimistic meta-induction', can be summarized as follows. Consider the history of scientific theories in any particular domain. From the perspective of the present, most past theories are considered false, strictly speaking. There is evidence of severe discontinuity over time, regarding both the entities and processes described. This evidence makes up a catalogue of instability in the things to which theories refer.[2] By induction based on these past cases, it is likely that present-day theories are also false and will be recognized as such in the future. Realists are generally keen to respond that not even they believe that theories are true *simpliciter*. Scientific theorizing is a complex business, replete with things like approximation, abstraction, and idealization. What is important is that successive theories get better with respect to the truth, coming closer to it over time. It is the progress sciences make in describing nature with increasing accuracy that fuels realism. Good theories, they say, are normally "approximately true", and more so as the

[2] Perhaps the most celebrated vision of discontinuity is found in Kuhn 1970/1962. More recent discussions often focus on the formulations of PI given in Laudan 1981.

sciences progress. Giving a precise account of what 'approximate truth' means, however, is no easy task.

So much for common sense. The promise of scientific realism is very much open to debate, and in light of IBE, UTD, and PI, this debate has spawned many positions. Let us take a look at the main players, so as to gain a better understanding of the context of realism.

1.2 A CONCEPTUAL TAXONOMY

Earlier I described realism as the view that scientific theories correctly describe the nature of a mind-independent world. This is shorthand for the various and more nuanced commitments realists tend to make. For example, many add that they are not realists about all theories, just ones that are genuinely successful. The clarification is supplied to dissolve the potential worry that realists must embrace theories that seem artificially successful – those that do not make novel predictions and simply incorporate past empirical data on an *ad hoc* basis, for instance. Realists often say that their position extends only to theories that are sufficiently "mature". Maturity is an admittedly vague notion, meant to convey the idea that a theory has withstood serious testing in application to its domain over some significant period of time, and some correlate the maturity of disciplines more generally with the extent to which their theories make successful, novel predictions.[3] Finally, as I have already mentioned, it is also standard to qualify that which theories are supposed to deliver: it is said that theoretical descriptions may not be true, *per se*, but that they are nearly or approximately true, or at least more so than earlier descriptions.

With these caveats in mind it may be instructive to situate scientific realism in a broader context, as a species of the genus of positions historically described as realisms. Traditionally, 'realism' simply denotes a belief in the reality of something – an existence that does not depend on minds, human or otherwise. Consider an increasingly ambitious sequence of items about which one might be a realist. One could begin with the objects of one's perceptions (goldfish, fishbowls), move on to objects beyond one's sensory abilities to detect (genes, electrons), and further still, beyond the realm of the concrete to the realm of the abstract, to non-spatiotemporal things such as numbers, sets, universals, and propositions. The sort of realist one is, if at all, can be gauged from the sorts of things one

[3] See Worrall 1989, pp. 153–4, on the notions of maturity and *ad hoc*ness, Psillos 1999, pp. 105–8, on *ad hoc* theories and novel predictions, and Leplin 1997 on novel predictions.

takes to qualify for mind-independent existence. Though I have just described these commitments as forming a sequence, it should be understood that realism at any given stage does not necessarily entail realism about anything prior to that stage. Some Platonists, for example, appear to hold that ultimately, the only real objects are abstract ones, the Forms, or that the Forms are in some sense "more real" than observables.[4] Scientific realism, in committing to something approaching the truth of scientific theories, makes a commitment to their subject matter: entities and processes involving their interactions, at the level of both the observable and the unobservable. Anything more detailed is a matter for negotiation, and realists have many opposing views beyond this shared, minimal commitment. My own more detailed proposals for realism are outlined in the chapters to come.

I said that 'realism' traditionally denotes a belief in the reality of something, but in the context of scientific realism the term has broader connotations. The most perspicuous way of understanding these aspects is in terms of three lines of inquiry: ontological, semantic, and epistemological. Ontologically, scientific realism is committed to the existence of a mind-independent world or reality. A realist semantics implies that theoretical claims about this reality have truth values, and should be construed literally, whether true or false. I will consider an example of what it might mean to construe claims in a non-literal way momentarily. Finally, the epistemological commitment is to the idea that these theoretical claims give us knowledge of the world. That is, predictively successful (mature, non-*ad hoc*) theories, taken literally as describing the nature of a mind-independent reality are (approximately) true. The things our best scientific theories tell us about entities and processes are decent descriptions of the way the world really is. Henceforth I will use the term 'realism' to refer to this scientific variety only. We are now ready to locate it and various other positions in a conceptual space.

