A METAPHYSICS FOR SCIENTIFIC REALISM

Knowing the Unobservable

ANJAN CHAKRAVARTTY

University of Toronto

For Arthur,
Who NOAs best,
Anjan

CAMBRIDGE
UNIVERSITY PRESS

CHAPTER 1

Realism and antirealism; metaphysics and empiricism

1.1 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 antirealist 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

house between realism and various kinds of scepticism. By scepticism here, I intend any position that agrees with the realist concerning ontology and semantics, but offers epistemic considerations to suggest that one does not have knowledge of the world, or at least that one is not in a position to know that one does. Constructive empiricism goes along with the sceptic part way, denying that one can have knowledge of the unobservable, but also with the realist part way, accepting that one can have knowledge of the observable. (More strictly, constructive empiricism is the view that the aim of science is true claims about observables, not truth more generally, but this is usually interpreted in the way I have suggested.) By adopting a realist semantics, constructive empiricism avoids the semantic difficulties that were in large part responsible for the demise of logical empiricism in the latter half of the twentieth century, and has taken its place as the main rival to realism today.

Table 1.1 does not exhaust the list of "isms" opposed to realism. It does, however, provide a fairly comprehensive list of the reasons and motivations one might have for being an antirealist. For example, the discipline known as the sociology of scientific knowledge is predominantly antirealist. This is not a logical consequence, however, of the desire to study science from a sociological perspective. Sociologists who are antirealists are usually so inclined because of commitments they share with one or more of the antirealist positions outlined in Table 1.1. Though I will not consider this approach to thinking about the sciences in any detail here, it is important to appreciate its influence. Sociological and related methodologies, which attempt to explicate scientific practice and its social, political, and economic relations, both internal and external, represent the major alternative approach to the study of the sciences today, contrasting with the more straightforwardly philosophical approach of realism and constructive empiricism.

Two last positions are worthy of note here, the first of which is actually a family of views belonging to the tradition of pragmatism. This is perhaps the most difficult position to situate with respect to realism, given that most pragmatists would answer 'yes' to all three of the questions posed in Table 1.1, but only some claim to be realists. The difficulty here is that pragmatists adopt a theory of truth that many see as incompatible with realism. For them, truth is an epistemic concept. To say that a statement or theory is true, or that it offers a correct description of the world, is simply to say that it has positive utility – it is useful in some way to believe it. Others hold that truths are what one would believe under epistemically ideal conditions, or in the ideal limit of inquiry. Many realists, however, are uncomfortable with epistemic theories of truth, and adopt instead some version of the correspondence theory, according to which truth is some sort of correspondence between things like theories and the world. But it is doubtful whether one must adopt a correspondence theory to be a realist. There are difficulties associated with explaining what correspondence means, and many prefer to do without. In any case, it does seem that in order to qualify as a realist, one must believe that good theories are reasonably successful in describing the nature of a mind-independent world, but whether this is understood in terms of correspondence truth or in some other way (for example, in terms of a theory of reference or representation) is an open question.

To complete this brief roundup, let me mention what Arthur Fine (1996, pp. 112–50) calls 'NOA', the natural ontological attitude. NOA shares certain motivations with pragmatism, though in addition to rejecting correspondence theories, it rejects all theories of truth including epistemic ones. Its most striking feature is a form of quietism with respect to issues concerning the unobservable that realists and antirealists are wont to contest. As an alternative, NOA prescribes a policy of non-engagement: all ontological claims are on a par, whether about observables or unobservables; beyond merely accepting statements regarding elephants and electrons (as both realists and antirealists do), one should refrain from interpreting such claims by adding that both sorts of objects are real, or that talking about electrons is simply a shorthand for talking about something observable, and so on. NOA rejects both realism *and* antirealism. It is intended as a neutral position for those who find nothing to be gained in debates surrounding them. From the perspective of these debates, however, NOA may seem too anti-philosophical a stand to take. Leaving aside intriguing questions about the potential value and cogency of quietism in this context, I will not consider it further here.

