

# **REPRESENTING AND INTERVENING**

---

**INTRODUCTORY TOPICS IN THE PHILOSOPHY OF  
NATURAL SCIENCE**

**IAN HACKING**



**CAMBRIDGE**  
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS  
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,  
São Paulo, Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press  
32 Avenue of the Americas, New York, NY 10013-2473, USA  
[www.cambridge.org](http://www.cambridge.org)  
Information on this title: [www.cambridge.org/9780521282468](http://www.cambridge.org/9780521282468)

© Cambridge University Press 1983

This publication is in copyright. Subject to statutory exception  
and to the provisions of relevant collective licensing agreements,  
no reproduction of any part may take place without the written  
permission of Cambridge University Press.

First published 1983  
25th printing 2010

*A catalog record for this publication is available from the British Library.*

ISBN 978-0-521-23829-8 Hardback  
ISBN 978-0-521-28246-8 Paperback

Cambridge University Press has no responsibility for the persistence or  
accuracy of URLs for external or third-party Internet Web sites referred to in  
this publication and does not guarantee that any content on such Web sites is,  
or will remain, accurate or appropriate.

**PART A**  
**REPRESENTING**



# 1 What is scientific realism?

*Scientific realism* says that the entities, states and processes described by correct theories really do exist. Protons, photons, fields of force, and black holes are as real as toe-nails, turbines, eddies in a stream, and volcanoes. The weak interactions of small particle physics are as real as falling in love. Theories about the structure of molecules that carry genetic codes are either true or false, and a genuinely correct theory would be a true one.

Even when our sciences have not yet got things right, the realist holds that we often get close to the truth. We aim at discovering the inner constitution of things and at knowing what inhabits the most distant reaches of the universe. Nor need we be too modest. We have already found out a good deal.

*Anti-realism* says the opposite: there are no such things as electrons. Certainly there are phenomena of electricity and of inheritance but we construct theories about tiny states, processes and entities only in order to predict and produce events that interest us. The electrons are fictions. Theories about them are tools for thinking. Theories are adequate or useful or warranted or applicable, but no matter how much we admire the speculative and technological triumphs of natural science, we should not regard even its most telling theories as true. Some anti-realists hold back because they believe theories are intellectual tools which cannot be understood as literal statements of how the world is. Others say that theories must be taken literally – there is no other way to understand them. But, such anti-realists contend, however much we may use the theories we do not have compelling reasons to believe they are right. Likewise anti-realists of either stripe will not include theoretical entities among the kinds of things that really exist in the world: turbines yes, but photons no.

We have indeed mastered many events in nature, says the anti-realist. Genetic engineering is becoming as commonplace as manufacturing steel, but do not be deluded. Do not suppose that

long chains of molecules are really there to be spliced. Biologists may think more clearly about an amino acid if they build a molecular model out of wire and coloured balls. The model may help us arrange the phenomena in our minds. It may suggest new microtechnology, but it is not a literal picture of how things really are. I could make a model of the economy out of pulleys and levers and ball bearings and weights. Every decrease in weight  $M$  (the ‘money supply’) produces a decrease in angle  $I$  (the ‘rate of inflation’) and an increase in the number  $N$  of ball bearings in this pan (the number of unemployed workers). We get the right inputs and outputs, but no one suggests that this is what the economy *is*.

### If you can spray them, then they are real

For my part I never thought twice about scientific realism until a friend told me about an ongoing experiment to detect the existence of fractional electric charges. These are called quarks. Now it is not the quarks that made me a realist, but rather electrons. Allow me to tell the story. It ought not to be a simple story, but a realistic one, one that connects with day to day scientific research. Let us start with an old experiment on electrons.

The fundamental unit of electric charge was long thought to be the electron. In 1908 J.A. Millikan devised a beautiful experiment to measure this quantity. A tiny negatively charged oil droplet is suspended between electrically charged plates. First it is allowed to fall with the electric field switched off. Then the field is applied to hasten the rate of fall. The two observed terminal velocities of the droplet are combined with the coefficient of viscosity of the air and the densities of air and oil. These, together with the known value of gravity, and of the electric field, enable one to compute the charge on the drop. In repeated experiments the charges on these drops are small integral multiples of a definite quantity. This is taken to be the minimum charge, that is, the charge on the electrons. Like all experiments, this one makes assumptions that are only roughly correct: that the drops are spherical, for instance. Millikan at first ignored the fact that the drops are not large compared to the mean free path of air molecules so they get bumped about a bit. But the idea of the experiment is definitive.

The electron was long held to be the unit of charge. We use  $e$  as the name of that charge. Small particle physics, however, increas-

ingly suggests an entity, called a quark, that has a charge of  $1/3 e$ . Nothing in theory suggests that quarks have independent existence; if they do come into being, theory implies, then they react immediately and are gobbled up at once. This has not deterred an ingenious experiment started by LaRue, Fairbank and Hebard at Stanford. They are hunting for 'free' quarks using Millikan's basic idea.

Since quarks may be rare or short-lived, it helps to have a big ball rather than a tiny drop, for then there is a better chance of having a quark stuck to it. The drop used, although weighing less than  $10^{-4}$  grams, is  $10^7$  times bigger than Millikan's drops. If it were made of oil it would fall like a stone, almost. Instead it is made of a substance called niobium, which is cooled below its superconducting transition temperature of  $9^\circ\text{K}$ . Once an electric charge is set going round this very cold ball, it stays going, forever. Hence the drop can be kept afloat in a magnetic field, and indeed driven back and forth by varying the field. One can also use a magnetometer to tell exactly where the drop is and how fast it is moving.

The initial charge placed on the ball is gradually changed, and, applying our present technology in a Millikan-like way, one determines whether the passage from positive to negative charge occurs at zero or at  $\pm 1/3 e$ . If the latter, there must surely be one loose quark on the ball. In their most recent preprint, Fairbank and his associates report four fractional charges consistent with  $+1/3 e$ , four with  $-1/3 e$ , and 13 with zero.

Now how does one alter the charge on the niobium ball? 'Well, at that stage,' said my friend, 'we spray it with positrons to increase the charge or with electrons to decrease the charge.' From that day forth I've been a scientific realist. *So far as I'm concerned, if you can spray them then they are real.*

Long-lived fractional charges are a matter of controversy. It is not quarks that convince me of realism. Nor, perhaps, would I have been convinced about electrons in 1908. There were ever so many more things for the sceptic to find out: There was that nagging worry about inter-molecular forces acting on the oil drops. Could that be what Millikan was actually measuring? So that his numbers showed nothing at all about so-called electrons? If so, Millikan goes no way towards showing the reality of electrons. Might there be minimum electric charges, but no electrons? In our quark example

we have the same sorts of worry. Marinelli and Morpurgo, in a recent preprint, suggest that Fairbank's people are measuring a new electromagnetic force, not quarks. What convinced me of realism has nothing to do with quarks. It was the fact that by now there are standard emitters with which we can spray positrons and electrons – and that is precisely what we do with them. We understand the effects, we understand the causes, and we use these to find out something else. The same of course goes for all sorts of other tools of the trade, the devices for getting the circuit on the supercooled niobium ball and other almost endless manipulations of the ‘theoretical’.

### **What is the argument about?**

The practical person says: consider what you use to do what you do. If you spray electrons then they are real. That is a healthy reaction but unfortunately the issues cannot be so glibly dismissed. Anti-realism may sound daft to the experimentalist, but questions about realism recur again and again in the history of knowledge. In addition to serious verbal difficulties over the meanings of ‘true’ and ‘real’, there are substantive questions. Some arise from an intertwining of realism and other philosophies. For example, realism has, historically, been mixed up with materialism, which, in one version, says everything that exists is built up out of tiny material building blocks. Such a materialism will be realistic about atoms, but may then be anti-realistic about ‘immaterial’ fields of force. The dialectical materialism of some orthodox Marxists gave many modern theoretical entities a very hard time. Lysenko rejected Mendelian genetics partly because he doubted the reality of postulated ‘genes’.

Realism also runs counter to some philosophies about causation. Theoretical entities are often supposed to have causal powers: electrons neutralize positive charges on niobium balls. The original nineteenth-century positivists wanted to do science without ever speaking of ‘causes’, so they tended to reject theoretical entities too. This kind of anti-realism is in full spate today.

Anti-realism also feeds on ideas about knowledge. Sometimes it arises from the doctrine that we can know for real only the subjects of sensory experience. Even fundamental problems of logic get

involved; there is an anti-realism that puts in question what it is for theories to be true or false.

Questions from the special sciences have also fuelled controversy. Old-fashioned astronomers did not want to adopt a realist attitude to Copernicus. The idea of a solar system might help calculation, but it does not say how the world really is, for the earth, not the sun, they insisted, is the centre of the universe. Again, should we be realists about quantum mechanics? Should we realistically say that particles do have a definite although unknowable position and momentum? Or at the opposite extreme should we say that the ‘collapse of the wave packet’ that occurs during microphysical measurement is an interaction with the human mind?

Nor shall we find realist problems only in the specialist natural sciences. The human sciences give even more scope for debate. There can be problems about the libido, the super ego, and the transference of which Freud teaches. Might one use psychoanalysis to understand oneself or another, yet cynically think that nothing answers to the network of terms that occurs in the theory? What should we say of Durkheim’s supposition that there are real, though by no means distinctly discernible, social processes that act upon us as inexorably as the laws of gravity, and yet which exist in their own right, over and above the properties of the individuals that constitute society? Could one coherently be a realist about sociology and an anti-realist about physics, or vice versa?

Then there are meta-issues. Perhaps realism is as pretty an example as we could wish for, of the futile triviality of basic philosophical reflections. The questions, which first came to mind in antiquity, are serious enough. There was nothing wrong with asking, once, Are atoms real? But to go on discussing such a question may be only a feeble surrogate for serious thought about the physical world.

That worry is anti-philosophical cynicism. There is also philosophical anti-philosophy. It suggests that the whole family of issues about realism and anti-realism is mickey-mouse, founded upon a prototype that has dogged our civilization, a picture of knowledge ‘representing’ reality. When the idea of correspondence between thought and the world is cast into its rightful place – namely, the grave – will not, it is asked, realism and anti-realism quickly follow?

### **Movements, not doctrines**

Definitions of ‘scientific realism’ merely point the way. It is more an attitude than a clearly stated doctrine. It is a way to think about the content of natural science. Art and literature furnish good comparisons, for not only has the word ‘realism’ picked up a lot of philosophical connotations: it also denotes several artistic movements. During the nineteenth century many painters tried to escape the conventions that bound them to portray ideal, romantic, historical or religious topics on vast and energetic canvases. They chose to paint scenes from everyday life. They refused to ‘aestheticize’ a scene. They accepted material that was trivial or banal. They refused to idealize it, refused to elevate it: they would not even make their pictures picturesque. Novelists adopted this realist stance, and in consequence we have the great tradition in French literature that passes through Flaubert and which issues in Zola’s harrowing descriptions of industrial Europe. To quote an unsympathetic definition of long ago, ‘a realist is one who deliberately declines to select his subjects from the beautiful or harmonious, and, more especially, describes ugly things and brings out details of the unsavoury sort’.

Such movements do not lack doctrines. Many issued manifestos. All were imbued with and contributed to the philosophical sensibilities of the day. In literature some latterday realism was called positivism. But we speak of movements rather than doctrine, of creative work sharing a family of motivations, and in part defining itself in opposition to other ways of thinking. Scientific realism and anti-realism are like that: they too are movements. We can enter their discussions armed with a pair of one-paragraph definitions, but once inside we shall encounter any number of competing and divergent opinions that comprise the philosophy of science in its present excited state.

### **Truth and real existence**

With misleading brevity I shall use the term ‘theoretical entity’ as a portmanteau word for all that ragbag of stuff postulated by theories but which we cannot observe. That means, among other things, particles, fields, processes, structures, states and the like. There are two kinds of scientific realism, one for theories, and one for entities.

The question about theories is whether they are true, or are true-or-false, or are candidates for truth, or aim at the truth.

The question about entities is whether they exist.

A majority of recent philosophers worries most about theories and truth. It might seem that if you believe a theory is true, then you automatically believe that the entities of the theory exist. For what is it to think that a theory about quarks is true, and yet deny that there are any quarks? Long ago Bertrand Russell showed how to do that. He was not, then, troubled by the truth of theories, but was worried about unobservable entities. He thought we should use logic to rewrite the theory so that the supposed entities turn out to be logical constructions. The term ‘quark’ would not denote quarks, but would be shorthand, via logic, for a complex expression which makes reference only to observed phenomena. Russell was then a realist about theories but an anti-realist about entities.

It is also possible to be a realist about entities but an anti-realist about theories. Many Fathers of the Church exemplify this. They believed that God exists, but they believed that it was in principle impossible to form any true positive intelligible theory about God. One could at best run off a list of what God is not – not finite, not limited, and so forth. The scientific-entities version of this says we have good reason to suppose that electrons exist, although no full-fledged description of electrons has any likelihood of being true. Our theories are constantly revised; for different purposes we use different and incompatible models of electrons which one does not think are literally true, but there are electrons, nonetheless.

## Two realisms

*Realism about entities* says that a good many theoretical entities really do exist. Anti-realism denies that, and says that they are fictions, logical constructions, or parts of an intellectual instrument for reasoning about the world. Or, less dogmatically, it may say that we have not and cannot have any reason to suppose they are not fictions. They may exist, but we need not assume that in order to understand the world.

*Realism about theories* says that scientific theories are either true or false independent of what we know: science at least aims at the truth, and the truth is how the world is. Anti-realism says that

theories are at best warranted, adequate, good to work on, acceptable but incredible, or what-not.

### Subdivisions

I have just run together claims about reality and claims about what we know. My realism about entities implies both that a satisfactory theoretical entity would be one that existed (and was not merely a handy intellectual tool). That is a claim about entities and reality. It also implies that we actually know, or have good reason to believe in, at least some such entities in present science. That is a claim about knowledge.

I run knowledge and reality together because the whole issue would be idle if we did not *now* have some entities that some of us think really do exist. If we were talking about some future scientific utopia I would withdraw from the discussion. The two strands that I run together can be readily unscrambled, as in the following scheme of W. Newton-Smith's.<sup>1</sup> He notes three ingredients in scientific realism:

1 An *ontological* ingredient: scientific theories are either true or false, and that which a given theory is, is in virtue of how the world is.

2 A *causal* ingredient: if a theory is true, the theoretical terms of the theory denote theoretical entities which are causally responsible for the observable phenomena.

3 An *epistemological* ingredient: we can have warranted belief in theories or in entities (at least in principle).

Roughly speaking, Newton-Smith's causal and epistemological ingredients add up to my realism about entities. Since there are two ingredients, there can be two kinds of anti-realism. One rejects (1); the other rejects (3).

You might deny the ontological ingredient. You deny that theories are to be taken literally; they are not either true or false; they are intellectual tools for predicting phenomena; they are rules for working out what will happen in particular cases. There are many versions of this. Often an idea of this sort is called *instrumentalism* because it says that theories are only instruments.

Instrumentalism denies (1). You might instead deny (3). An

<sup>1</sup> W. Newton-Smith, 'The underdetermination of theory by data', *Proceedings of the Aristotelian Society*, Supplementary Volume 52 (1978), p. 72.

example is Bas van Fraassen in his book *The Scientific Image* (1980). He thinks theories are to be taken literally – there is no other way to take them. They are either true or false, and which they are depends on the world – there is no alternative semantics. But we have no warrant or need to believe any theories about the unobservable in order to make sense of science. Thus he denies the epistemological ingredient.

My realism about theories is, then, roughly (1) and (3), but my realism about entities is not exactly (2) and (3). Newton-Smith's causal ingredient says that if a theory is true, then the theoretical terms denote entities that are causally responsible for what we can observe. He implies that belief in such entities depends on belief in a theory in which they are embedded. But one can believe in some entities without believing in any particular theory in which they are embedded. One can even hold that no general deep theory about the entities could possibly be true, for there is no such truth. Nancy Cartwright explains this idea in her book *How the Laws of Physics Lie* (1983). She means the title literally. The laws are deceitful. Only phenomenological laws are possibly true, but we may well know of causally effective theoretical entities all the same.

Naturally all these complicated ideas will have an airing in what follows. Van Fraassen is mentioned in numerous places, especially Chapter 3. Cartwright comes up in Chapter 2 and Chapter 12. The overall drift of this book is away from realism about theories and towards realism about those entities we can use in experimental work. That is, it is a drift away from representing, and towards intervening.