If by 'antirealism' one means any view opposed to realism, many different positions will fit the bill. Exploiting differences in commitments along our three lines of inquiry, one may construct a taxonomy of views discussed in connection with these debates. Table 1.1 lists the most prominent of these, and for each notes how it stands on the existence of a mind-independent world, on whether theoretical statements should be taken literally, and on whether such claims yield knowledge of their putative

[4] For a nice summary of the connections between scientific and other realisms, see Kukla 1998, pp. 3–11.

Table 1.1. *Scientific realism and antirealisms*

	The ontological question: mind-independent reality?	The semantic question: theories literally construed?	The epistemological question: knowledge?
Realism	yes	yes	yes
Constructive empiricism	yes	yes	observables: yes unobservables: no
Scepticism	yes	yes	no
Logical positivism/empiricism	yes/no/?	observables: yes unobservables: no	yes
Traditional instrumentalism	yes	observables: yes unobservables: no	observables: yes unobservables: no
Idealism	no	no	yes

subject matter. This is a blunt instrument; an impressive array of viewpoints is not adequately reflected in this simple classificatory scheme, and the reflections present are imprecise. There are many ways, for example, in which to be a sceptic. But the core views sketched in Table 1.1 offer some basic categories for locating families of related commitments.

Traditionally and especially in the early twentieth century, around the time of the birth of modern analytic philosophy, realist positions were contrasted with idealism, according to which there is no world external to and thus independent of the mental. The classic statement of this position is credited to Bishop George Berkeley, for whom reality is constituted by thoughts and ultimately sustained by the mind of God. Idealism need not invoke a deity, though. A phenomenalist, for instance, might be an idealist without appealing to the divine. Given an idealist ontology, it is no surprise that scientific claims cannot be construed literally, since they are not about what they seem to describe at face value, but this of course does not preclude knowledge of a mind-*dependent* reality. As Table 1.1 shows, idealism is the only position considered here to take an unambiguous antirealist stand with respect to ontology.

Instrumentalism is a view shared by a number of positions, all of which have the following contention in common: theories are merely instruments for predicting observable phenomena or systematizing observation reports. Traditional instrumentalism is an even stronger view according to which, furthermore, claims involving unobservable entities and processes have no meaning at all. Such 'theoretical claims', as they are called ('claims about unobservables' is better, I think, since theories describe observables too), do not have truth values. They are not even capable of being true or false; rather, they are mere tools for prediction. In common usage, however, some now employ the term in a weaker sense, to describe views that grant truth values to claims involving unobservables while maintaining that one is not in a position, for whatever reason, to determine what these truth values are. In this latter, weaker sense, constructive empiricism is sometimes described as a form of instrumentalism. And though I have represented instrumentalists in Table 1.1 as subscribing to realism in ontology, some would include those who do not.

Logical positivism, famously associated with the philosophers and scientists of the Vienna Circle, and its later incarnation, logical empiricism, are similar to traditional instrumentalism in having a strict policy regarding the unobservable. But where traditional instrumentalism holds that claims about unobservables are meaningless, logical empiricism assigns meaning to some of these claims by interpreting them non-literally. Rather than taking these claims at face value as describing the things they appear to describe, claims about unobservables are meaningful for logical empiricists if and only if their unobservable terms are linked in an appropriate way to observable terms. The unobservable vocabulary is then treated as nothing more than a shorthand for the observation reports to which they are tied. 'Electron', for example, might be shorthand in some contexts for 'white streak in a cloud chamber', given the path of water droplets one actually sees in a cloud chamber experiment, along what is theoretically described as the trajectory of an electron. It is by means of such 'correspondence rules' or 'bridge principles' that talk of the unobservable realm is interpreted. Given a translation manual of this sort, theories construed non-literally are thought to yield knowledge of the world. The label 'logical positivism / empiricism' covers vast ground, however, and views regarding the ontological status of the world described by science are far from univocal here. Rudolph Carnap (1950), for instance, held that while theories furnish frameworks for systematizing knowledge, ontological questions 'external' to such frameworks are meaningless, or have no cognitive content.

While traditional instrumentalism banishes meaningful talk about unobservables altogether and logical empiricism interprets it non-literally, constructive empiricism, the view advocated by van Fraassen, adopts a realist semantics. The antirealism of this latter position is thus wholly manifested in its epistemology. For the constructive empiricist the observable–unobservable distinction is extremely important, but only in the realm of knowledge, and this feature marks the position as an interesting half-way