1.3 METAPHYSICS, EMPIRICISM, AND SCIENTIFIC KNOWLEDGE

Armed with a basic summary of realism and its principal rivals, let us turn to the central focus of this work. Earlier I said that much of the controversy surrounding these positions concerns the question of how one should understand scientific claims, in light of the distinctions between the observable and unobservable on one hand, and between the two categories of the unobservable, the detectable and undetectable, on the other. By examining these distinctions one may begin to shed some light

on the roles that metaphysics and empiricism play in the interpretation of scientific claims, and the dialectic between them.

The first distinction, between observables and unobservables, concerns things that one can under favourable circumstances perceive with one's unaided senses, and things one cannot. Note that this use of 'observable' and 'unobservable' is different from what is often the case in the sciences themselves. In scientific practice the label 'observable' is usually applied permissively to anything with which one can forge some sort of causal contact, as one does when one uses instruments (such as microscopes) for detection.[5] In the present discussion, however, observables are strictly things one can perceive with the unmediated senses. As Table 1.1 attests, almost everyone thinks one can have knowledge of the observable. This is not to say, however, that interpreting claims about observables is necessarily straightforward. It may be, for example, that the categories of objects and processes one employs to express one's knowledge of the observable are interestingly shaped by the theories one adopts. Indeed, that this is the case for both observables and unobservables is a central tenet of the influential views of Thomas Kuhn (1970/1962), the sociological approaches that followed him, and even some of the logical empiricists who preceded him, who held that "conceptual schemes" shape one's knowledge of the world. I will return to this issue in Chapter 6, but otherwise, for the most part, will take the idea that one has knowledge of observables for granted.

It is the status of the unobservable that has proved most controversial. Logical positivism was, in effect, the founding movement of modern philosophy of science, and the radical empiricism of the positivists has had a lasting impact. It will be useful in what follows to clarify my second distinction, between unobservables that are detectable and those that are not. Let me reserve the word 'detectable' for unobservables one can detect using instruments but not otherwise, and 'undetectable' for those one cannot detect at all (see Figure 1.1). The mitochondrion, for example, is a cellular organelle in which substances are oxidized to produce energy. Though unobservable, one can detect mitochondria using microscopy. A celebrated historical example of a more indirect case of detection is the neutrino, a subatomic particle originally posited by Wolfgang Pauli and theorized about by Enrico Fermi in the 1930s. The neutrino was hypothesized to allow for the conservation of mass-energy and angular

[5] See Shapere 1982 for a discussion of the differences between philosophers' and scientists' notions of observation. Shapere examines the conditions under which astrophysicists speak of "observing" solar neutrinos, and also (amazingly) core regions of stars, by means of neutrino detection.

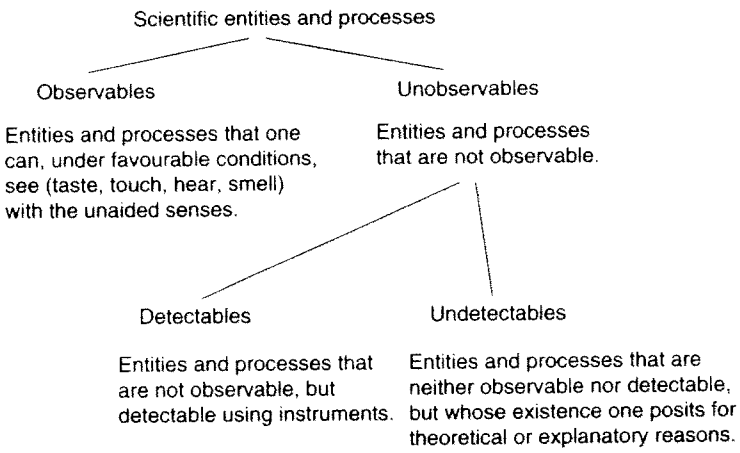

Figure 1.1. Observables and unobservables

momentum in certain subatomic interactions, such as the β-decay of radium-210, and detections of such interactions might thus be viewed as indirect detections of neutrinos. It was not until 1956 that Frederick Reines and Clyde Cowan successfully performed an experiment in which neutrinos were detected more directly. Now consider unobservables whose putative existence cannot be the subject of empirical investigation, whether in practice or in principle. Examples include Newton's conceptions of position and velocity with respect to absolute space, and causally inefficacious entities such as mathematical objects. Even if they exist, such things are undetectable.