### Metaphysics and the special sciences

We should also distinguish realism-in-general from realism-in-particular.

To use an example from Nancy Cartwright, ever since Einstein's work on the photoelectric effect the photon has been an integral part of our understanding of light. Yet there are serious students of optics, such as Willis Lamb and his associates, who challenge the reality of photons, supposing that a deeper theory would show that the photon is chiefly an artifact of our present theories. Lamb is not saying that the extant theory of light is plain false. A more profound theory would preserve most of what is now believed about light, but

would show that the effects we associate with photons yield, on analysis, to a different aspect of nature. Such a scientist could well be a realist in general, but an anti-realist about photons in particular.

Such localized anti-realism is a matter for optics, not philosophy. Yet N.R. Hanson noticed a curious characteristic of new departures in the natural sciences. At first an idea is proposed chiefly as a calculating device rather than a literal representation of how the world is. Later generations come to treat the theory and its entities in an increasingly realistic way. (Lamb is a sceptic proceeding in the opposite direction.) Often the first authors are ambivalent about their entities. Thus James Clerk Maxwell, one of the creators of statistical mechanics, was at first loth to say whether a gas really is made up of little bouncy balls producing effects of temperature pressure. He began by regarding this account as a ‘mere’ model, which happily organizes more and more macroscopic phenomena. He became increasingly realist. Later generations apparently regard kinetic theory as a good sketch of how things really are. It is quite common in science for anti-realism about a particular theory or its entities to give way to realism.

Maxwell’s caution about the molecules of a gas was part of a general distrust of atomism. The community of physicists and chemists became fully persuaded of the reality of atoms only in our century. Michael Gardner has well summarized some of the strands that enter into this story.<sup>2</sup> It ends, perhaps, when Brownian motion was fully analysed in terms of molecular trajectories. This feat was important not just because it suggested in detail how molecules were bumping into pollen grains, creating the observable movement. The real achievement was a new way to determine Avogadro’s number, using Einstein’s analysis of Brownian motion and Jean Perrin’s experimental techniques.

That was of course a ‘scientific’, not a ‘philosophical’, discovery. Yet realism about atoms and molecules was once the central issue for philosophy of science. Far from being a local problem about one kind of entity, atoms and molecules were the chief candidates for real (or merely fictional) theoretical entities. Many of our present positions on scientific realism were worked out then, in connection

<sup>2</sup> M. Gardner, ‘Realism and instrumentalism in 19th century atomism’, *Philosophy of Science* 46 (1979), pp. 1–34.

with that debate. The very name ‘scientific realism’ came into use at that time.

Realism-in-general is thus to be distinguished from realism-in-particular, with the proviso that a realism-in-particular can so dominate discussion that it determines the course of realism-in-general. A question of realism-in-particular is to be settled by research and development of a particular science. In the end the sceptic about photons or black holes has to put up or shut up. Realism-in-general reverberates with old metaphysics and recent philosophy of language. It is vastly less contingent on facts of nature than any realism-in-particular. Yet the two are not fully separable and often, in formative stages of our past, have been intimately combined.

### **Representation and intervention**

Science is said to have two aims: theory and experiment. Theories try to say how the world is. Experiment and subsequent technology change the world. We represent and we intervene. We represent in order to intervene, and we intervene in the light of representations. Most of today’s debate about scientific realism is couched in terms of theory, representation, and truth. The discussions are illuminating but not decisive. This is partly because they are so infected with intractable metaphysics. I suspect there can be no final argument for or against realism at the level of representation. When we turn from representation to intervention, to spraying niobium balls with positrons, anti-realism has less of a grip. In what follows I start with a somewhat old-fashioned concern with realism about entities. This soon leads to the chief modern studies of truth and representation, of realism and anti-realism about theories. Towards the end I shall come back to intervention, experiment, and entities.

The final arbitrator in philosophy is not how we think but what we do.

## 2 Building and causing

Does the word ‘real’ have any use in natural science? Certainly. Some experimental conversations are full of it. Here are two real examples. The cell biologist points to a fibrous network that regularly is found on micrographs of cells prepared in a certain way. It looks like chromatin, namely the stuff in the cell nucleus full of fundamental proteins. It stains like chromatin. But it is not real. It is only an artifact that results from the fixation of nucleic sap by glutaraldehyde. We do get a distinctive reproduction pattern, but it has nothing to do with the cell. It is an artifact of the preparation.<sup>1</sup>

To turn from biology to physics, some critics of quark-hunting don’t believe that Fairbank and his colleagues have isolated long-lived fractional charges. The results may be important but the free quarks aren’t real. In fact one has discovered something quite different; a hitherto unknown new electromagnetic force.

What does ‘real’ mean, anyway? The best brief thoughts about the word are those of J.L. Austin, once the most powerful philosophical figure in Oxford, where he died in 1960 at the age of 49. He cared deeply about common speech, and thought we often prance off into airy-fairy philosophical theories without recollecting what we are saying. In Chapter 7 of his lectures, *Sense and Sensibilia*, he writes about reality: ‘We must not dismiss as beneath contempt such humble but familiar phrases as “not real cream”.’ That was his first methodological rule. His second was not to look for ‘one single specifiable always-the-same *meaning*’. He is warning us against looking for synonyms, while at the same time urging systematic searches for regularities in the usage of a word.

He makes four chief observations about the word ‘real’. Two of these seem to me to be important even though they are expressed somewhat puckishly. The two right remarks are that the word ‘real’

<sup>1</sup> For example, R.J. Skaer and S. Whytock, ‘Chromatin-like artifacts from nuclear sap’, *Journal of Cell Science* 26 (1977), pp. 301–5.

is substantive-hungry: hungry for nouns. The word is also what Austin, in a genially sexist way, calls a trouser-word.

The word is hungry for nouns because ‘that’s real’ demands a noun to be properly understood: real cream, a real constable, a real Constable.

‘Real’ is called a trouser-word because of negative uses of the words ‘wear the trousers’. Pink cream is pink, the same colour as a pink flamingo. But to call some stuff real cream is not to make the same sort of positive assertion. Real cream is, perhaps, not a non-dairy coffee product. Real leather is hide, not naugehyde, real diamonds are not paste, real ducks are not decoys, and so forth. The force of ‘real S’ derives from the negative ‘not (a) real S’. Being hungry for nouns and being a trouser-word are connected. To know what wears the trousers we have to know the noun, in order that we can tell what is being denied in a negative usage. Real telephones are, in a certain context, not toys, in another context, not imitations, or not purely decorative. This is not because the word is ambiguous, but because whether or not something is a real N depends upon the N in question. The word ‘real’ is regularly doing the same work, but you have to look at the N to see what work is being done. The word ‘real’ is like a migrant farm worker whose work is clear: to pick the present crop. But what is being picked? Where is it being picked? How is it being picked? That depends on the crop, be it lettuce, hops, cherries or grass.

In this view the word ‘real’ is not ambiguous between ‘real chromatin’, ‘real charge’, and ‘real cream’. One important reason for urging this grammatical point is to discourage the common idea that there *must* be different kinds of reality, just because the word is used in so many ways. Well, perhaps there are different kinds of reality. I don’t know, but let not a hasty grammar force us to conclude there are different kinds of reality. Moreover we now must force the philosopher to make plain what contrast is being made by the word ‘real’ in some specialized debate. If theoretical entities are, or are not, real entities, what contrast is being made?

## Materialism

J.J.C. Smart meets the challenge in his book, *Philosophy and Scientific Realism* (1963). Yes, says Smart, ‘real’ should mark a contrast. Not all theoretical entities are real. ‘Lines of force, unlike

electrons, *are* theoretical fictions. I wish to say that this table is composed of electrons, etc., just as this wall is composed of bricks' (p. 36). A swarm of bees is made up of bees, but nothing is made up of lines of force. There is a definite number of bees in a swarm and of electrons in a bottle, but there is no definite number of lines of magnetic force in a given volume; only a convention allows us to count them.

With the physicist Max Born in mind, Smart say that the anti-realist holds that electrons do not occur in the series: 'stars, planets, mountains, houses, tables, grains of wood, microscopic crystals, microbes'. On the contrary, says Smart, crystals *are* made up of molecules, molecules of atoms, and atoms are made up of electrons, among other things. So, infers Smart, the anti-realist is wrong. There are at least some real theoretical entities. On the other hand, the word 'real' marks a significant distinction. In Smart's account, lines of magnetic force are not real.

Michael Faraday, who first taught us about lines of force, did not agree with Smart. At first he thought that lines of force are indeed a mere intellectual tool, a geometrical device without any physical significance. In 1852, when he was over 60, Faraday changed his mind. 'I cannot conceive curved lines of force without the condition of physical existence in that intermediate space.'<sup>2</sup> He had come to realize that it is possible to exert a stress on the lines of force, so they had, in his mind, to have real existence. 'There can be no doubt,' writes his biographer, 'that Faraday was firmly convinced that lines of force were real.' This does not show that Smart is mistaken. It does however remind us that some physical conceptions of reality pass beyond the rather simplistic level of building blocks.

Smart is a *materialist* – he himself now prefers the term *physicalist*. I do not mean that he insists that electrons are brute matter. By now the older ideas of matter have been replaced by more subtle notions. His thought remains, however, based on the idea that material things like stars and tables are built up out of electrons and so forth. The anti-materialist, Berkeley, objecting to the corpuscles of Robert Boyle and Isaac Newton, was rejecting just such a picture. Indeed Smart sees himself as opposed to phenomenism, a modern version of Berkeley's immaterialism. It is perhaps

<sup>2</sup> All quotations from and remarks about Faraday are from L. Pearce Williams, *Michael Faraday, A biography*, London and New York, 1965.

significant that Faraday was no materialist. He is part of that tradition in physics that downplays matter and emphasizes fields of force and energy. One may even wonder if Smart's materialism is an empirical thesis. Suppose that the model of the physical world, due to Leibniz, to Boscovič, to the young Kant, to Faraday, to nineteenth-century energeticists, is in fact far more successful than atomism. Suppose that the story of building blocks gives out after a while. Would Smart then conclude that the fundamental entities of physics are theoretical fictions?

*La Réalité Physique*, the most recent book by the philosophical quantum theorist, Bernard d'Espagnat, is an argument that we can continue to be scientific realists without being materialists. Hence 'real' must be able to mark other contrasts than the one chosen by Smart. Note also that Smart's distinction does not help us say whether the theoretical entities of social or psychological science are real. Of course one can to some extent proceed in a materialistic way. Thus we find the linguist Noam Chomsky, in his book *Rules and Representations* (1980), urging realism in cognitive psychology. One part of his claim is that structured material found in the brain, and passed down from generation to generation, helps explain language acquisition. But Chomsky is not asserting only that the brain is made up of organized matter. He thinks the structures are responsible for some of the phenomenon of thought. Flesh and blood structures in our heads cause us to think in certain ways. This word 'cause' prompts another version of scientific realism.

### Causalism

Smart is a materialist. By analogy say that someone who emphasizes the causal powers of real stuff is a *causalist*. David Hume may have wanted to analyse causality in terms of regular association between cause and effect. But good Humeans know there must be more than mere correlation. Every day we read this sort of thing:

While the American College of Obstetricians and Gynecologists recognizes that an association has been established between toxic-shock-syndrome and menstruation-tampon use, we should not assume that this means there is a definite cause-and-effect relationship until we better understand the mechanism that creates this condition. (Press release, October 7, 1980.)

A few young women employing a new brand ('Everything you've ever wanted in a tampon . . . or napkin') vomit, have diarrhoea and

a high fever, some skin rash, and die. It is not just fear of libel suits that makes the College want a better understanding of mechanisms before it speaks of causes. Sometimes an interested party does deny that an association shows anything. For example, on September 19, 1980, a missile containing a nuclear warhead blew up after someone had dropped a pipe wrench down the silo. The warhead did not go off, but soon after the chemical explosion the nearby village of Guy, Arkansas, was covered in reddish-brown fog. Within an hour of the explosion the citizens of Guy had burning lips, shortness of breath, chest pains, and nausea. The symptoms continued for weeks and no one anywhere else in the world had the same problem. Cause and Effect? ‘The United States Air Force has contended that no such correlation has been determined.’ (Press release, October 11, 1980).

The College of Obstetricians and Gynecologists insists that we cannot talk of causes until we find out how the causes of toxic-shock syndrome actually work. The Air Force, in contrast, is lying through its teeth. It is important to the causalist that such distinctions arise in a natural way. We distinguish ludicrous denials of any correlation, from assertions of correlations. We also distinguish correlations from causes. The philosopher C.D. Broad once made this anti-Humeian point in the following way. We may observe that every day a factory hooter in Manchester blows at noon, and exactly at noon the workers in a factory in Leeds lay down their tools for an hour. There is a perfect regularity, but the hooter in Manchester is not the cause of the lunch break in Leeds.

Nancy Cartwright advocates causalism. In her opinion one makes a very strong claim in calling something a cause. We must understand why a certain type of event regularly produces an effect. Perhaps the clearest proof of such understanding is that we can actually use events of one kind to produce events of another kind. Positrons and electrons are thus to be called real, in her vocabulary, since we can for example spray them, separately, on the niobium droplet and thereby change its charge. It is well understood why this effect follows the spraying. One made the experimental device because one knew it would produce these effects. A vast number of very different causal chains are understood and employed. We are entitled to speak of the reality of electrons not because they are building blocks but because we know that they have quite specific causal powers.

This version of realism makes sense of Faraday. As his biographer put it:

The magnetic lines of force are visible if and when iron filings are spread around a magnet, and the lines are supposedly denser where the filings are thicker. But no one had assumed that the lines of force are there, in reality, even when the iron filings are removed. Faraday now did: we can cut these lines and get a real effect (for example with the electric motor that Faraday invented) – hence they are real.

The true story of Faraday is a little more complicated. Only long after he had invented the motor did he set out his line of force realism in print. He began by saying ‘I am now about to leave the strict line of reasoning for a time, and enter upon a few speculations respecting the physical character of lines of force’. But whatever the precise structure of Faraday’s thought, we have a manifest distinction between a tool for calculation and a conception of cause and effect. No materialist who follows Smart will regard lines of force as real. Faraday, tinged with immaterialism, and something of a causalist, made just that step. It was a fundamental move in the history of science. Next came Maxwell’s electrodynamics that still envelops us.

### **Entities not theories**

I distinguished *realism about entities* and *realism about theories*. Both causalists and materialists care more for entities than theories. Neither has to imagine that there is a best true theory about electrons. Cartwright goes further; she denies that the laws of physics state the facts. She denies that the models that play such a central role in applied physics are literal representations of how things are. She is an anti-realist about theories and a realist about entities. Smart could, if he chose, take a similar stance. We may have no true theory about how electrons go into the build-up of atoms, then of molecules, then of cells. We will have models and theory sketches. Cartwright emphasizes that in several branches of quantum mechanics the investigator regularly uses a whole battery of models of the same phenomena. No one thinks that one of these is the whole truth, and they may be mutually inconsistent. They are intellectual tools that help us understand phenomena and build bits and pieces of experimental technology. They enable us to intervene in processes and to create new and hitherto unimagined phen-

omena. But what is actually ‘making things happen’ is not the set of laws, or true laws. There are no exactly true laws to make anything happen. It is the electron and its ilk that is producing the effects. The electrons are real, they produce the effects.

This is a striking reversal of the empiricist tradition going back to Hume. In that doctrine it is only the regularities that are real. Cartwright is saying that in nature there are no deep and completely uniform regularities. The regularities are features of the ways in which we construct theories in order to think about things. Such a radical doctrine can only be assessed in the light of her detailed treatment in *How the Laws of Physics Lie*. One aspect of her approach is described in Chapter 12 below.

The possibility of such a reversal owes a good deal to Hilary Putnam. As we shall find in Chapters 6 and 7, he had readily modified his views. What is important here is that he rejects the plausible notion that theoretical terms, such as ‘electron’, get their sense from within a particular theory. He suggests instead that we can name kinds of things that the phenomena suggest to an inquiring and inventive mind. Sometimes we shall be naming nothing, but often one succeeds in formulating the idea of a kind of thing that is retained in successive elaborations of theory. More importantly one begins to be able to do things with the theoretical entity. Early in the day one may start to measure it; much later, one may spray with it. We shall have all sorts of incompatible accounts of it, all of which agree in describing various causal powers which we are actually able to employ while intervening in nature. (Putnam’s ideas are often run together with ideas about essence and necessity more attributable to Saul Kripke: I attend only to the practical and pragmatic part of Putnam’s account of naming.)