Historically, the most pressing challenges to realism have come from those adopting some form of empiricism. This is not to say, however, that all empiricists are antirealists! It may be helpful here to note that empiricism is traditionally associated with two strands of thought which often come together, interwoven. One strand is the idea that sensory experience is the *source* of all knowledge of the world, and this by itself does not preclude an empiricist from being a realist. A realist might accept this first strand while further believing that one can infer the existence of certain unobservables on the basis of the evidence of one's senses. The second strand of empiricism is the idea that all knowledge of the world is *about* experience, and it is this tenet that conflicts with realism, since realists believe claims about things that transcend experience in addition to claims about observables. So an empiricist of the first strand alone may be

a realist, but not one of the first and second strands combined. The most adamant critiques of realism stem from those who are committed to the second strand of empiricism, and to violate this commitment is to engage in what its advocates view as a fruitless and misconceived philosophical activity: speculative metaphysics.

What is this metaphysics, then, of which so many empiricists disapprove? To say that there is a conflict between metaphysics and empiricism *simpliciter* is too strong, since many empiricists do metaphysics as it is understood most broadly, as the study of the first or basic principles of philosophy, being *qua* being, and the natures of things that exist. The metaphysics that empiricists disavow concerns the unobservable, and thus any position that endorses speculation of this sort, leading to substantive beliefs about detectables or undetectables, is unacceptable to them. This includes not only speculations about things like universals and causal necessity, which are familiar topics within metaphysics, but also speculations about mitochondria and neutrinos, which are familiar topics within the sciences. But empiricists are generally happy to do metaphysics so long as it does not involve believing speculations about the unobservable. Thus Hume gives an account of causation, not in terms of undetectable necessary connections, but solely in terms of observable events that follow one another. And thus nominalists speak of properties, not as abstract entities like universals, but as sets of observable things to which the predicates associated with these properties apply. The unobservable is likewise an anathema to many empiricist accounts of science. The scientific realist, in maintaining that one can have knowledge of scientific unobservables, engages in the very sort of metaphysical speculation these empiricists reject.[6]

Logical positivism and logical empiricism lost their way, but constructive empiricism has emerged as the main empiricist rival to realism today. Van Fraassen argues for a reconceptualization of empiricism, one of whose goals is to demonstrate the superiority of empiricism over speculative metaphysics. In the remainder of this chapter I will consider his recasting of empiricism, and the question of whether it succeeds in the task of banishing its old adversary. I will argue that it does not, thus opening

[6] Van Fraassen (1980, p. 8) defines realism in terms of aspiration: 'Science aims to give us, in its theories, a literally true story of what the world is like; and acceptance of scientific theory involves the belief that it is true.' On this view realism is not necessarily metaphysical, since one might adopt it without endorsing the approximate truth of any claims about unobservables. It seems to me that this is too weak. Realists do believe claims about unobservables, subject to the various caveats I have described, and consequently realism is (on my view) metaphysical.

the door to a detailed consideration of the foundational beliefs of a thoroughly updated scientific realism.

1.4 THE RISE OF STANCE EMPIRICISM

Van Fraassen's reformulation of empiricism occurs within a general framework for thinking about epistemology, the core of which can be described in terms of a tripartite distinction between "levels" of epistemological analysis. At the ground level there are matters of putative fact, or claims about the nature of the world; these are potential objects of belief. Consider, for example, the claim that mammals typically give birth to live young, or that positrons have charge, or that possible worlds exist, or that the only source of knowledge of the world is experience. These are claims about aspects of reality, and if one believes them one takes them to describe these aspects correctly. Factual beliefs do not generate themselves, however. Knowing subjects must acquire them, and when one reflects on how that is done, one arrives at the second level of analysis, the level of stances.

The notion of a stance is intended to be construed rather broadly, but I will use the term to refer to epistemic stances in particular. A stance is a cluster of commitments and strategies for generating factual beliefs. It makes no claim about reality, at least not directly. One might think of them partially, after Paul Teller (2005), as combinations of epistemic "policies" with respect to the methodologies one adopts in order to generate factual beliefs. For example, consider the idea that one should think of explanatory virtue as an important desideratum in determining what to believe, or that one should privilege the methods of the sciences. These are policies regarding the generation of factual beliefs, and policies are not themselves true or false. Certainly, it may be true or false that adopting a particular stance is likely to produce facts as opposed to likely falsehoods, but stances are not themselves propositional for the most part. They furnish guidelines for ways of acting. One does not believe a stance in the way that one believes a fact. Rather, one commits to a stance, or adopts it – they are possible means to realms of possible facts. Crucially, holding a stance is a function of one's *values* as opposed to one's factual beliefs, and though values may be well or ill advised, they are not true or false. (For those critical of the fact–value distinction, it may be possible to speak here simply in terms of different sorts of beliefs.) On van Fraassen's view, as we shall see, metaphysics and empiricism are stances.

The third and final level of epistemological analysis is what I will call the level of meta-stances. Here one finds various attitudes towards the nature

of frontline, epistemic stances, and thus ultimately towards the putative facts they generate. One issue at the level of meta-stances is particularly important to the present discussion: the question of which of innumerable possible stances one should adopt. Van Fraassen advocates a view according to which it is rationally permissible to hold any stance and believe any set of facts that meet certain minimal constraints; for example, but not exclusively, those that harbour no logical inconsistency or probabilistic incoherence. This account of rationality, which he calls 'voluntarism', is opposed to the idea that any one stance (and associated set of beliefs) is rationally compelled. I will return to the matter of voluntarism shortly, but first let us come to some understanding of what stance metaphysics and stance empiricism are, precisely.

Earlier I described metaphysical approaches with which empiricists are unhappy as those that endorse speculations about unobservables as a route to belief concerning the unobservable realm. Van Fraassen identifies this with a tradition of analytic metaphysics stretching from seventeenth-century philosophers such as Descartes and Leibniz to contemporary ones such as David Armstrong and David Lewis: 'characterized by the attempted construction of a theory of the world, of the same form as a fundamental science and continuous with (as extension or foundation of) the natural sciences' (2002, p. 231, footnote 1). Henceforth I will simply use the term 'metaphysics' for this sort of speculative approach and 'empiricism' for views that oppose it. The claims of metaphysics annoy the empiricist, but this annoyance is most economically understood at the level of stances. Rather than list the countless factual claims of which empiricists disapprove, one can simply observe that metaphysics is a stance of which empiricists disapprove, which generates annoying factual claims. On van Fraassen's account, stances are generally rich fabrics of interwoven commitments and attitudes, but let me summarize the basic elements of metaphysics very concisely. The core of the metaphysical stance comprises the following epistemic policies:

- M1 Accept demands for explanation in terms of things underlying the observable.
- M2 Attempt to answer such demands by speculating about the unobservable.

Why should anyone disapprove of these policies? Empiricists hold that via M1, metaphysicians seek to explain things one already understands! Via M2, metaphysicians generate explanantia that are less comprehensible than the explananda with which they begin! These are, it turns out,

familiar responses of empiricist philosophers to metaphysics throughout the ages. The empiricist wonders, for example, why she should accept the demand for a deeper explanation of why and how green things form an identifiable group – as she already knows, they are green. And postulating the existence of universals such as greenness, and mysterious relations such as instantiation, is surely more obscure than the fact that some things are green. So argues the empiricist.

Empiricism, conversely, is a stance opposed to the excesses of metaphysics, shared by many historical positions. Again, let me summarize very concisely the core of this position, in terms of the following epistemic policies:

- E1 Reject demands for explanation in terms of things underlying the observable.
- E2 *A fortiori*, reject attempts to answer such demands by speculating about the unobservable.
- E3 Follow, as a model of inquiry, the methods of the sciences.