### Beyond physics

Unlike the materialist, the causalist can consider whether the superego or late capitalism is real. Each case has to stand on its own: one might conclude that Jung’s collective unconscious is not real while Durkheim’s collective consciousness is real. Do we sufficiently understand what these objects or processes do? Can we intervene and redeploy them? Measurement is not enough. We can measure IQ and boast that a dozen different techniques give the same stable array of numbers, but we have not the slightest causal

understanding. In a recent polemic Stephen Jay Gould speaks of the 'fallacy of reification' in the history of IQ: I agree.

Causalism is not unknown in the social sciences. Take Max Weber (1864–1920), one of the founding fathers. He has a famous doctrine of ideal types. He was using the word 'ideal' fully aware of its philosophical history. In his usage it contrasts with 'real'. The ideal is a conception of the human mind, an instrument of thought (and none the worse for that). Just like Cartwright in our own day, he was 'quite opposed to the naturalistic prejudice that the goal of the social sciences must be the reduction of reality to "laws"'. In a cautious observation about Marx, Weber writes,

All specifically Marxian 'laws' and developmental constructs, in so far as they are theoretically sound, are ideal types. The eminent, indeed *heuristic* significance of these ideal-types when they are used for the *assessment* of reality is known to everyone who has ever employed Marxian concepts and hypotheses. Similarly their perniciousness, as soon as they are thought of as empirically valid or real (i.e. truly metaphysical) 'effective forces', 'tendencies', etc., is likewise known to those who have used them.<sup>3</sup>

One can hardly invite more controversy than by citing Marx and Weber in one breath. The point of the illustration is, however, a modest one. We may enumerate the lessons:

- 1 The materialist, such as Smart, can attach no direct sense to the reality of social science entities.
- 2 The causalist can.
- 3 The causalist may in fact reject the reality of any entities yet proposed in theoretical social science; materialist and causalist may be equally sceptical – although no more so than the founding fathers.
- 4 Weber's doctrine of ideal types displays a causalist attitude to social science laws. He uses it in a negative way. He holds that for example Marx's ideal types are not real precisely because they do not have causal powers.
- 5 The causalist may distinguish some social science from some physical science on the ground that the latter has found some entities whose causal properties are well understood, while the former has not.

<sup>3</sup> 'Objectivity in social science and social policy', German original 1904, in Max Weber, *The Methodology of the Social Sciences* (E.A. Shils and H.A. Finch, eds. and trans.), New York, 1949, p. 103.

My chief lesson here is that at least some scientific realism can use the word ‘real’ very much the same way that Austin claims is standard. The word is not notably ambiguous. It is not particularly deep. It is a substantive-hungry trouser-word. It marks a contrast. What contrast it marks depends upon the noun or noun phrase  $N$  that it modifies or is taken to modify. Then it depends upon the way that various candidates for being  $N$  may fail to be  $N$ . If the philosopher is suggesting a new doctrine, or a new context, then one will have to specify why lines of force, or the id, fail to be real entities. Smart says entities are for building. Cartwright says they are for causing. Both authors will deny, although for different reasons, that various candidates for being real entities are, in fact, real. Both are scientific realists about some entities, but since they are using the word ‘real’ to effect different contrasts, the contents of their ‘realisms’ are different. We shall now see that the same thing can happen for anti-realists.

# 3 Positivism

One anti-realist tradition has been around for a long time. At first sight it does not seem to worry about what the word ‘real’ means. It says simply: there *are* no electrons, nor any other theoretical entities. In a less dogmatic mood it says we have no good reason to suppose that any such things exist; nor have we any expectation of showing that they do exist. Nothing can be known to be real except what might be observed.

The tradition may include David Hume’s *A Treatise of Human Nature* (1739). Its most recent distinguished example is Bas van Fraassen’s *The Scientific Image* (1980). We find precursors of Hume even in ancient times, and we shall find the tradition continuing long into the future. I shall call it *positivism*. There is nothing in the name, except that it rings a few bells. The name had not even been invented in Hume’s day. Hume is usually classed as an empiricist. Van Fraassen calls himself a constructive empiricist. Certainly each generation of philosophers with a positivist frame of mind gives a new form to the underlying ideas and often chooses a new label. I want only a handy way to refer to those ideas, and none serves me better than ‘positivism’.

## Six positivist instincts

The key ideas are as follows: (1) An emphasis upon *verification* (or some variant such as *falsification*): Significant propositions are those whose truth or falsehood can be settled in some way. (2) *Pro-observation*: What we can see, feel, touch, and the like, provides the best content or foundation for all the rest of our non-mathematical knowledge. (3) *Anti-cause*: There is no causality in nature, over and above the constancy with which events of one kind are followed by events of another kind. (4) *Downplaying explanations*: Explanations may help organize phenomena, but do not provide any deeper answer to *Why* questions except to say that the phenomena regularly occur in such and such a way. (5) *Anti-theoretical entities*:

Positivists tend to be non-realists, not only because they restrict reality to the observable but also because they are against causes and are dubious about explanations. They won't infer the existence of electrons from their causal effects because they reject causes, holding that there are only constant regularities between phenomena. (6) Positivists sum up items (1) to (5) by being *against metaphysics*. Untestable propositions, unobservable entities, causes, deep explanation – these, says the positivist, are the stuff of metaphysics and must be put behind us.

I shall illustrate versions of these six themes by four epochs: Hume (1739), Comte (1830–42), logical positivism (1920–40) and van Fraassen (1980).

### **Self-avowed positivists**

The name 'positivism' was invented by the French philosopher Auguste Comte. His *Course of Positive Philosophy* was published in thick installments between 1830 and 1842. Later he said that he had chosen the word 'positive' to capture a lot of values that needed emphasis at the time. He had, he tells us, chosen the word 'positive' because of its happy connotations. In the major West European languages 'positive' had overtones of reality, utility, certainty, precision, and other qualities that Comte held in esteem.

Nowadays when philosophers talk of 'the positivists' they usually mean not Comte's school but rather the group of logical positivists who formed a famous philosophy discussion group in Vienna in the 1920s. Moritz Schlick, Rudolf Carnap, and Otto Neurath were among the most famous members. Karl Popper, Kurt Gödel, and Ludwig Wittgenstein also came to some of the meetings. The Vienna Circle had close ties to a group in Berlin of whom Hans Reichenbach was a central figure. During the Nazi regime these workers went to America or England and formed a whole new philosophical tradition there. In addition to the figures that I have already mentioned, we have Herbert Feigl and C.G. Hempel. Also the young Englishman A.J. Ayer went to Vienna in the early 1930s and returned to write his marvellous tract of English logical positivism, *Language, Truth and Logic* (1936). At the same time Willard V.O. Quine made a visit to Vienna which sowed the seeds of his doubt about some logical positivist theses, seeds which blossomed into Quine's famous denials of the analytic–synthetic

distinction and the doctrine of the indeterminacy of translation.

Such widespread influence makes it natural to call the logical positivists simply positivists. Who remembers poor old Comte, longwinded, stuffy, and not a success in life? But when I am speaking strictly, I shall use the full label ‘logical positivism’, keeping ‘positivism’ for its older sense. Among the distinctive traits of logical positivism, in addition to items (1) to (6), is an emphasis on logic, meaning, and the analysis of language. These interests are foreign to the original positivists. Indeed for the philosophy of science I prefer the old positivism just because it is not obsessed by a theory of meaning.

The usual Oedipal reaction has set in. Despite the impact of logical positivism on English-speaking philosophy, no one today wants to be called a positivist. Even logical positivists came to favour the label of ‘logical empiricist.’ In Germany and France ‘positivism’ is, in many circles, a term of opprobrium, denoting an obsession with natural science and a dismissal of alternative routes to understanding in the social sciences. It is often wrongly associated with a conservative or reactionary ideology.

In *The Positivist Dispute in German Sociology*, edited by Theodore Adorno, we see German sociology professors and their philosophical peers – Adorno, Jürgens Habermas and so forth – lining up against Karl Popper, whom they call a positivist. He himself rejects that label because he has always dissociated himself from logical positivism. Popper does not share enough of my features (1) to (6) for me to call him a positivist. He is a realist about theoretical entities, and he holds that science tries to discover explanations and causes. He lacks the positivist obsession with observation and the raw data of sense. Unlike the logical positivists he thought that the theory of meaning is a disaster for the philosophy of science. True, he does define science as the class of testable propositions, but far from decrying metaphysics, he thinks that untestable metaphysical speculation is a first stage in the formation of more testable bold conjectures.

Why then did the anti-positivist sociology professors call Popper a positivist? *Because he believes in the unity of scientific method.* Make hypotheses, deduce consequences, test them: that is Popper’s method of conjecture and refutation. He denies that there is any peculiar technique for the social sciences, any *Verstehen* that is

different from what is best for natural science. In this he is at one with the logical positivists. But I shall keep ‘positivism’ for the name of an anti-metaphysical collection of ideas (1) to (6), rather than dogma about the unity of scientific methodology. At the same time I grant that anyone who dreads an enthusiasm for scientific rigour will see little difference between Popper and the members of the Vienna Circle.

### **Anti-metaphysics**

Positivists have been good at slogans. Hume set the tone with the ringing phrases with which he concludes his *An Enquiry Concerning Human Understanding*:

When we run over libraries, persuaded of these principles, what havoc must we make? If we take in our hand any volume; of divine or school metaphysics, for instance; let us ask, *Does it contain any abstract reasoning concerning quantity or number?* No. *Does it contain any experimental reasoning concerning matter of fact and existence?* No. Commit it then to the flames: for it can contain nothing but sophistry and illusion.

In the introduction to his anthology, *Logical Positivism*, A.J. Ayer says that this ‘is an excellent statement of the positivists’ position. In the case of the logical positivists the epithet “logical” was added because they wished to annex the discoveries of modern logic.’ Hume, then, is the beginning of the criterion of verifiability intended to distinguish nonsense (metaphysics) from sensible discourse (chiefly science). Ayer began his *Language, Truth and Logic* with a powerful chapter, called ‘The elimination of metaphysics’. The logical positivists, with their passion for language and meanings, combined their scorn for idle metaphysics with a meaning-oriented doctrine called ‘the verification principle’. Schlick announced that the meaning of a statement is its method of verification. Roughly speaking, a statement was to be meaningful, or to have ‘cognitive meaning’, if and only if it was verifiable. Surprisingly, no one was ever able to define verifiability so as to exclude all bad metaphysical conversation and include all good scientific talk.

Anti-metaphysical prejudices and a verification theory of meaning are linked largely by historical accident. Certainly Comte was a great anti-metaphysician with no interest in the study of ‘meanings’. Equally in our day van Fraassen is as opposed to metaphysics.

He is of my opinion that, whatever be the interest in the philosophy of language, it has very little value for understanding science. At the start of *The Scientific Image*, he writes: 'My own view is that empiricism is correct, but could not live in the linguistic form the [logical] positivists gave it.' (p. 3)

### Comte

Auguste Comte was very much a child of the first half of the nineteenth century. Far from casting empiricism into a linguistic form, he was an historicist: that is, he firmly believed in human progress and in the near-inevitability of historical laws. It is sometimes thought that positivism and historicism are at odds with each other: quite the contrary, they are, for Comte, complementary parts of the same ideas. Certainly historicism and positivism are no more necessarily separated than positivism and the theory of meaning are necessarily connected.

Comte's model was a passionate *Essay on the Development of the Human Mind*, left as a legacy to progressive mankind by the radical aristocrat, Condorcet (1743–94). This document was written just before Condorcet killed himself in the cell from which, the following morning, he was to be taken to the guillotine. Not even the Terror of the French Revolution, 1794, could vanquish faith in progress. Comte inherited from Condorcet a structure of the evolution of the human spirit. It is defined by The Law of Three Stages. First we went through a theological stage, characterized by the search for first causes and the fiction of divinities. Then we went through a somewhat equivocal metaphysical stage, in which we gradually replaced divinities by the theoretical entities of half-completed science. Finally we now progress to the stage of positive science.

Positive science allows propositions to count as true-or-false if and only if there is some way of settling their truth values. Comte's *Course of Positive Philosophy* is a grand epistemological history of the development of the sciences. As more and more styles of scientific reasoning come into being, they thereby constitute more and more domains of positive knowledge. Propositions cannot have 'positivity' – be candidates for truth-or-falseness – unless there is some style of reasoning which bears on their truth value and can at least in principle determine that truth value. Comte, who invented

the very word ‘sociology’, tried to devise a new methodology, a new style of reasoning, for the study of society and ‘moral science’. He was wrong in his own vision of sociology, but correct in his meta-conception of what he was doing: creating a new style of reasoning to bring positivity – truth-or-falsehood – to a new domain of discourse.

Theology and metaphysics, said Comte, were earlier stages in human development, and must be put behind us, like childish things. This is not to say that we must inhabit a world denuded of values. In the latter part of his life Comte founded a Positivist Church that would establish humanistic virtues. This Church is not quite extinct; some buildings still stand, a little tatty, in Paris, and I am told that Brazil still possesses strongholds of the institution. Long ago it did flourish in collaboration with other humanistic societies, in many parts of the world. Thus positivism was not only a philosophy of scientism but a new, humanistic, religion.

### **Anti-cause**

Hume notoriously taught that cause is only constant conjunction. To say that *A* caused *B* is not to say that *A*, from some power or character within itself, brought about *B*. It is only to say that things of type *A* are regularly followed by things of type *B*. The details of Hume’s argument are analysed in hundreds of philosophy books. We may, however, miss a good deal if we read Hume out of his historical context.

Hume is in fact not responsible for the widespread philosophical acceptance of a constant-conjunction attitude to causation. Isaac Newton did it, unintentionally. The greatest triumph of the human spirit in Hume’s day was held to be the Newtonian theory of gravitation. Newton was so canny about the metaphysics of gravity that scholars will debate to the end of time what he really thought. Immediately before Newton, all progressive scientists thought that the world must be understood in terms of mechanical pushes and pulls. But gravity did not seem ‘mechanical’, for its was action at a distance. For that very reason, Newton’s only peer, Leibniz, quite rejected Newtonian gravitation: it was a reactionary reversion to inexplicable occult powers. A positivist spirit triumphed over Leibniz. We learned to think that the laws of gravity are regularities that describe what happens in the world. Then we decided that all causal laws are mere regularities!

For empirically minded people the post-Newtonian attitude was, then, this: we should not seek for causes in nature, but only regularities. We should not think of laws of nature revealing what must happen in the universe, but only what does happen. The natural scientist tries to find universal statements – theories and laws – which cover all phenomena as special cases. To say that we have found the explanation of an event is only to say that the event can be deduced from a general regularity.

There are many classic statements of this idea. Here is one from Thomas Reid's *Essays on the Active Powers of the Human Mind* of 1788. Reid was the founder of what is often called the Scottish School of Common Sense Philosophy, which was imported to form the main American philosophy until the advent of pragmatism at the end of the nineteenth century.

Natural philosophers, who think accurately, have a precise meaning to the terms they use in the science; and, when they pretend to show the cause of any phenomenon of nature, they mean by the cause, a law of nature of which that phenomenon is a necessary consequence.

The whole object of natural philosophy, as Newton expressly teaches, is reducible to these two heads: first, by just induction from experiment and observation, to discover the laws of nature; and then to apply those laws to the solution of the phenomena of nature. This was all that this great philosopher attempted, and all that he thought attainable. (I. vii. 6.)

Comte tells a similar story in his *Cours de philosophie positive*:

The first characteristic of the positive philosophy is that it regards all phenomena as subjected to invariable natural *laws*. Our business is – seeing how vain is any research into what are called *causes*, whether first or final – to pursue an accurate discovery of these laws, with a view to reducing them to the smallest possible number. By speculating upon causes, we could solve no difficulty about origin and purpose. Our real business is to analyze accurately the circumstances of phenomena, and to connect them by the natural relations of succession and resemblance. The best illustration of this is in the case of the doctrine of gravitation. We say that the general phenomena of the universe are *explained* by it, because it connects under one head the whole immense variety of astronomical facts; exhibiting the constant tendency of atoms towards each other in direct proportion to their masses, and in inverse proportion to the squares of their distances; while the general fact itself is a mere extension of one that is perfectly familiar to us and that we therefore say that we know – the weight of bodies on the surface of the earth. As to what weight and attraction are, these are questions that we regard as insoluble, which are not part of positive philosophy and which we rightly abandon to the imagination of the theologians or the subtlety of the metaphysicians. (Paris, 1830, pp. 14–16.)