E1 and E2 are directly opposed to the metaphysical stance.[7] E3, on the other hand, is somewhat puzzling. It is not obvious that the sciences share any particular, substantive, methodological principles, or if they do that they are unique to the empirical stance. Van Fraassen does suggest, however, that one aspect of the sciences of which empiricists approve is a certain tolerance for different beliefs. Scientists routinely disagree, but conflicting beliefs are tolerated and respected as rivals worthy of consideration. One reason he is concerned to portray empiricism as a stance is that he is wary of the charge that, understood as a factual claim, such as 'the only source of knowledge of the world is experience', empiricism may defeat itself. For if empiricism is a factual thesis it will be contrary to other, perhaps metaphysical theses, and though any statement of empiricism would be inconsistent with statements of other views, the principle of tolerance in accordance with E3 demands that one respect contrary factual claims as rivals worthy of consideration. So much for the rejection of metaphysics by empiricists! By ascending to the level of

[7] These policies must be qualified if this is to be consistent with van Fraassen's earlier work. There (1980), he distinguishes between belief (taking a theory to be true) and mere acceptance (believing only its observable consequences). Presumably E1 and E2 concern taking explanations to be true, for there may be pragmatic reasons for pursuing metaphysics in some cases. Speculating about unobservables may facilitate the construction of more empirically adequate theories. Without this qualification, there is a tension between E1/E2 and E3, since the methods of the sciences generally favour M1/M2, not E1/E2.

stances, van Fraassen hopes to rid empiricism of any worry of incoherence in its critique of metaphysics.

In any case, E3 is puzzling, not least because a tolerance of contrary factual claims seems too liberal an attitude for the empiricist. Some factual claims are metaphysical, and it is the very business of an empiricist to be intolerant of these claims. Statements about the existence and nature of universals, causal necessity, and possible worlds may be mistaken, but they are putatively factual, and a position that takes such claims as rivals worthy of consideration would be a strange sort of empiricism. Nevertheless, rising to the level of stances does I think help the empiricist to avoid a form of self-defeat. Any plausible definition of empiricism in factual terms, such as 'the only source of knowledge of the world is experience', is likely to make a claim that reaches beyond that which is established in experience. Experience by itself does not rule out the *possibility* of other sources of knowledge. When she defines empiricism as a factual doctrine, the empiricist commits the same sin as the metaphysician: she speculates about the world in such a way as to reach beyond the observable. But this is to engage in metaphysics, and that is why van Fraassen's empiricism cannot be understood as a factual thesis, on pain of defeating itself. One can hardly oppose metaphysics by embracing a metaphysical thesis. The empirical *stance*, conversely, is not part of the metaphysical stance, and to adopt the empirical stance is not to do metaphysics in disguise. Recasting empiricism at the level of stances is thus a means of formulating the position in a way that is not obviously self-defeating.

We are now in a position to ask the question whose answer will determine the very legitimacy of an investigation into the nature of realism. Why should anyone adopt the empirical stance as opposed to its metaphysical counterpart? The reasons had better not make recourse to arguments employing metaphysical premises, or the empiricist will again find herself opposing metaphysics by doing metaphysics. And thus we find ourselves with two stances, the empirical and the metaphysical, and wanting an argument for why the former is preferable to the latter. What, then, is the case against metaphysics?

1.5 THE FALL OF THE CRITIQUE OF METAPHYSICS

I submit that there can be no case against metaphysics, or more correctly, no case for a fair-minded, non-dogmatic metaphysician to address. To understand why this is so, one must engage a specific concern at the level of meta-stances: identifying an appropriate criterion or criteria with

which to facilitate choosing a stance. Van Fraassen suggests two criteria: one that is uniformly applicable to anyone's choice of stance, and another whose application varies across stance holders. The uniform criterion is rationality. One should adopt a stance that is rational and reject those that are not. The variable criterion is the set of values that leads an agent to adopt one stance over another.

I will return to the issue of values momentarily, but first let us consider van Fraassen's conception of rationality, which is famously thin. It is rationally permissible, he says, to hold any stance or believe any set of facts that is logically consistent and probabilistically coherent. Incoherence was originally explicated (1989) in terms of holding combinations of beliefs that are exploitable by Dutch books to the detriment of the belief-holder (making bets all of whose possible outcomes are unfavourable), and consistency and coherence are usually understood as logical constraints, straightforwardly applicable to propositional things like factual beliefs. Stances, however, are in large part non-propositional, so in this context mere *logical* consistency and coherence will not suffice. At least part of what is intended by incoherence here must have a pragmatic dimension, and indeed, van Fraassen (2005, p. 184) holds that the 'defining hallmark' of irrationality more generally is 'self-sabotage by one's own lights'. Self-sabotage is broad enough to include such unfortunate circumstances as believing contradictions and probabilistically incoherent combinations, as one might do on the level of facts, but it may also include circumstances in which the stance one adopts has pragmatic failings, such as a combination of attitudes or policies that tend to undermine or conflict with one another. Note that on this view, different and mutually incompatible stances may be rational – no one stance and resultant set of beliefs are compelled. Van Fraassen calls this meta-stance 'voluntarism'.

Let us now return to values. Recall that in addition to rationality, agents' values furnish criteria for their choice of stance. If one's values promote a commitment to the empirical stance, one will reject metaphysics. After all, E1 and E2 are directly opposed to M1 and M2. The empiricist rejects metaphysics by committing to epistemic policies that are incompatible with it. But does this offer a case against metaphysics? To the consternation of the empiricist, it does not. For if rationality is the only constraint that applies uniformly to all agents adopting stances, and different, mutually incompatible stances are rational, then the framework for debate on the level of stances is relativistic. Relativism is premised on the idea that there is no view from nowhere, no view that cuts across perspectives so as to serve as a sufficient common ground from which to debate. If it turns out

that metaphysics is rational, empiricists may nevertheless claim that it is wrong-headed from their perspective. The qualifying phrase 'from one's perspective', however, is inseparable from any statement of the correctness of adopting a stance. Saying that different communities have different values is shorthand for saying that correctness and incorrectness are relativized to perspectives, and have no meaning otherwise.

Comparing M1 and M2 to E1 and E2, one finds different policies supported by different intuitions, or values, concerning two things: what needs explaining; and what counts as obscure or unilluminating. Many criticisms of stances that meet the constraint of rationality are cogent only from within the confines of some other stance, and this cogency is not preserved "outside". Thus, if empiricists hope to offer a case against metaphysics that is telling for the metaphysician, not merely for someone who adopts empiricist values that metaphysicians need not share, they must demonstrate the *irrationality* of metaphysics, because rationality is the only stance-transcendent criterion for choosing a stance. In other words, the empiricist must show that metaphysics sabotages itself, or more specifically, that if one adopts the epistemic policies of metaphysics, there are derivable consequences of which even metaphysicians would disapprove.

The task, then, is to demonstrate that metaphysics fails by its own lights, but how? Perhaps one could argue that the factual claims of metaphysics are problematic. Van Fraassen (1989) himself argues, for example, that the concept of a law of nature is incoherent. But even if it turned out that *every* current metaphysical concept was incoherent, this would not amount to a demonstration of the irrationality of metaphysics. One interesting consequence of understanding metaphysics and empiricism as stances is that they are not (exclusively) identifiable with any one set of factual beliefs. Stances underdetermine the factual beliefs they produce. Over philosophical time, both metaphysics and empiricism have survived many changes in the beliefs with which they are associated and no doubt will again. For this reason, van Fraassen (2002, p. 62) is clear that stances are not identical to the factual claims with which they may be associated at any given time. Thus, no demonstration of the irrationality of believing such factual claims can entail the irrationality of adopting a stance.

Let us focus, then, on the stance itself. Perhaps there are commitments, standards, or principles accepted by metaphysicians that the metaphysical stance itself fails to meet or exemplify. If so, this would constitute the sort of pragmatic incoherence the empiricist requires in order to demonstrate that metaphysics is irrational. There are suggestions to this effect

throughout van Fraassen's critique. Let me summarize the relevant principles as follows:

- P1 No form of inquiry into the nature of the world should be immune to the possibility of error, or failure.
- P2 Correct logical or grammatical form should not be considered sufficient to render claims about the world substantive.
- P3 The epistemic status of one's criteria for theory choice should be linked to the epistemic status of one's theories.

It seems reasonable that both metaphysicians and empiricists should accept P1–P3, so let us examine each in turn, and consider why one might think metaphysics fails to satisfy them.