Logical positivism was also to accept Hume's constant conjunction account of causes. Laws of Nature, in Mortitz Schlick's maxim, *describe* what happens, but do not *prescribe* it. They are accounts of regularities only. The logical positivist account of explanation was finally summed up in C.G. Hempel's 'deductive-nomological' model of explanation. To explain an event whose occurrence is described by the sentence *S* is to present some laws of nature (i.e. regularities) *L*, and some particular facts *F* and to show that the sentence *S* is deducible from sentences stating *L* and *F*. Van Fraassen, who has an interestingly more sophisticated account of explanation, shares the traditional positivist hostility to causes. 'Flights of fancy' he dismissively calls them in his book (for causes are even worse, in his book, than explanation).

### **Anti-theoretical-entities**

Opposition to unobservable entities goes hand in hand with an opposition to causes. Hume's scorn for the entity-postulating sciences of his day is, as always, stated in an ironic prose. He admires the seventeenth-century chemist Robert Boyle for his experiments and his reasoning, but not for his corpuscular and mechanical philosophy that imagines the world to be made up of little bouncy balls or springlike tops. In Chapter LXII of his great *History of England* he tells us that, 'Boyle was a great partisan of the mechanical philosophy, a theory which, by discovering some of the secrets of nature and allowing us to imagine the rest, is so agreeable to the natural vanity and curiosity of men.' Isaac Newton, 'the greatest and rarest genius that ever arose for the ornament and instruction of the species', is a better master than Boyle: 'While Newton seemed to draw off the veil from some of the mysteries of nature, he showed at the same time the imperfections of the mechanical philosophy, and thereby restored her ultimate secrets to that obscurity in which they ever did and ever will remain.'

Hume seldom denies that the world is run by hidden and secret causes. He denies that they are any of our business. The natural vanity and curiosity of our species may let us seek fundamental particles, but physics will not succeed. Fundamental causes ever did and ever will remain cloaked in obscurity.

Opposition to theoretical entities runs through all positivism. Comte admitted that we cannot merely generalize from observations, but must proceed through hypotheses. These must, how-

ever, be regarded only as hypotheses, and the more that they postulate, the further they are from positive science. In practical terms, Comte was opposed to the Newtonian aether, soon to be electromagnetic aether, filling all space. He was equally opposed to the atomic hypothesis. You win one, you lose one.

The logical positivists distrusted theoretical entities in varying degrees. The general strategy was to employ logic and language. They took a leaf from Bertrand Russell's notebook. Russell thought that whenever possible, inferred entities should be replaced by logical constructions. That is, a statement involving an entity whose existence is merely inferred from data is to be replaced by a logically equivalent statement about the data. In general these data are closely connected with observation. Thereby arose a great programme of reductionism for the logical positivists, who hoped that all statements involving theoretical entities would by means of logic be 'reduced' to statements that did not make reference to such entities. The failure of this project was greater even than the failure to state the verification principle.

Van Fraassen continues the positivist antipathy to theoretical entities. Indeed he will not even let us speak of theoretical entities: we mean, he writes, simply unobservable entities. These, not being seen, must be inferred. It is van Fraassen's strategy to block every inference to the truth of our theories or the existence of their entities.

### **Believing**

Hume did not believe in the invisible bouncy balls or atoms of Robert Boyle's mechanical philosophy. Newton had showed us that we ought only to seek natural laws that connect the phenomena. We should not allow our natural vanity to imagine that we can successfully seek out causes.

Comte equally disbelieved in the atoms and aether of the science of his time. We need to make hypotheses in order to tell us where to investigate nature, but positive knowledge must lie at the level of the phenomena whose laws we may determine with precision. This is not to say that Comte was ignorant of science. He was trained by the great French theoretical physicists and applied mathematicians. He believed in their laws of phenomena and distrusted any drive towards postulating new entities.

Logical positivism had no such simplistic opportunities.

Members of the Vienna Circle believed the physics of their day: some had made contributions to it. Atomism and electromagnetism had long been established, relativity was a proven success and the quantum theories were advancing by leaps and bounds. Hence arose, in the extreme version of logical positivism, a doctrine of reductionism. It was proposed that in principle there are logical and linguistic transformations in the sentences of theories that will reduce them to sentences about phenomena. Perhaps when we speak of atoms and currents and electric charges we are not to be understood quite literally, for the sentences we use are reducible to sentences about phenomena. Logicians did to some extent oblige. F.P. Ramsey showed how to leave out the names of theoretical entities in the theories, using instead a system of quantifiers. William Craig proved that for any axiomatizable theory involving both observational and theoretical terms, there exists an axiomatizable theory involving only the observational terms. But these results did not do quite what logical positivism wanted, nor was there any linguistic reduction for any genuine science. This was in terrible contrast to the remarkable partial successes by which more superficial scientific theories have been reduced to deeper ones, for example, the ways in which analytic chemistry is founded upon quantum chemistry, or the theory of the gene has been transformed into molecular biology. Attempts at scientific reduction – reducing one empirical theory to a deeper one – have scored innumerable partial successes, but attempts at linguistic reduction have got nowhere.

### **Accepting**

Hume and Comte took all that stuff about fundamental particles and said: We don't believe it. Logical positivism believed it, but said in a sense that it must not be taken literally; our theories are really talking about phenomena. Neither option is open to a present-day positivist, for the programmes of linguistic reduction failed, while on the other hand one can hardly reject the whole body of modern theoretical science. Yet van Fraassen finds a way through this impasse by distinguishing belief from acceptance.

Against the logical positivists, van Fraassen says that theories are to be taken literally. There is no other way to take them! Against the realist he says that we need not believe theories to be true. He invites us instead to use two further concepts: *acceptance* and *empirical*

*adequacy*. He defines scientific realism as the philosophy that maintains that, ‘Science aims to give us, in its theories, a literally true story of what the world is like; and acceptance of a scientific theory involves the belief that it is true’ (p. 8). His own *constructive empiricism* asserts instead that, ‘Science aims to give us theories which are empirically adequate; and acceptance of a theory involves as belief only that it is empirically adequate’ (p. 12).

‘There is,’ he writes, ‘no need to believe good theories to be true, nor to believe *ipso facto* that the entities they postulate are real.’ The ‘*ipso facto*’ reminds us that van Fraassen does not much distinguish realism about theories from realism about entities. I say that one could believe entities to be real, not ‘in virtue of the fact’ that one believes some theory to be true, but for other reasons.

A little later van Fraassen explains as follows: ‘to accept a theory is (for us) to believe that it is empirically adequate – that what the theory says *about what is observable* (by us) is true’ (p. 18). Theories are intellectual instruments for prediction, control, research and sheer enjoyment. Acceptance means commitment, among other things. To accept a theory in your field of research is to be committed to developing the programme of inquiry that it suggests. You may even accept that it provides explanations. But you must reject what has been called inference to the best explanation: to accept a theory because it makes something plain is not thereby to think that what the theory says is literally true.

Van Fraassen’s is the most coherent present-day positivism. It has all six features by which I define positivism, and which are shared by Hume, Comte and the logical positivists. Naturally it lacks Hume’s psychology, Comte’s historicism, and logical positivism’s theories of meaning, for those have nothing essential to do with the positivist spirit. Van Fraassen shares with his predecessors the *anti-metaphysics*: ‘The assertion of empirical adequacy is a great deal weaker than the assertion of truth, and the restraint to acceptance delivers us from metaphysics’ (p. 69). He is *pro-observation*, and *anti-cause*. He *downplays explanation*; he does not think explanation leads to truth. Indeed, just like Hume and Comte, he cites the classic case of Newton’s inability to explain gravity as proof that science is not essentially a matter of explanation (p. 94). Certainly he is *anti-theoretical-entities*. So he holds five of our six positivist doctrines. The only one left is the emphasis

on *verification* or some variant. Van Fraassen does not subscribe to the logical positivist verifiability theory of meaning. Nor did Comte. Nor, I think, did Hume, although Hume did have an unverifiability maxim for burning books. The positivist enthusiasm for verifiability was only temporarily connected with meaning, in the days of logical positivism. More generally it represents a desire for positive science, for knowledge that can be settled as true, and whose facts are determined with precision. Van Fraassen's constructive empiricism shares this enthusiasm.

### **Anti-explanation**

Many positivist theses were more attractive in Comte's day than our own. In 1840, theoretical entities were thoroughly hypothetical, and distaste for the merely postulated is the starting point for some sound philosophy. But increasingly we have come even to see what was once merely postulated: microbes, genes, even molecules. We have also learned how to use many theoretical entities in order to manipulate other parts of the world. These grounds for realism about entities are discussed in Chapters 10 and 16 below. However one positivist theme stands up rather well: caution about explanation.

The idea of 'inference to the best explanation' is quite old. C.S. Peirce (1839–1914) called it the method of hypothesis, or abduction. The idea is that if, confronted by some phenomenon, you find one explanation (perhaps with some initial plausibility) that makes sense of what is otherwise inexplicable, then you should conclude that the explanation is probably right. At the start of his career Peirce thought that there are three fundamental modes of scientific inference: deduction, induction and hypothesis. The older he got the more sceptical he became of the third category, and by the end of his life he attached no weight at all to 'inference to the best explanation'.

Was Peirce right to recant so thoroughly? I think so, but we need not decide that now. We are concerned only with inference to the best explanation as an argument for realism. The basic idea was enunciated by H. Helmholtz (1821–94), the great nineteenth-century contributor to physiology, optics, electrodynamics and other sciences. Helmholtz was also a philosopher who called realism

'an admirably useful and precise hypothesis'.<sup>1</sup> By now there appear to be three distinct arguments in circulation. I shall call them the simple inference argument, the cosmic accident argument, and the success of science argument.

I am sceptical of all three. I should begin by saying that explanation may play a less central a role in scientific reasoning than some philosophers imagine. Nor is *the* explanation of a phenomenon one of the ingredients of the universe, as if the Author of Nature had written down various things in the Book of the World – the entities, the phenomena, the quantities, the qualities, the laws, the numerical constants, and also the explanations of events. Explanations are relative to human interests. I do not deny that explaining – 'feeling the key turn in the lock' as Peirce put it – does happen in our intellectual life. But that is largely a feature of the historical or psychological circumstances of a moment. There are times when we feel a great gain in understanding by the organization of new explanatory hypotheses. But that feeling is not a ground for supposing that the hypothesis is true. Van Fraassen and Cartwright urge that being an explanation is never a ground for belief. I am less stringent than they: it seems to me like Peirce to be merely a feeble ground. In 1905 Einstein explained the photoelectric effect with a theory of photons. He thereby made attractive the notion of quantized bundles of light. But the ground for believing the theory is its predictive success, and so forth, not its explanatory power. Feeling the key turn in the lock makes you feel that you have an exciting new idea to work with. It is not a ground for the truth of the idea: that comes later.

### Simple inference

The simple inference argument says it would be an absolute miracle if for example the photoelectric effect went on working while there were no photons. The explanation of the persistence of this phenomenon – the one by which television information is converted from pictures into electrical impulses to be turned into electromagnetic waves in turn to be picked up on the home receiver – is

<sup>1</sup> 'On the aim and progress of physical science' (German original 1871) in H. von Helmholtz, *Popular Lectures and Addresses on Scientific Subjects* (D. Atkinson trans.), London, 1873, p. 247.

that photons do exist. As J.J.C. Smart expresses the idea: ‘One would have to suppose that there were innumerable lucky accidents about the behavior mentioned in the observational vocabulary, so that they behaved miraculously *as if* they were brought about by the non-existent things ostensibly talked about in the theoretical vocabulary.’<sup>2</sup> The realist then infers that photons are real because otherwise we could not understand how scenes are turned into electronic messages.

Even if, contrary to what I have said, explanation were a ground for belief, this seems not to be an inference to the best explanation at all. That is because the *reality* of photons is no part of the explanation. There is not, after Einstein, some further explanation, namely ‘and photons are real’, or ‘there exist photons’. I am inclined to echo Kant, and say that existence is a merely logical predicate that adds nothing to the subject. To add ‘and photons are real’, after Einstein has finished, is to add nothing to the understanding. It is not in any way to increase or enhance the explanation.

If the explainer protests, saying that Einstein himself asserted the existence of photons, then he is begging the question. For the debate between realist and anti-realist is whether the adequacy of Einstein’s theory of the photon does require that photons be real.

### Cosmic accidents

The simple inference argument considers just one theory, one phenomenon and one kind of entity. The cosmic accident argument notes that often in the growth of knowledge a good theory will explain diverse phenomena which had not hitherto been thought of as connected. Conversely, we often come at the same brute entities by quite different modes of reasoning. Hans Reichenbach called this the common cause argument, and it has been revived by Wesley Salmon.<sup>3</sup> His favoured example is not the photoelectric effect but another of Einstein’s triumphs. In 1905 Einstein also explained the Brownian movement – the way in which, as we now say, pollen particles are bounced around in a random way by being hit by molecules in motion. When Einstein’s calculations are combined

<sup>2</sup> J.J.C. Smart, ‘Difficulties for realism in the philosophy of science’, in *Logic, Methodology and Philosophy of Science VI*, Proceedings of the 6th International Congress of Logic, Methodology and Philosophy of Science, Hannover, 1979, pp. 363–75.

<sup>3</sup> Wesley Salmon, ‘Why ask, “Why?” An Inquiry Concerning Scientific Explanation’, *Proceedings and Addresses of the American Philosophical Association* 51 (1978), pp. 683–705.

with the results of careful experimenters, we are able, for example, to compute Avogadro's number, the number of molecules of an arbitrary gas contained in a given volume at a set temperature and pressure. This number had been computed from numerous quite different sources ever since 1815. What is remarkable is that we always get essentially the same number, coming at it from different routes. The only explanation must be that there *are* molecules, indeed, some  $6.023 \times 10^{23}$  molecules per gram-mole of any gas.

Once again, this seems to me to beg that realist/anti-realist issue. The anti-realist agrees that the account, due to Einstein and others, of the mean free path of molecules is a triumph. It is empirically adequate – wonderfully so. The realist asks why is it empirically adequate – is that not because there just are molecules? The anti-realist retorts that explanation is no hall-mark of truth, and that all our evidence points only to empirical adequacy. In short the argument goes around in circles (as, I contend, do all arguments conducted at this level of discussion of theories).

### The success story

The previous considerations bear more on the existence of entities; now we consider the truth of theories. We reflect not on one bit of science but on 'Science' which, Hilary Putnam tells us, is a Success. This is connected with the claim that Science is converging on the truth, as urged by many, including W. Newton-Smith in his book *Rationality* (1982). Why is Science Successful? It must be because we are converging on the truth. This issue has now been well aired, and I refer you to a number of recent discussions.<sup>4</sup> The claim that here we have an 'argument' drives me to the following additional expostulations:

1 The phenomenon of growth is at most a monotonic increase in knowledge, not convergence. This trivial observation is important, for 'convergence' implies somewhat that there is *one* thing being converged on, but 'increase' has no such implication. There can be heapings up of knowledge without there being any unity of

<sup>4</sup> Among many arguments in favour of this idea of convergence, see R.N. Boyd, 'Scientific realism and naturalistic epistemology', in P.D. Asquith and R. Giere (eds.), *PSA 1980*, Volume 2, Philosophy of Science Assn., East Lansing, Mich., pp. 613–62, and W.H. Newton-Smith, *The Rationality of Science*, London, 1981. For a very powerful statement of the opposite point of view, see L. Laudan, 'A confutation of convergent realism', *Philosophy of Science* 48 (1981), pp. 19–49.

science to which they all add up. There can also be an increasing depth of understanding, and breadth of generalization, without anything properly called convergence. Twentieth-century physics is a witness to this.

2 There are numerous merely sociological explanations of the growth of knowledge, free of realist implications. Some of these deliberately turn the ‘growth of knowledge’ into a pretence. On Kuhn’s analysis in *Structure*, when normal science is ticking over nicely, it is solving the puzzles that it creates as solvable, and so growth is built in. After revolutionary transition, the histories are rewritten so that early successes are sometimes ignored as uninteresting, while the ‘interesting’ is precisely what the post-cataclysmic science is good at. So the miraculously uniform growth is an artifact of instruction and textbooks.