First, consider P1. Van Fraassen and empiricists generally are sometimes heard to complain that metaphysics has the character of a particularly futile game. Its futility is evidenced by the fact that no one ever wins or loses, and perhaps most damagingly, it never ends! If some part of metaphysics is shown to be inconsistent, it simply reinvents itself. One always has the option, it seems, of retreating to another position within the game of metaphysics that is immune to the criticism applied, and this violates P1, the idea that no form of inquiry should have this kind of immunity. It would not be difficult, I suspect, to find some measure of sympathy for this complaint among those who are interested in the sciences. A great deal of speculation in metaphysics is too far removed from the sciences to generate much interest or care on the part of realists, for instance, at least in the context of thinking about scientific knowledge. This, however, merely expresses a taste, and expressions of taste are not demonstrations of irrationality. Metaphysicians should accept P1, since metaphysics is fallible, but one must take care not to conflate metaphysical claims and theories with the metaphysical stance itself, any more than one would conflate the empirical stance with any particular empirical claim or theory. When metaphysical claims are found to be problematic, one tries something else. Clearly, then, particular theories *can* lose out, and it is not a pointless game after all. It is in the nature of the stances that generate these candidates for knowledge, however, to go on. Thus it seems that P1 is no threat to the metaphysical stance.

Consider P2. Van Fraassen challenges metaphysicians to show that their claims are substantive. They should amount to more, he says, than 'coherent nonsense'. Merely correct logical or grammatical form is insufficient to demonstrate that metaphysical claims exemplify reasonable

attempts to say something substantive about the world. Again, I suspect that metaphysicians would agree with P2, but it is an odd sort of thing to be asked to prove the substance of one's claims, especially in the context of one's own inquiry. In response to the question of how anyone could think that M1 and M2 lead to substantive contentions, one might legitimately wonder what sort of answer would suffice. There is an interesting question here of the burden of proof. In just the same way that the empiricist wants to know what reason anyone might have for thinking that metaphysical claims are substantive, the sceptic might well ask constructive empiricists to show, for example, that their claims about the world are, in fact, something more than coherent nonsense, and so on and so forth. Perhaps only the solipsist of the present moment is safe from this line of questioning. At the end of the day, the only thing anyone can do in response to this sort of question is to point to his or her own epistemic practices, and the values that favour them, and this takes us to P3.

Metaphysical theories, says van Fraassen, are evaluated in terms of purely subjective values and probabilities of success. These values, however, such as preferences for theories that maximize simplicity, scope, or explanatory power, are not *epistemic* values. That is, they are not linked to truth, or at least one has no reason to think they are. Metaphysicians thus suffer from a form of 'false consciousness': they apply their subjective values and probabilities of success in pursuit of truths, but there is no reason to think that such application leads to anything other than theories they like. I submit, however, that van Fraassen is not in a position to make this charge, given his voluntarism. Once again, it seems reasonable that metaphysicians should accept P3, but they disagree with the empiricist's evaluation of the epistemic status of their criteria for theory choice. Consider the case of scientific theories and their epistemic status. Under certain conditions, realists think it is reasonable to infer the approximate truth of our best theories involving unobservables, and their criteria for theory choice include such things as maximizing simplicity, scope, explanatory power, etc. Empiricists demur. These criteria are at best indicative of truths about observables, they say. But does this disagreement entail that at least one of these parties is being irrational? It is hard to see how it could – neither position is rationally compelled, and neither, it seems, is guilty of inconsistency or incoherence.

Both metaphysicians and empiricists make a leap from what is strictly entailed by the observable data, as a matter of faith, perhaps, but in different ways, consistent with the values to which they subscribe. It should thus be clear that P3 is no threat to metaphysics. It is precisely

because metaphysicians think their criteria for theory choice are epistemically significant, as a result either of a voluntaristic choice or of reasons to be adduced, that they believe our theories might well be close to the truth. There is no pragmatic incoherence in this. There is, no doubt, a difference in degree between the speculation about unobservables that is most commonly part of realism, and much of what takes place in metaphysics more generally. In both cases, speculations about unobservable entities and processes are intended to be consistent with the observable data, but often scientific theories seem to take a greater risk, because they often make *novel* predictions and other metaphysical theories do not. Not all sciences make novel predictions, however, and differences in risk are differences in degree, not kind. There is no rationally compelled answer to the question of how much is required in order to make a form of inquiry acceptable. Where one draws the line here will depend on the values one has, not on matters of rationality.