3 What grows is not particularly the strictly increasing body of (nearly true) *theory*. Theory-minded philosophers fixate on accumulation of theoretical knowledge – a highly dubious claim. Several things do accumulate. (a) Phenomena accumulate. For example, Willis Lamb is trying to do optics without photons. Lamb may kill off the photons but the photoelectric effect will still be there. (b) Manipulative and technological skills accumulate – the photoelectric effect will still be opening the doors of supermarkets. (c) More interestingly to the philosopher, styles of scientific reasoning tend to accumulate. We have gradually accumulated a horde of methods, including the geometrical, the postulational, the model-building, the statistical, the hypothetico-deductive, the genetic, the evolutionary, and perhaps even the historicist. Certainly there is growth of types (a), (b), and (c), but in none of them is there any implication about the reality of theoretical entities or the truth of theories.

4 Perhaps there is a good idea, which I attribute to Imre Lakatos, and which is foreshadowed by Peirce and the pragmatism soon to be described. It is a route open to the post-Kantian, post-Hegelian, who has abandoned a correspondence theory of truth. One takes the growth of knowledge to be a given fact, and tries to characterize truth in terms of it. This is not explanation by assuming a reality, but a definition of reality as ‘what we grow to’. That may be a mistake, but at least it has an initial cogency. I describe it in Chapter 8 below.

5 Moreover, there are genuine conjectural inferences to be drawn from the growth of knowledge. To cite Peirce again, our talents at forming roughly the right expectations about the human-sized world may be accounted for by the theory of evolution. If we regularly formed the wrong expectations, we would all be dead. But we seem to have an uncanny ability to formulate structures that explain and predict both the inner constitution of nature, and the most distant realms of cosmology. What can it have benefited us, in terms of survival, that we have a brain so tooled for the lesser and the larger universe? Perhaps we should guess that people are indeed rational animals that live in a rational universe. Peirce made a more instructive if implausible proposal. He asserted that strict materialism and necessitarianism are false. The whole world is what he called ‘effete mind’, which is forming habits. The habits of inference that we form about the world are formed according to the same habits that the world used as it acquired its increased spectrum of regularities. That is a bizarre and fascinating metaphysical conjecture that might be turned into an explanation of ‘the success of science’.

How Peirce’s imagination contrasts with the banal emptiness of the Success Story or convergence argument for realism! Popper, I think, is a wiser self-professed realist than most when he writes that it never makes sense to ask for the explanation of our success. We can only have the faith to hope that it will continue. If you must have an explanation of the success of science, then say what Aristotle did, that we are rational animals that live in a rational universe.

## 4 Pragmatism

Pragmatism is the American philosophy founded by Charles Sanders Peirce (1839–1914), and made popular by William James (1842–1910). Peirce was a cantankerous genius who obtained some employment in the Harvard Observatory and the US Coast and Geodesic survey, both thanks to his father, then one of the few distinguished mathematicians in America. In an era when philosophers were turning into professors, James got him a job at Johns Hopkins University. He created a stir there by public misbehaviour (such as throwing a brick at a ladyfriend in the street), so the President of the University abolished the whole Philosophy Department, then created a new department and hired everyone back – except Peirce. Peirce did not like James's popularization of pragmatism, so he invented a new name for his ideas – pragmaticism – a name ugly enough, he would say, that no one would steal it. The relationship of pragmaticism to reality is well stated in his widely reprinted essay, 'Some consequences of four incapacities' (1868).

And what do we mean by the real? It is a conception which we must first have had when we discovered that there was an unreal, an illusion; that is, when we first corrected ourselves. . . . *The real, then, is that which, sooner or later, information and reasoning would finally result in*, and which is therefore independent of the vagaries of me and you. Thus, the very origin of the conception of reality shows that this conception essentially involves the notion of a COMMUNITY, without definite limits, and capable of a definite increase of knowledge. And so those two series of cognition – the real and the unreal – consist of those which, at a time sufficiently future, the community will always continue to reaffirm; and of those which, under the same conditions, will ever after be denied. Now, a proposition whose falsity can never be discovered, and the error of which therefore is absolutely incognizable, contains, upon our principle, absolutely no error. Consequently, that which is thought in these cognitions is the real, as it really is. There is nothing, then, to prevent our knowing outward things as they really are, and it is most likely that we do thus know them in numberless cases, although we can never be absolutely certain of doing so in any special case. (*The Philosophy of Peirce*, J. Buchler (ed.), pp. 247f.)

Precisely this notion is revived in our day by Hilary Putnam, whose ‘internal realism’ is the topic of Chapter 7.

### The road to Peirce

Peirce and Nietzsche are the two most memorable philosophers writing a century ago. Both are the heirs of Kant and Hegel. They represent alternative ways to respond to those philosophers. Both took for granted that Kant had shown that truth cannot consist in some correspondence to external reality. Both took for granted that process and possibly progress are essential characteristics of the nature of human knowledge. They had learned that from Hegel.

Nietzsche wonderfully recalls how the true world became a fable. An aphorism in his book, *The Twilight of the Idols*, starts from Plato’s ‘true world – attainable for the sage, the virtuous man’. We arrive, with Kant, at something ‘elusive, pale, Nordic, Königsbergian’. Then comes Zarathustra’s strange semblance of subjectivism. That is not the only post-Kantian route. Peirce tried to replace truth by method. Truth is whatever is in the end delivered to the community of inquirers who pursue a certain end in a certain way.

Thus Peirce is finding an objective substitute for the idea that truth is correspondence to a mind-independent reality. He sometimes called his philosophy objective idealism. He is much impressed with the need for people to attain a stable set of beliefs. In a famous essay on the fixation of belief, he considers with genuine seriousness the notion that we might fix our beliefs by following authority, or by believing whatever first comes into our heads and sticking to it. Modern readers often have trouble with this essay, because they do not for a moment take seriously that Peirce held an Established (and powerful) Church to be a very good way to fix beliefs. If there is nothing to which true belief has to correspond, why not have a Church fix your beliefs? It can be very comforting to know that your Party has the truth. Peirce rejects this possibility because he holds as a fact of human nature (not of pre-human truth) that there will in the end always be dissidents. So you want a way to fix beliefs that will fit in with this human trait. If you can have a method which is internally self-stabilizing, which acknowledges permanent fallibility and yet at the same time tends to settle down, then you will have found a better way to fix belief.

### **Repeated measurements as the model of reasoning**

Peirce is perhaps the only philosopher of modern times who was quite a good experimenter. He made many measurements, including a determination of the gravitational constant. He wrote extensively on the theory of error. Thus he was familiar with the way in which a sequence of measurements can settle down to one basic value. Measurement, in his experience, converges, and what it converges on is by definition correct. He thought that all human beliefs would be like that too. Inquiry continued long enough would lead to a stable opinion about any issue we could address. Peirce did not think that truth is correspondence to the facts: the truths are the stable conclusions reached by that unending COMMUNITY of inquirers.

This proposal to substitute method for truth – which would still warrant scientific objectivity – has all of a sudden become popular again. I think that it is the core of the methodology of research programmes of Imre Lakatos, and explained in Chapter 8. Unlike Peirce, Lakatos attends to the motley of scientific practices and so does not have the simplistic picture of knowledge settling down by a repeated and slightly mindless process of trial and error. More recently Hilary Putnam has become Peircian. Putnam does not think that Peirce's account of the method of inquiry is the last word, nor does he propose that there is a last word. He does think that there is an evolving notion of rational investigating, and that the truth is what would result from the results to which such investigation tends. In Putnam there is a double limiting process. For Peirce, there was one method of inquiry, based on deduction, induction, and, to some small degree, inference to the best explanation. Truth was, roughly, whatever hypothesizing, inducing, and testing settled down upon. That is one limiting process. For Putnam the methods of inquiry can themselves grow, and new styles of reasoning can build on old ones. But he hopes that there will be some sort of accumulation here, rather than abrupt displacement of one style of reasoning just replacing another one. There can then be two limiting processes: the long term settling into a 'rationality' of accumulated modes of thinking, and the long term settling into facts that are agreed to by these evolving kinds of reason.

## Vision

Peirce wrote on the whole gamut of philosophical topics. He has gathered about him a number of coteries who hardly speak to each other. Some regard him as a predecessor of Karl Popper, for nowhere else do we find so trenchant a view of the self-correcting method of science. Logicians find that he had many premonitions of how modern logic would develop. Students of probability and induction rightly see that Peirce had as deep an understanding of probabilistic reasoning as was possible in his day. Pierce wrote a great deal of rather obscure but fascinating material on signs, and a whole discipline that calls itself semiotics reveres him as a founding father. I think him important because of his bizarre proposal that one just is one's language, a proposal that has become a centrepiece of modern philosophy. I think him important because he was the first person to articulate the idea that we live in a universe of chance, chance that is both indeterministic, but which because of the laws of probability accounts for our false conviction that nature is governed by regular laws. A glance at the index at the end of this book will refer you to other things that we can learn from Peirce. Peirce has suffered from readers of narrow vision, so he is praised for having had this precise thought in logic, or that inscrutable idea about signs. We should instead see him as a wild man, one of the handful who understood the philosophical events of his century and set out to cast his stamp upon them. He did not succeed. He finished almost nothing, but he began almost everything.

## The branching of the ways

Peirce emphasized rational method and the community of inquirers who would gradually settle down to a form of belief. Truth is whatever in the end results. The two other great pragmatists, William James and John Dewey, had very different instincts. They lived, if not for the now, at least for the near future. They scarcely addressed the question of what might come out in the end, if there is one. Truth is whatever answers to our present needs, or at least those needs that lie to hand. The needs may be deep and various, as attested in James's fine lectures, *The Varieties of Religious Experience*. Dewey gave us the idea that truth is warranted acceptability. He thought of language as an instrument that we use to

mould our experiences to suit our ends. Thus the world, and our representation of it, seems to become at the hands of Dewey very much of a social construct. Dewey despised all dualisms – mind/matter, theory/practice, thought/action, fact/value. He made fun of what he called the spectator theory of knowledge. He said it resulted from the existence of a leisure class, who thought and wrote philosophy, as opposed to a class of entrepreneurs and workers, who had not the time for just looking. My own view, that realism is more a matter of intervention in the world, than of representing it in words and thought, surely owes much to Dewey.

There is, however, in James and Dewey, an indifference to the Peircian vision of inquiry. They did not care what beliefs we settle on in the long run. The final human fixation of belief seemed to them a chimaera. That is partly why James's rewriting of pragmatism was resisted by Peirce. This same disagreement is enacted at the very moment. Hilary Putnam is today's Peircian. Richard Rorty, in his book *Philosophy and the Mirror of Nature* (1979), plays some of the parts acted by James and Dewey. He explicitly says that recent history of American philosophy has got its emphases wrong. Where Peirce has been praised, it has been only for small things. (My section above on Peirce's vision, obviously disagrees.) Dewey and James are the true teachers, and Dewey ranks with Heidegger and Wittgenstein as the three greats of the twentieth century. However Rorty does not write only to admire. He has no Peirce/Putnam interest in the long run nor in growing canons of rationality. Nothing is more reasonable than anything else, in the long run. James was right. Reason is whatever goes in the conversation of our days, and that is good enough. It may be sublime, because of what it inspires within us and among us. There is nothing that makes one conversation intrinsically more rational than another. Rationality is extrinsic: it is whatever we agree on. If there is less persistence among fashionable literary theories than among fashionable chemical theories, that is a matter of sociology. It is not a sign that chemistry has a better method, nor that it is nearer to the truth.

Thus pragmatism branches: there are Peirce and Putnam on the one hand, and James, Dewey and Rorty on the other. Both are anti-realist, but in somewhat different ways. Peirce and Putnam optimistically hope that there is something that sooner or later,

information and reasoning would finally result in. That, for them, is the real and the true. It is interesting for Peirce and Putnam both to define the real and to know what, within our scheme of things, will pan out as real. This is not of much interest to the other sort of pragmatism. How to live and talk is what matters, in those quarters. There is not only no external truth, but there are no external or even evolving canons of rationality. Rorty's version of pragmatism is yet another language-based philosophy, which regards all our life as a matter of conversation. Dewey rightly despised the spectator theory of knowledge. What might he have thought of science as conversation? In my opinion, the right track in Dewey is the attempt to destroy the conception of knowledge and reality as a matter of thought and of representation. He should have turned the minds of philosophers to experimental science, but instead his new followers praise talk.

Dewey distinguished his philosophy from that of earlier philosophical pragmatists by calling it *instrumentalism*. This partly indicated the way in which, in his opinion, things we make (including all tools, including language as a tool) are instruments that intervene when we turn our experiences into thoughts and deeds that serve our purposes. But soon 'instrumentalism' came to denote a philosophy of science. An instrumentalist, in the parlance of most modern philosophers, is a particular kind of anti-realist about science – one who holds that theories are tools or calculating devices for organizing descriptions of phenomena, and for drawing inferences from past to future. Theories and laws have no truth in themselves. They are only instruments, not to be understood as literal assertions. Terms that seemingly denote invisible entities do not function as referential terms at all. Thus instrumentalism is to be contrasted with van Fraassen's view, that theoretical expressions are to be taken literally – but not believed, merely 'accepted' and used.

### **How do positivism and pragmatism differ?**

The differences arise from the roots. Pragmatism is an Hegelian doctrine which puts all its faith in the process of knowledge. Positivism results from the conception that seeing is believing. The pragmatist claims no quarrel with common sense: surely chairs and electrons are equally real, if indeed we shall never again come to

doubt their value to us. The positivist says electrons cannot be believed in, because they can never be seen. So it goes through all the positivist litany. Where the positivist denies causation and explanation, the pragmatist, at least in the Peircian tradition, gladly accepts them – so long as they turn out to be both useful and enduring for future inquirers.

## 5 Incommensurability

Why is so shop-soiled a topic as scientific realism once again prominent in the philosophy of science? Realism fought a great battle when Copernican and Ptolemaic world views were at issue long ago. Towards the end of the nineteenth century worries about atomism strongly contributed to anti-realism among philosophers of science. Is there a comparable scientific issue today? Maybe. One way to understand quantum mechanics is to take an idealist line. Some people argue that human observation plays an integral role in the very nature of a physical system, so that the system changes simply when it is measured. Talk of ‘the measurement problem in quantum mechanics’, the ‘ignorance interpretation’, and ‘the collapse of the wave packet’ make it no accident that contributions to the philosophy of quantum mechanics play an important part in the writings of the more original figures in the realist debate. A number of the ideas of Hilary Putnam, Bas van Fraassen or Nancy Cartwright seem to result from taking quantum mechanics as the model of all science.

Conversely, numerous physicists wax philosophical. Bernard d’Espagnat has made one of the most important recent contributions to a new realism. He is partly motivated by the dissolution, in some parts of modern physics, of old realist concepts such as matter and entity. He is especially driven by some recent results, that bear the general name of Bell’s inequality, and which have been thought to call in question concepts as various as logic, the temporal order of causation, and action at a distance. In the end he defends a realism different from any discussed in this book.

There are, then, problems within science that spur present thinking about realism. But problems of a particular science are never the whole story of a philosophical disturbance. Notoriously the Ptolemy/Copernicus debate that climaxed in the condemnation of Galileo had roots in religion. It involved our conception of the status of humanity in the universe: are we at its centre or on the

periphery? Anti-realist anti-atomism was part of late-nineteenth-century positivism. Likewise, in our time, Kuhn's historico-philosophical work has been a major element in the redisussion of realism. It is not that he single-handedly wrought a transformation in the history and philosophy of science. When his book *The Structure of Scientific Revolutions* came out in 1962, similar themes were being expressed by a number of voices. Moreover a new discipline, the history of science, was forming itself. In 1950 it was mostly the province of gifted amateurs. By 1980 it was an industry. Young Kuhn, training as a physicist, was attracted to history just at the moment when many other people were looking that way. As I have said in my Introduction, the fundamental transformation in philosophical perspective was this: science became an historical phenomenon.

This revolution had two interconnected effects on philosophers. There was the crisis of rationality that I described. There was also a great wave of doubt about scientific realism. With each paradigm shift, we come, so Kuhn hints, to see the world differently – perhaps we live in a different world. Nor are we converging on a true picture of the world, for there is none to be had. There is no progress towards the truth, but only increased technology and perhaps progress 'away from' ideas that we shall never again find tempting. Is there then a real world at all?

Within this family of ideas one catchword has had a special vogue – *incommensurability*. It has been said that successive and competing theories within the same domain 'speak different languages'. They cannot strictly be compared to each other nor translated into each other. The languages of different theories are the linguistic counterparts of the different worlds we may inhabit. We can pass from one world or one language to another by a gestalt-switch, but not by any process of understanding.

The realist about theories cannot welcome this view, in which the aim of discovering the truth about the world is dispersed. Nor is the realist about entities pleased, for all theoretical entities seem totally theory-bound. There may be electrons within our present theory, but no sense is left for the claim that there just are electrons, regardless of what we think. There have been numerous theories about electrons professed by distinguished scientists: R.A. Millikan, H.A. Lorentz, and Niels Bohr had very different ideas. The

incommensurabilist says that they meant something different, in each case, by the word ‘electron’. They were talking about different things, says the incommensurabilist, whereas the realist about entities thinks they were talking about electrons.

Hence, although incommensurability is an important topic for discussions of rationality, it also opposes scientific realism. A little care, however, makes it seem less of a dragon than is sometimes supposed.

### Kinds of incommensurability

The new philosophical use of the word ‘incommensurable’ is the product of conversations between Paul Feyerabend and Thomas Kuhn on Berkeley’s Telegraph Avenue around 1960. What did it mean before these two men refashioned it? It has an exact sense in Greek mathematics. It means ‘no common measure’. Two lengths have common measure if you can lay  $m$  of the first lengths against exactly  $n$  of the second, and thus measure the one by the other. Not all lengths are commensurable. The diagonal on a square has no common measure with the length of the sides, or, as we now express this fact,  $\sqrt{2}$  is not a rational fraction,  $m/n$ .

Philosophers have nothing so precise in mind when they use the metaphor of incommensurability. They are thinking of comparing scientific theories, but of course there could be no *exact* measure for that purpose. After twenty years of heated debate the very word ‘incommensurable’ seems to point to three distinguishable things. I shall call them *topic-incommensurability*, *dissociation*, and *meaning-incommensurability*. The first two may be fairly straightforward but the third is not.

### Accumulation and subsumption

Ernest Nagel’s *The Structure of Science* of 1961 was a classic statement of much philosophy of science that had recently been written in English. (Titles can say so much. The hit of 1962 was *The Structure of Scientific Revolutions*.) Nagel tells of stable structures and continuity. He took for granted that knowledge tends to accumulate. From time to time one theory  $T$  is replaced by a successor  $T^*$ . When is it rational to switch theories? Nagel’s idea was that the new  $T^*$  ought to be able to explain the phenomena that  $T$  explains, and it should also make whatever true predictions are

made by  $T$ . In addition, it should either exclude some part of  $T$  that is erroneous, or cover a wider range of phenomena and predictions. Ideally  $T^*$  does both. In that case  $T^*$  *subsumes*  $T$ .

When  $T^*$  subsumes  $T$  there is, loosely speaking, a common measure for comparing the two; at any rate, the correct part of  $T$  is included in  $T^*$ . So we might, by metaphor, say that  $T$  and  $T^*$  are commensurable. This very commensurability provides a basis for the rational comparison of theories.

### **Topic-incommensurability**

Feyerabend and Kuhn made it clear that Nagel did not exhaust the possibilities for theory change. A successor theory may attack different problems, use new concepts and have applications different from the old theory. It may simply forget many former successes. The ways in which it recognizes, classifies, and above all produces phenomena may not match up well with the older account. For example, the oxygen theory of burning and bleaching did not at first apply to all the phenomena that fitted nicely into phlogiston. As an historical fact it was just not true that the new theory subsumed the old one.

In Nagel's opinion  $T^*$  ought to cover the same topics as  $T$ , and cover them at least as well as  $T$ ; it should also cover some new topics. Such a sharing and extension of topics makes for commensurability between  $T$  and  $T^*$ . Kuhn and Feyerabend said that often there is a radical shift in topics. We cannot say that successor  $T^*$  does the same job better than  $T$ , because they do different jobs.

Kuhn's picture of normal science, crisis, revolution, normal science makes such topic-incommensurability quite plausible. A crisis occurs in  $T$  when a family of counterexamples attracts widespread attention, but refuses to yield to revisions in  $T$ . A revolution redescribes the counterexamples, and produces a theory, that explains previously troublesome phenomena. The revolution succeeds if the new concepts resolve certain old problems and produces new approaches and topics to investigate. The resulting normal science may ignore a lot of triumphs of the preceding normal science. Hence although there will be some overlap between  $T^*$  and  $T$ , there may be nothing like Nagel's picture of subsumption. Moreover even where there is overlap, the ways in which  $T^*$  describes some phenomena may be so different from the description

furnished by  $T$  that we may feel that these are not even understood in the same way.

In 1960, when most philosophers writing in English would have agreed with Nagel, Kuhn and Feyerabend came as a great shock. But by now topic-incommensurability by itself seems quite straightforward. It is an historical question whether the oxygen theory mostly moved on to a set of topics different from those studied by phlogiston. Doubtless there will be a great range of historical examples, starting on the one end from pure Nagelian subsumption, and arriving at the opposite extreme in which we wish to say that the successor theory totally replaced the topics, concepts and problems of  $T$ . In the extreme, students of a later generation educated on  $T^*$  may find  $T$  simply unintelligible until they play the role of historians and interpreters, relearning  $T$  from scratch.

### Dissociation

A long enough time, and radical enough shifts in theory, may make earlier work unintelligible to a later scientific audience. Here it is important to make a distinction. An old theory may be forgotten, but still be intelligible to the modern reader who is willing to spend the time relearning it. On the other hand some theories indicate so radical a change that one requires something far harder than mere learning of a theory. Two examples suffice to make the contrast.

The five-volume *Celestial Mechanics* is a great Newtonian physics book written by Laplace around 1800. The modern student of applied mathematics can understand it. This is true even toward the end of the work where Laplace writes on caloric. Caloric is a substance, the substance of heat, and it is supposed to consist of small particles with a repulsive force that decays very rapidly with distance. Laplace is proud to solve some important problems with his caloric model. He is able to provide the first derivation of the speed of sound in air. Laplace gets roughly the observed velocity, while Newton's derivations gave quite the wrong answer. We no longer believe there is such a substance as caloric, and we have entirely replaced Laplace's theory of heat. But we can work it out and understand what he is doing.

For a contrast turn to the many volumes of Paracelsus, who died in 1541. He exemplifies a Northern European Renaissance tradition

of a bundle of hermetic interests: medicine, physiology, alchemy, herbals, astrology, divination. Like many another ‘doctor’ of the day, he practised all of these as part of a single art. The historian can find in Paracelsus anticipations of later chemistry and medicine. The herbalist can retrieve some forgotten lore from his remarks. But if you try to read him you will find someone utterly different from us.

It is not that we cannot understand his words, one by one. He wrote in dog-Latin and proto-German, but that is no serious problem. He is now translated into modern German and some of his work is available in English. The tone is well suggested by passages like this: ‘Nature works through other things, such as pictures, stones, herbs, words, or when she makes comets, similitudes, halos and other unnatural products of the heavens.’ It is the ordering of thought that we cannot grasp here, for it is based on a whole system of categories that is hardly intelligible to us.

Even when we seem to be able to understand the words perfectly well, we are left in a fog. Many a Renaissance writer of high seriousness and intelligence makes extraordinary statements about the origins of ducks or geese or swans. Rotting logs floating in the Bay of Naples will generate geese. Ducks are generated from barnacles. People then knew all about ducks and geese: they had them in their barnyards nearby. Swans were kept in semi-cultivation by the ruling classes. What is the force of these absurd propositions about barnacles and logs? We do not lack sentences to express these thoughts. We have words, such as this one to be found alike in Johnson’s *Dictionary* (1755) and the *Oxford English Dictionary*: ‘*Anatiferous* – producing ducks or geese, that is producing barnacles, formerly supposed to grow on trees, and, dropping off into the water below, to turn into tree-geese.’ The definition is plain enough, but what is the point of the idea?

Paracelsus is not a closed book. One can learn to read him. One can even imitate him. There were in his day many imitations that we now call pseudo-Paracelsus. You could get sufficiently into his way of thinking to forge another volume of pseudo-Paracelsus. But to do that you would have to recreate an alien system of thought that we now only barely recall, for example, in homeopathic medicine. The trouble is not just that we think Paracelsus wrote falsely, but that we cannot attach truth or falsehood to a great many of his sentences.

His style of reasoning is alien. Syphilis is to be treated by a salve of mercury and by internal administration of the metal, because the metal mercury is the sign of the planet, Mercury, and that in turn signs the market place, and syphilis is contracted in the market place. Understanding this is an entirely different exercise from learning Laplace's theory of caloric.

Paracelsus's discourse is incommensurable with ours, because there is no way to match what he wanted to say against anything we want to say. We can express him in English, but we cannot assert or deny what is being said. At best one can start talking his way only if one becomes alienated or dissociated from the thought of our own time. Hence I shall say that the contrast between ourselves and Paracelsus is *dissociation*.

We do not strain a metaphor if we say that Paracelsus lived in a different world from ours. There are two strong linguistic correlates of dissociation. One is that numerous Paracelsan statements are not among our candidates for truth-or-falsehood. The other is that forgotten styles of reasoning are central to his thought. I argue elsewhere that these two aspects are closely connected. An interesting proposition is in general true-or-false only if there is a style of reasoning that helps one settle its truth value.<sup>1</sup> Quine and others write of conceptual schemes, by which they mean a body of sentences held for true. That is, I think, a mistaken characterization. A conceptual scheme is a network of possibilities, whose linguistic formulation is a class of sentences up for grabs as true or false. Paracelsus viewed the world as a different network of possibilities, embedded in different styles of reasoning from ours, and that is why we are dissociated from him.

Although Paul Feyerabend has spoken of incommensurability in many domains of science, his mature thoughts in *Against Method* are mostly about what I call dissociation. His prize example is the shift from archaic to classical Greece. Drawing chiefly on epic poetry and paintings on urns, he contends that Homeric Greeks literally saw things differently from Athenians. Whether or not this is correct, it is a much less surprising claim than one that says, for example, that each cohort of physicists has been referring to different things, when speaking of electrons.

<sup>1</sup> See I. Hacking, 'Language, truth and reason', in M. Hollis and S. Lukes (eds.), *Rationality and Relativism*, Oxford, 1982, pp. 48–66.

Many examples lie between the extremes of Laplace and Paracelsus. The historian soon learns that old texts constantly conceal from us the extent to which they are dissociated from our ways of thought. Kuhn tells us, for example, that Aristotle's physics relies on ideas of motion that are dissociated from ours, and one can understand him only by recognizing the network of his words. Kuhn is one of many historians to teach the need to rethink the works of our predecessors in their way, not ours.

### **Meaning-incommensurability**

The third kind of incommensurability is not historical but philosophical. It starts from asking about the meaning of terms that stand for theoretical, unobservable entities.

How do names for theoretical entities or processes get their meaning? We may have the idea that a child could grasp the use of words like 'hand' and 'sick' and 'sad' and 'horrible' by being shown things to which these words apply (including his own hands, his own sadness). Whatever be our theory of language acquisition, the manifest presence or absence of hands or sadness must be a help in catching on to what the words mean. But theoretical terms refer – almost by definition – to what cannot be observed. How do they get their meaning?

We can give some meanings by definitions. But in the case of deep theories, any definition would itself involve other theoretical terms. Moreover we seldom use definitions for starting an understanding. We explain theoretical terms by talking theory. This has long suggested that the sense of the terms is given by a string of words from the theory itself. The meaning of individual terms in the theory is given by their position within the structure of the entire theory.

On this view of meaning, it would follow that 'mass' in Newtonian theory would not mean the same 'mass' in relativistic mechanics. 'Planet' in Copernican theory will not mean the same as 'planet' in Ptolemaic theory, and indeed the sun is a planet for Ptolemy but not for Copernicus. Such conclusions are not necessarily problematic. Did not the sun itself mean something different when Copernicus put it at the centre of our system of planets? Why should it matter if we say that 'planet' or 'mass' evolved new

meanings as people thought more about planets and mass? Why should we fuss about meaning change? Because it seems to matter when we start comparing theories.

Let  $s$  be a sentence about mass, asserted by relativistic mechanics and denied by Newtonian mechanics. If the word ‘mass’ gets its meaning from its place in a theory, it will mean something different depending on whether it is used in Newtonian or relativistic mechanics. Hence the sentence  $s$ , asserted by Einstein, must differ in meaning from the sentence  $s$  denied by Newton. Indeed, let  $r$  be another sentence using the word ‘mass’, but which unlike  $s$ , is asserted by both Newton and Einstein. We cannot say that the sentence  $r$ , which occurs in the Newtonian theory, is subsumed in the relativistic theory. For ‘mass’ will not mean the same in both contexts. There will be no one proposition, the shared meaning of  $r$ , which is common to both Newton and Einstein.

That is incommensurability with a vengeance. There is no common measure for any two theories that employ theoretical terminology because in principle they can never discuss the same issues. There cannot be theoretical propositions that one theory shares with its successor. Nagel’s doctrine of subsumption then becomes logically impossible, simply because what  $T$  says cannot even be asserted (or denied) in the successor theory  $T^*$ . Such are the remarkable claims for meaning-incommensurability. One can even begin to wonder whether crucial experiments are logically possible. If an experiment is to decide between theories, would there not have to be a sentence asserting what one theory predicts and what the other denies? Can there be such a sentence?

The doctrine of meaning-incommensurability was met by cries of outrage. The whole idea was said to be incoherent. For example: no one would deny that astronomy and genetics are incommensurable – they are about different domains. But meaning-incommensurability says that competing or successive theories are incommensurable. How could we even call them competing or successive if we did not recognize them to be about the same subjects, and hence be making a comparison between them? There are other equally shallow responses to meaning-incommensurability. Then there are deep ones, of which the best is Donald Davidson’s. Davidson implies that incommensurability

makes no sense, because it rests on the idea of different and incomparable conceptual schemes. But, he urges, the very idea of a conceptual scheme is incoherent.<sup>2</sup>

At a more straightforward level it has been carefully argued, for example by Dudley Shapere, that there is enough sameness of meaning between successive theories to allow for theory comparison.<sup>3</sup> Shapere is among those, now including Feyerabend, who suppose that such matters are best discussed without bringing in the idea of meaning at all. I agree. But at the root of meaning-incommensurability is a question about how terms denoting theoretical entities get their meaning. The question presupposes a rough conception of meaning. Given that the question has been raised and such a storm provoked, we are obliged to produce a better rough conception of meaning. Hilary Putnam has honoured that obligation, and we now turn to his theory of reference in order to evade meaning-incommensurability altogether.

<sup>2</sup> D. Davidson, 'On the very idea of a conceptual scheme', *Proceedings and Addresses of the American Philosophical Association* 57 (1974), pp. 5–20.

<sup>3</sup> D. Shapere, 'Meaning and scientific change', in R. Colodny (ed.), *Mind and Cosmos: Essays in Contemporary Science and Philosophy*, Pittsburgh, 1966, pp. 41–85.

# 6 Reference

If only philosophers of science had never troubled themselves about meaning we should have no doctrine of meaning-incommensurability. As it is, we need an alternative account of meaning which allows that people holding competing or successive theories may still be talking about the same thing. The most viable alternative is Hilary Putnam's.<sup>1</sup> He intended it as a part of his former scientific realism. He has since become increasingly anti-realist, but that is a story I reserve for the next chapter. For the present consider his meaning of 'meaning'.

## Sense and reference

The word 'meaning' has many uses, many of which are more evocative than precise. Even if we stick to the commonplace meaning of words, as opposed to poems, there are at least two distinct kinds of meaning. They are distinguished in a famous 1892 essay by Gottlob Frege, 'On sense and reference'.

Consider two different kinds of answer to the question, What do you mean? Suppose I have just told you that the glyptodon brought by Richard Owen from Buenos Aires has now been restored. Most people do not know the meaning of the word 'glyptodon' and so may ask, What do you mean?

If we are standing in the museum I may simply point to a largish and preposterously shaped skeleton. *That* is what I mean. In Frege's parlance, that very skeleton is the reference of my words, 'The glyptodon brought by Richard Owen from Buenos Aires.'