In concluding this chapter I believe we are now in a position to appreciate why the realist cannot be arrested by the empiricist critique of metaphysics. The critique is subject to a form of relativism that renders it effective only to the ears of empiricists. It appeals to values and policies that empiricists share, but that need not be shared by other rational agents. Only if it could be demonstrated that the metaphysical stance is incoherent by its own lights would the empiricist have a critique that escapes this conundrum, but this is asking too much. At one point, van Fraassen (2000, p. 277) characterizes what it is to be rational in terms that I think, despite his deep commitment to empiricism, embrace the metaphysical stance:

> Nothing more than staying within the bounds of reason is needed for [the] status of rationality. Not good reasons, not a rationale, not support of any special sort . . . nothing is needed above and beyond coherence. Thus any truly coherent position is rational.

On his conception of epistemology the threshold for rationality is low, and as a consequence the threshold of irrationality is very high indeed. When the sceptic challenges the constructive empiricist to prove that it is not irrational to believe the observable content of our best theories, I do not think the latter has much to answer for. Empiricists choose forms of inquiry that fit with their values, epistemic and otherwise, and some of these tell them the sceptic's life is not worth living. The same applies to the metaphysician. One may decide, in accordance with one's values, what forms of inquiry to pursue. That is our prerogative, after all. But few

if any prerogatives transcend all possible stances, and there can be no radical critique of metaphysics by empiricism.

The metaphysics of the sciences concerns the observable and unobservable parts of the world described by scientific theories, both explicitly and implicitly. The epistemology of the sciences concerns the specific methods used to generate scientific claims, the justification or confirmation of these claims, whether they constitute knowledge, and if so, what sort. The influence of logical positivism during the birth of the philosophy of science as a separate discipline in the late nineteenth century, and throughout most of the twentieth century, led to a vestigial neglect of metaphysical questions in connection with realism. Those investigating problems such as the nature of causation, laws of nature, and conceptions of natural kinds have done so largely in isolation from debates between realists and antirealists. Metaphysical issues have been the purview of the philosophy of particular sciences: space and time, evolutionary biology, quantum mechanics, and so on. The neglect of metaphysics in the context of realism, however, is a mistake. For there is a sense in which the metaphysics of science is a precursor to its epistemology. One cannot fully appreciate what it might mean to be a realist until one has a clear picture of what one is being invited to be a realist about.

In the further chapters of this book, I will propose an answer to the question of how to construe realism by developing a metaphysics that underpins it. The aim of this endeavour is an integrated account of the unobservables of which scientists speak in detail, like mitochondria and neutrinos, and those features of reality that realists sometimes take for granted but say little about, like causation and laws. The result, I hope, is a reunion of arguments about the natures of things in the world with those about how one can know these things – a reunion that redresses the separation of metaphysics and epistemology in the context of scientific knowledge.

CHAPTER 2

Selective scepticism: entity realism, structural realism, semirealism

2.1 THE ENTITIES ARE NOT ALONE

Scientific realists invite questions about their metaphysical beliefs, often perhaps unwittingly. In their accounts of scientific knowledge, they routinely invoke not only unobservable entities and processes commonly discussed by scientists, but also things whose natures generally fall outside the remit of the sciences, such as causation, laws of nature, and the idea that scientific taxonomies divide the world into natural categories, or kinds. The recourse to these latter metaphysical notions in support of realism is not problematic *per se*, but a lack of attention given to spelling them out can have problematic consequences. While these topics are central to metaphysics and many realists investigate them, few offer unified accounts in connection with specific proposals for realism. In the absence of such details, the views of realists are sometimes associated by default with the metaphysical speculations of great, systematic philosophers of the past, from ancient and medieval, up to and including early modern times. Unfortunately for the realist, some of these speculations are outmoded today, especially in a modern scientific context. In this chapter I will begin the process of spelling out in a more detailed way what I believe scientific realism has become, and thereby initiate a proposal for its metaphysical foundations.

Chapter 1 began with a rough, first-approximation definition of realism: scientific theories correctly describe the nature of a mind-independent world. This first approximation, however, is naïve in several respects, and this leaves it open to several immediate objections. In order to remedy this situation a number of qualifications are usually made. These include the idea that realists should only commit to theories that are genuinely successful and not merely *ad hoc*, as evidenced by the nature of their predictions, retrodictions, and novel predictions. Another restriction is to theories that are sufficiently mature. A mature theory is one that has