On the other hand, since you probably do not have a clue what the word 'glyptodon' means, I may tell you that a glyptodon is an enormous, extinct South American mammal akin to the armadillo, but with fluted teeth. With this definition I indicate what Frege would have called the *sense* of the word 'glyptodon'.

<sup>1</sup> All references to Hilary Putnam are to 'The meaning of "meaning"' and other essays reprinted in Volume 2 of his *Philosophical Papers. Mind, Language and Reality*, Cambridge, 1979.

It is natural to think of a phrase as having a sense, namely what we understand by it, that enables us to pick out the reference, if there is one. Hearing the definition of ‘glyptodon’ I can go to a museum and try to find their skeletons, if any, without looking at the labels beneath the specimens. Frege thought that a word has a standard sense, which is what makes a scientific tradition possible. The sense is what is shared by all communicators, and may be passed down from generation to generation of students.

### **Sense and meaning-incommensurability**

Frege would have despised meaning-incommensurability but his way of looking at things helped lead into that trap. He taught us that an expression should have a definite fixed sense, which we apprehend, and which enables us to pick out the reference. Now add to this the unFregeian idea that we can grasp the sense of theoretical terms only by considering their place in a network of theoretical propositions. It seems to follow that the sense of such a term must change as the theory undergoes change.

We can evade this conclusion in several ways. One is to avoid breaking up meaning into just two components, sense and reference, with all the work being done by abstract, objective, senses. After all, the idea of meaning does not come in two nice packages that nature has labelled sense and reference. The sorting and the wrapping is the work of logicians and linguists. J.S. Mill did it in a slightly different way (connotation and denotation). So did the scholastic grammarians (intension and extension). French writers following the linguist Ferdinand de Saussure have a quite different split (signifier and signified). We may loosen Frege’s strings and tie up the parcels differently. Doubtless there are many ways to do so. Hilary Putnam’s is especially useful because, unlike all the other writers, he does not have just a pair of components of ‘meaning’.

### **Putnam’s meaning of ‘meaning’**

Dictionaries are mines of information. They do not state only abstract Fregeian senses, omitting all empirical non-linguistic facts about the world. Open one at random, and you’ll learn, say, that the French gold coin, the Louis d’or, was first struck in 1640, and continued up to the Revolution. You’ll learn that ancient Egyptian and Hindu religious art includes ritual representation of a water lily

called a lotus – and that the fruit of the mythical lotus tree is held to produce dreamy contentment. A dictionary begins an entry with some pronunciation and grammar, proceeds past etymology to a lot of information, and may conclude with examples of usage. My concise dictionary ends the entry for ‘it’ with the example: ‘It’s a dirty business, this meat-canning.’

Putnam builds his account of meaning from an analogous string of components. We may think of him as leading a back-to-the-dictionary movement. I shall use two words as examples. One is his own choice, ‘water’, and the other is our word, ‘glyptodon’.

Putnam’s first component of meaning is grammatical. He calls it a *syntactic marker*. ‘Glyptodon’ is a count noun, and ‘water’ is a mass noun. That has to do with for example the formation of plurals. We say there is some water in the pit, but either, there is a glyptodon in the pit, or, there are some glyptodons in the pit. The words have different grammars. Putnam would also include among the syntactic markers indications that both words are concrete (as opposed to abstract) names.

Putnam’s second component is a *semantic marker*. In our cases this will show the category of items to which the words apply. Both ‘water’ and ‘glyptodon’ are names of things found in nature, so Putnam enters ‘natural kind term’ among the semantic markers. Under ‘water’ he lists ‘liquid’. Under ‘glyptodon’, he would put ‘mammal’.

### Stereotypes

Putnam’s more original contribution is the third component, the *stereotype*. A stereotype is a conventional idea associated with a word, which might well be inaccurate. To use his example, a person who understands the word ‘tiger’ in our community must know that tigers are thought of as striped. Illustrations in children’s books emphasize the stripiness of tigers; that is important for showing them to be pictures of *tigers*. Even if one thought that being striped is a sort of accident, and that tigers will soon adapt to the destruction of their forests by becoming a uniform desert-tan colour, it is still true that our standard tigers are striped. You need to know that to communicate at length about tigers. But it is not a self-contradiction to speak of a tiger that has lost its stripes. An entirely white tiger has been authentically recorded. Likewise, it is part of

the stereotype of dogs that they are four-legged, even though my dog Bear has only three legs.

As part of the stereotype for ‘water’ Putnam gives us colourless, transparent, tasteless, thirst-quenching, etc. Under ‘glyptodon’ we might have enormous, extinct, South American, akin to the armadillo, with fluted teeth.

Notice that some of these elements may be mistaken. The word ‘glyptodon’ comes from the Greek for flute + tooth. It was invented by the paleontologist who discovered glyptodon remains in 1839, Richard Owen. But maybe the eponymous fluted teeth are a feature of only some glyptodonts. Every single element in the stereotype could be wrong. Maybe we shall find small glyptodonts. There were glyptodonts in North America too. Perhaps the species is not extinct, but survives far up the Amazon or the Andes. Maybe Owen was wrong about the evolutionary tree, and the animal is not akin to the armadillo.

Likewise, we may add things to the stereotype. Glyptodonts lived in the Pleistocene era. They had spiky tails with knobs on the end that could be used as clubs. They ate anything they could get their fluted teeth into. I have noticed that reference books written 70 years ago emphasize quite different features of the glyptodon from what I find today.

### **The division of linguistic labour**

The elements of Putnam’s stereotypes are not permanent criteria for the use of the word in question. A person may know the meaning of the word, and know how to use it in many situations, without knowing the present best criteria for the application of the word. I may know how to tell a glyptodon skeleton when I see one, but not be up on the criteria current among paleontologists. Putnam speaks of the division of linguistic labour. We rely on experts to know the best criteria and how to apply them. That kind of expertise is not a matter of knowing the meaning but of knowing the world.

Putnam suggests something of a hierarchy in our understanding. It is similar to the one presented by Leibniz long ago, in his *Meditations Concerning Truth and Ideas* (1684).

In the worst state, a person may simply not know what a word means. Thus in one of his papers Putnam asserts that ‘heather’ is a

synonym for ‘gorse’. That is an innocent slip that charmingly illustrates Putnam’s own distinctions. Gorse and heather are both plants characteristic of Scotland, for example, but gorse is a big shrub, spiky, with bright yellow flowers. Heather is low, soft, with tiny purple bell shaped flowers. Putnam must have not known, or forgotten, even the stereotypes for these shrubs. But it is doubtless a slip: he should have said that ‘furze’ is a synonym for ‘gorse’. Fowler’s *Modern English Usage* says that these two words are the rarest of pairs, perfect synonyms, used in the same regions interchangeably by the same speakers with no shade of difference in meaning.

Next, one may know what a word means and yet not be able to apply it correctly. Putnam, continuing his candid botanical confessions, tells us that he cannot tell a beech tree from an elm. Hence, he has what Leibniz called an *obscure* idea of a beech tree: in Leibniz’s words, ‘when my vague idea of a flower or animal which I once previously saw does not suffice for me to recognize a new instance when I encounter one’.

Next, one may be able to tell beech trees from elms, or to tell gold from other substances, without knowing the standard criteria or how to apply them. This is what Leibniz calls having a *clear* idea. One has a *distinct* idea when one knows the criteria and how to use them. Putnam and Leibniz use the same example: an assayer is an expert who knows the principles for distinguishing gold, and can apply the tests. The assayer has a distinct idea of gold.

Only a few experts have distinct ideas, that is, know the criteria appropriate in some domain. But in general we all know the meanings of the common words like ‘gold’ or ‘beech’ for which there do exist definite criteria. Perhaps these words would not have quite their present currency were there not experts in the offing. Putnam conjectures that the division of linguistic labour is an important part of any linguistic community. Note too that expert criteria may change. Assayers do not use the same techniques now as they did in the time of Leibniz. It is also common for the first stab at defining a species to flounder. Stereotypic features are recognized, but not enough is known about the things to know what is important. What then is constant in meaning? Putnam makes everything turn on reference and extension.

## Reference and extension

The *reference* of a natural kind term is the natural kind in question – if indeed there is such a natural kind. The reference of ‘water’ is a certain kind of stuff, namely H<sub>2</sub>O. The *extension* of a term is the set of things that it is true of. Thus the extension of the term glyptodon is the set of all past, present and future glyptodons. What if ‘glyptodon’ is not a natural kind? Imagine that the paleontologists made a terrible mistake, and all the fluted teeth were from one sort of animal, while the armadillo-like shell was from another. There never was a glyptodon. Then ‘glyptodon’ is not a natural kind term and the question of its extension does not arise. If it must arise, the extension is the empty set.

The Putnamian account of meaning differs from previous ones in that it includes the extension, or the reference (or both) as part of the meaning. These, and not Fregeian sense, are what are held constant from generation to generation.

## The meaning of ‘meaning’

What is the meaning of the word ‘glyptodon’? Putnam’s answer is a vector with four components: syntactic markers, semantic markers, stereotype, extension. In practice, then, we should have:

*Glyptodon*: [Concrete count noun]. [Names a natural kind, a mammal]. [Extinct, primarily South American, enormous, akin to the armadillo, has a gigantic solid shell up to five feet long with no movable rings or parts, lived during the Pleistocene era, ate anything]. [. . . . .].

Here we have nothing other than an uglified dictionary entry, except for the final square brackets that cannot be filled in. We cannot put all the glyptodons on to the page of the dictionary. Nor can we put in the natural kind. Pictorial dictionaries do their best, because they give us a photograph of a real glyptodon skeleton, or a sketch of how a glyptodon must have looked. Let us call the final [. . . . .] the *dots of extension*.

## Reference and incommensurability

Stereotypes may change as we find out more about a certain kind of thing or stuff. If we do have a genuine natural kind term, the reference of the term will remain the same, even though stereotyp-

ical opinions of the kind may change. *Thus the fundamental principle of identity for a term shifts from Fregeian sense to Putnamian reference.*

Putnam has always objected to meaning-incommensurability. The meaning-incommensurabilist says, implausibly, that whenever a theory changes, we cease to be talking about the same thing. Putnam realistically replies, that's absurd. Of course we are talking about the same thing, namely, the stable extension of the term.

When Putnam developed his theory of reference he was still a scientific realist. Meaning-incommensurability is bad for scientific realism, so it behooved Putnam to develop a theory of meaning that avoided the pitfalls of incommensurability. That is a negative result. There is also a positive one. For example, van Fraassen is an anti-realist who, like myself, thinks that the theory of meaning should have very little place in the philosophy of science. Still, he does tease the realist, who is confident that there are electrons: 'Whose electron did Millikan observe; Lorentz's, Rutherford's, Bohr's or Schrödinger's?' (*The Scientific Image*, p. 214). Putnam's account of reference provides the realist with the obvious reply: Millikan measured the charge on the electron. Lorentz, Rutherford, Bohr, Schrödinger and Millikan were all talking about electrons. They had different theories about electrons. Different stereotypes of electrons have been in vogue but it is the reference that fixes the sameness of what we are talking about.

This reply goes one dangerous step beyond what has thus far been said. In the case of water and glyptodons, there appears to be a good way of hooking up words and the world. We can at least point to some of that stuff, water; we can point to, photograph, or reconstruct a skeleton from a member of that species, glyptodon. We cannot point to electrons. We must show how Putnam's theory works on theoretical entities.

In the next few sections I describe some real-life namings. One ought to have a sense of the odd things that happen in science, as opposed to the limited range of unimaginative events that populate science fiction. It is a defect of Putnam's essays that he favours fictions over facts. The facts reveal some flaws in Putnam's simplified meaning of 'meaning'. Yet he has relieved us of the pseudo-problem of meaning-incommensurability. We do not need any theory about names in order to name electrons. (I secretly hold,

on philosophical grounds, that in principle there can be no complete, general, theory of meaning or of naming.) We need only be assured that an obviously false theory is not the only possible theory. Putnam has done that.

I should also warn of some optional extras that are sometimes added to Putnam's account. Putnam's ideas evolved at the time that Saul Kripke independently presented a remarkable set of lectures now published under the title *Naming and Necessity*. Kripke holds that when one succeeds in naming a natural kind of thing, a thing of that kind must, as part of its very essence, of its very nature, be that kind. This harks back to a philosophy due to Aristotle, called essentialism. According to Kripke, if water is in fact H<sub>2</sub>O then water is necessarily H<sub>2</sub>O. As a matter of metaphysical necessity, it cannot be anything else. Of course for all we know, it might be something else, but that is an epistemic matter. This essentialism is only accidentally connected with Putnam's meaning of 'meaning'. His references need not be 'essences'. D.H. Mellor has given strong reasons to resist that idea, at least in so far as concerns the philosophy of science.<sup>2</sup> (That is another instance of the need for philosophers of science to be chary of theories of meaning.) Despite the intrinsic interest of Kripke's ideas for students of logic they are *not* to be added here to my version of Putnam's notions.

### Dubbing the electron

New natural kinds, such as electrons, are often the result of initial speculations which are gradually articulated into theory and experiment.

Putnam urges that it is not necessary to point to an instance of a natural kind in order to pick it out and name it. Moreover, pointing is never enough. It is a well-known claim, often attributed to Wittgenstein, that any amount of pointing at examples and calling them apples is consistent with several – or indefinitely many – ways of applying the word 'apple' thereafter. Moreover, no amount of defining apples precludes, in principle, the possibility that the rule for using the word 'apple' will branch in indefinitely many different ways – not to mention our curious metaphors, such as that bit of the human neck called Adam's apple, or the oak apple, a large hard ball

<sup>2</sup> D.H. Mellor, 'Natural kinds', *British Journal for the Philosophy of Science* 28 (1977), pp. 299–312.

on Californian oaks built as the nest of a parasite. No matter how we may feel about this supposedly Wittgensteinian doctrine, it is at least clear that pointing is never enough. What pointing does do is to provide us with a causal, historical, connection between our word ‘apple’ and a certain kind of fruit, namely apples. That connection could be established in other ways, as is illustrated by the historical development of theory and experiment around the word ‘electron’.

Putnam tells a story of Bohr and the electron. Bohr, according to Putnam, had a theory about electrons. It was not a strictly correct theory, but he did draw our attention to this natural kind of thing. We should, says Putnam, apply a sort of principle of charity. Putnam calls it the principle of the benefit of the doubt, or, as he playfully puts it, the benefit of the dubbed. We may have doubts about what Bohr was doing, but given his place in our historical tradition, we should allow that he was indeed talking about electrons, albeit with an inadequate theory.

As usual I prefer truth to science fiction. Bohr did not invent the word ‘electron’, but took over a standard usage. He speculated about an already quite well-understood particle. The correct story is as follows. ‘Electron’ was the name suggested in 1891 for the natural unit of electricity. Johnstone Stoney had been writing about such a natural unit as early as 1874, and he dubbed it ‘electron’ in 1891. In 1897 J.J. Thomson showed that cathode rays consist of what were then called ‘ultratomic particles’ bearing a minimal negative charge. These particles were for long called ‘corpuscles’ by Thomson, who rightly thought he had got hold of some ultimate stuff. He determined their mass. Meanwhile, Lorentz was elaborating a theory of a particle of minimum charge which he quickly called the electron. Around 1908 Millikan measured this charge. The theory of Lorentz and others was shown to tie in rather nicely with the experimental work.

In my opinion Johnstone Stoney was speculating when he said there is a minimum unit of electric charge. We give him the benefit of the doubt, or rather, the benefit of the dubber, for he made up the name. If you like he, too, was talking about electrons (does it matter?) I have no doubt, however, in connection with Thomson and Millikan. They were well on the way to establishing the reality of these charged ultratomic particles, by experimentally determining their mass and charge. Thomson did have a false picture of the

atom, often called the pudding picture. His atom had electrons in it like currants in a British pudding. But the incommensurabilist would be out of his mind if he said that Thomson measured the mass of something other than the electron – our electron, Millikan’s electron, Bohr’s electron.

The electron provides a happy illustration of Putnam’s view of reference. We know ever so much more about electrons than Thomson did. We have regularly found that speculations about electrons and experiment on electrons can be made to mesh. In the early 1920s an experiment by O. Stern and W. Gerlach suggested that electrons have angular momentum, and soon after, in 1925, S.A. Goudsmit and G.E. Uhlenbeck had the theory of electron spin. No one at present doubts that the electron is a natural kind of fundamental importance. Many now imagine that the electron is not charged with the minimum unit of electric charge. Quarks, it is conjectured, have a charge of  $1/3 e$ , but this doesn’t hurt the reality or genuineness of electrons. It means only that one bit of the long-lived stereotype must be revised.

### Acids: bifurcating kinds

One of Putnam’s earliest examples concerns acids. ‘Acid’ does not denote a theoretical entity, but is a natural kind term like ‘water’. The incommensurabilist would say that we mean something different by the word ‘acid’ than did Lavoisier or Dalton around 1800. Our theories about acids have changed substantially, but, Putnam says, we are still talking about the same kind of stuff as those pioneers of the new chemistry.

Is Putnam correct? Certainly there is an important cluster of properties in the professional stereotype for acids: acids are substances that in water solution taste sour, and change the colour of indicators such as litmus paper. They react with many metals to form hydrogen, and react with bases to form salts.

Lavoisier and Dalton would agree completely with this stereotype. Lavoisier happened to have a false theory about such substances, for he thought every acid had oxygen in it. Indeed he defined acids in that way, but in 1810 Davy showed that was a mistake, because muriatic acid is just HCl, what we now call hydrochloric acid. But there is no doubt that Lavoisier and Davy were talking about the same stuff.

Unfortunately for Putnam's choice of example, acids are not quite such a success story as electrons. Everything went along fine until 1923. In that year J.N. Brønsted in Norway and T.M. Lowry in Britain produced one new definition of 'acid', while G.N. Lewis in the United States produced a different definition. Today there are two natural kinds: Brønsted-Lowry acids and Lewis acids. Naturally these two 'kinds' both include all the standard acids, but some substances are acids of only one of the two kinds.

A Brønsted-Lowry acid is a member of a species that has a tendency to lose a proton (while bases have a tendency to gain one). A Lewis acid belongs to a species that can accept an electron pair from a base by forming a chemical bond composed of a shared electron pair. The two definitions happen to agree about bases but not about acids, because typical Lewis acids do not contain protons, which are a precondition of being a Brønsted-Lowry acid. As I understand it, most chemists prefer the Brønsted-Lowry account for most purposes, because it appears to provide a more satisfactory explanation of many facets of acidity. On the other hand the Lewis account is used for some purposes and was originally motivated by certain analogies with the older phenomenal characteristics of acids. One authority writes: 'Numerous lengthy polemical exchanges have taken place regarding the relative merits of the Brønsted-Lowry and Lewis definitions of acids and bases. The difference is essentially one of nomenclature and has little scientific content'. Still, the philosopher of naming must ask if Lavoisier meant Brønsted-Lowry acids or Lewis acids when he spoke of acids. Obviously he meant neither. Must we now mean one or the other species? No, only for certain specialized purposes. I think this example is somewhat in the spirit of Putnam's approach to meaning. There is, however, a problem if we take him literally. The meaning of 'acid' in 1920 (i.e. before 1923) ought to have the dots of extension filled in. By Brønsted and Lowry? Or by Lewis? Since both schools of chemistry were in part enlarging the theory of acids, we could try 'all the things agreed to be acids in 1920, before the enlarging got under way'. But that is almost certainly *not* a natural kind! We could try the intersection of the two definitions, but I doubt that is a natural kind either. This example reminds us that the notion of meaning is ill-adapted to philosophy of science. We should worry about kinds of acids, not kinds of meaning.

### Caloric: the nonentity

People talk about phlogiston when they want a non-existent natural kind. Caloric is more interesting. When Lavoisier had done down the phlogiston theory, he still needed some account of heat. This was provided by caloric. Just as with the word ‘electron’, we know exactly when a substance was dubbed as caloric. It did not happen in a casual way. In 1785 there was a French chemical dubbing commission which decreed what things should be called. Many substances have been so called since that day. One new name was *calorique*, a precise term to replace one sense of the old word *chaleur*. Caloric was supposed to have no (or imponderable?) mass, and to be the substance we call heat. Not everyone accepted the official French definition. British writers would speak scathingly of ‘what the French persist in calling calorific when there is a perfectly good English word, namely fire’.

There is a tendency to regard stuff like caloric as simply stupid. That is a mistake. As I remarked in Chapter 5, it plays a real role in the final volume of Laplace’s great *Celestial Mechanics*, and not as ‘fire’ either. Laplace was a great Newtonian, and in the *Optics* Newton had speculated that the fine structure of the universe is composed of particles with forces of attraction and repulsion. The rates of extinction of these forces would vary from case to case (the rate of extinction for gravitational force is as the square of the distance). Laplace postulated different rates of extinction for both the attraction and repulsion of caloric directed to other particles. From this he was able to solve one of the outstanding problems of the century. Newtonian physics had hitherto made a complete hash of explaining the velocity of sound in air. From his assumptions about caloric Laplace was able to get a reasonable figure, very close to available experimental determinations. Laplace was justly proud of his achievement. Yet even before he published, Rumford was convincing some people that there could not be such a thing as caloric.

Caloric may seem to be no problem for Putnam’s meaning of ‘meaning’. This is a rare occasion in which we can fill in the dots of extension. The extension is the empty set. But this is too simple. Remember that Putnam was trying to explain how we and Lavoisier could both be talking about acids. Most of the answer was provided

by the dots of extension. What about caloric? The community of French revolutionary scientists – men like Berthollet, Lavoisier, Biot and Laplace – all had different theories about caloric. They were still able to talk to each other, and it seems to me they were talking about the same thing. The glib remark is, yes, the same thing, namely nothing. But these four great men were not talking about the same thing as their predecessors, who discussed phlogiston, also of zero extension. They were very glad to know that caloric is not phlogiston. Putnam's theory does not give a very good account of why 'caloric' has the same meaning for all these people: a meaning different from phlogiston. Their stereotypes for caloric were different from those for phlogiston – but not *that* different. Nor, on Putnam's theory, is it stereotypes that fix meaning. I think the lesson is that the language game of naming hypothetical entities can occasionally work well even if no real thing is being named.

### **Mesons and muons: how theories steal names from experiments**

It is easier to give old examples than recent ones because many old examples have become general knowledge. But philosophy of science loses in richness by sticking to the past. So my concluding example will be a little more up to date, and correspondingly harder to understand. It illustrates a simple point. You can baptize  $x$ 's with the new name  $N$ , and then it is decided that completely different things  $y$  are  $N$ . Some other name has to be found for  $x$ 's. Namings need not stick; they can be stolen. Anybody who thinks that reference works by a causal and historical connection to the thing named ought to reflect on the following example.

A meson is a medium weight particle, heavier than an electron, lighter than a proton. There are many kinds of mesons. A muon is rather like an electron, but 207 times heavier. Mesons are very unstable. They decay into lighter mesons and muons, and then into electrons, neutrinos and photons. Muons decay into electrons and two kinds of neutrinos. Most muons come from meson decay. Since muons are charged, they have to lose the charge when decaying. They do this by ionization, that is, by knocking electrons off atoms. Since this dissipates little energy, muons are very penetrating. They occur in cosmic rays and are that part of the ray that can travel miles under the surface of the earth to be detected deep in a mineshaft.

The fundamental fact about these two kinds of entity has to do with forces and interactions. There are four kinds of forces in the universe: electromagnetic, gravitation, weak and strong. More explanation of the latter two will be given in Chapter 16. For the present, they are just suggestive names. Strong forces bind together electrons and protons in the atom, while weak forces can be illustrated by radioactive decay. Mesons have to do with strong forces, and were originally postulated to explain how the atom stays together. They enter into strong interactions. Muons enter only into weak interactions.

As quantum mechanics became applied to electrodynamics, about 1930, there arose quantum electrodynamics, or QED for short. It has since proven to be the best theory of the universe yet devised, applying over a far wider range of phenomena and sizes of entities than anything previously known. (Perhaps it is the fulfilment of Newton's dream in the *Optics*.) In the beginning, like all physics, it would make simplifying assumptions, for example, that the electron occupies a point. It was taken for granted that some of its equations would have singularities with no solution to a real physical problem, and that one would rectify this by various *ad hoc* approximations, for example, adding extra terms to an equation.

It was at first thought that the available QED did not apply to the very penetrating particles in cosmic rays. They must be highly energized electrons, and electrons with that much energy would produce a singularity in the equations of QED. No one was much worried by this, for physics is mostly a matter of such adjustments in equations.

In 1934, H.A. Bethe and W.H. Heitler derived an important consequence of QED. It is called the energy-loss formula and it applies to electrons. In 1936 two groups of workers (C.D. Anderson and S.H. Neddermeyer; J.C. Street and E.C. Stevenson), studying cosmic rays with cloud chambers, were able to show that the energetic particles in cosmic rays did not obey the Bethe–Heitler energy-loss formula. In fact at that time QED was confirmed, contrary to expectations. The equations of QED were fine; however there was a new particle hitherto undreamt of. This was named the mesotron, because its mass lay between the electron and proton. This name was soon shortened to meson.

Meanwhile in 1935 H. Yukawa had been speculating about what

holds the atom together. He postulated that there must be a new kind of object, also intermediate in mass between electron and proton. Evidently, he was addressing a problem entirely different from cosmic rays, and there is no reason to suppose that Anderson, Neddermeyer, Street or Stevenson knew about the problems of strong forces. The speculation and the experiment were quickly put together by people like Niels Bohr and it was supposed that Yukawa's theory applied to the mesons discovered by the experimenters.

We know exactly when and how the dubbing of the experimental particle took place. Millikan wrote to the *Physical Review* as follows:<sup>3</sup>

After reading Professor Bohr's address at the British Association last September in which he tentatively suggested the name 'yucon' for the newly discovered particle, I wrote to him incidentally mentioning the fact that Anderson and Neddermeyer had suggested the name 'mesotron' (intermediate particle) as the most appropriate name. I have just received Bohr's reply to this letter in which he says:

'I take pleasure in telling you that every one at a small conference on cosmic-ray problems, including Auger, Blackett, Fermi, Heisenberg, and Rossi, which we have just held in Copenhagen, was in complete agreement with Anderson's proposal of the name "mesotron" for the penetrating cosmic-ray particles.'

Robert A. Millikan

California Institute of Technology

Pasadena, California

December 7, 1938

Note that Bohr had suggested the name 'yucon' in honour of Yukawa, but the experimentalist's name took hold by unanimous consent. Indeed there were problems from the start about the 1936 particle being what Yukawa needed – the calculated and actual lifetimes were utterly discrepant. Much later, in 1947, another particle was found in cosmic rays, while the new accelerators were starting to verify the existence of a range of related particles in scattering experiments. These were the kind of thing that Yukawa had wanted, and they came to be called  $\pi$ -mesons. The 1936 particle became a  $\mu$ -meson. After a while it was evident that they were totally different kinds of thing – a  $\pi$ -meson and a  $\mu$ -meson being

<sup>3</sup> This letter was published in *The Physical Reviews* 55 (1939) p. 105. The papers using the Bethe-Heitler energy-loss formula to reveal the original mesons (muons) are S.H. Neddermeyer and C.D. Anderson, *ibid.* 51 (1937), pp. 884–6, relying on data and photographs in *ibid.* 50 (1936), pp. 263–7. Also J.C. Street and E.C. Stevenson, *ibid.* 51 (1937), pp. 1005A.

about as unlike as any pair of entities in nature can be. The name ‘meson’ stuck with the post-1947 particles, and the 1936 particle became a muon. Histories of the subject now imply that Anderson *et al.* were actually out looking for an object to fit Yukawa’s conjecture – a conjecture they had never even heard of!

I shall return later to the question, Which comes first, theory or experiment? Chapter 9 has more examples of how theory-obsessed histories turn experimental explorations into investigations of a theory which was totally unknown to the experimenters.<sup>4</sup> For the present our concern is reference. The meson/muon story does not fit well with Putnam’s meaning of ‘meaning’. Putnam wanted to make the reference, in the end, the lynchpin of meaning. The name would apply to an entity that had been dubbed by that name on a particular historical occasion, at a baptism as it were. In our case, there was such a baptism, in 1938. However, the very name ‘mesotron’ or ‘meson’ came to mean, for the theoreticians, ‘whatever it is that satisfies Yukawa’s conjecture’. In short, the name acquired a sort of Fregeian sense. That is what took hold, baptism or no. When it was realized that this sense did not apply to the baptized object, the baptism was annulled, and a new dubbing took place.

### Meaning

Putnam’s theory of meaning works well for success stories like electrons. It is imperfect around the edges. It leaves us unhappy about bifurcating concepts, such as acidity. It does not explain how people with different theories about a nonentity such as caloric can communicate just as well with each other as people with different theories about real entities, say electrons. It relies in part on historical dubbings, the benefit of the dubbed, and a causal chain of the right sort passing from the first baptism to our present use of a name. Real communities cheerfully disregard baptisms if they want to. Those who wish a theory of meaning for scientific terms will have to improve on Putnam. They will also pay attention to the contrast between Putnam’s story and what happens, in real life, in the life sciences. This contrast has been well described by John

<sup>4</sup> In a letter to C.W.F. Everitt about our joint project, ‘Theory or experiment, which comes first?’, the Nobel laureate physicist E. Purcell suggested numerous examples of the way theory rewrites experimental history. Checking his case of the  $\mu$ -meson led me to use the example to illustrate reference-stealing as above.

Dupré.<sup>5</sup> I have only one admonition. When philosophers turn to this topic, let them not wave their hands, henceforth, about dubbings and baptisms and so forth. Let them, like Dupré, look for example at taxonomy. Let us not speak about dubbing in the abstract, but about those events in which glyptodons, caloric, electrons or mesons were named. There is a true story to be told of each. There is a real letter written by Millikan. There is a real getting together of Frenchmen to name substances, including caloric. There was even a real Johnstone Stoney. The truths about those events beat philosophical fiction any day.

I have not wanted to advance a philosophical theory of meaning. I have had only the negative purpose, of describing a theory of meaning that is pretty natural for a wide range of linguistic practice, and which does not invite talk of incommensurability. It is the kind of theory that scientific realists about entities need. It is especially attractive if one is rather cool about realism about theories. For if one expects that our theories are not strictly true, one will not want to use them to define entities in any permanent way. Rather one wants a notion of reference that is not tied by any specific, binding theory about what is referred to. A Putnamian account of reference does not, however, force you to be a realist. We must now consider why Putnam has abandoned out-and-out realism.

<sup>5</sup> 'Natural kinds and biological taxa', *The Philosophical Review* 90 (1981), pp. 66–90.

# 7 Internal realism

This chapter is probably irrelevant to scientific realism and so can well be omitted. It is about Putnam's important new 'internal realism', apparently a species of idealism.<sup>1</sup> A switch from realism to idealism sounds central to our discussion, but it is not. Putnam is no longer engaged in the debate between the scientific realist and the anti-realist about science. That debate makes a keen distinction between theoretical and observable entities. Everything Putnam now says ignores that. So it should be. His is a philosophy founded upon reflections on language, and no such philosophy can teach anything positive about natural science.

To omit Putnam's developments would nevertheless be to bypass issues of current interest. Moreover, since he finds a predecessor in Kant, we can bring in Kant's own kind of realism and idealism. Kant is a useful foil to Putnam. If we simplify and pretend that Kant too is an 'internal realist' (or that Putnam is a 'transcendental idealist') we can imagine a Kant who, unlike Putnam, emphasizes the difference between observed and inferred entities. Putnam seems to be a scientific realist within his internal realism, while we can invent a Kant who is an anti-realist about theoretical entities within a similar setting.

## Internal and external realism

Putnam distinguishes two philosophical points of view. One is 'metaphysical realism', with an 'externalist perspective' about entities and truth: 'the world consists of some fixed totality of mind-independent objects. There is exactly one true and complete description of "the way the world is". Truth involves some sort of correspondence between words or thought-signs and external things and sets of things' (p. 49).

<sup>1</sup> All references to Hilary Putnam in this chapter are to his *Reason, Truth and History*, Cambridge, 1982.

Putnam proposes instead an ‘internalist perspective’ which holds that the question,

*what objects does the world consist of?* is a question that it only makes sense to ask *within* a theory or description. . . . ‘Truth’, in an internalist view, is some sort of (idealized) rational acceptability – some sort of ideal coherence of our beliefs with each other and with our experiences *as those experiences are represented in our belief system.*

At this level internalism and pragmatism have much in common. Putnam’s position depends additionally on ideas about reference. He rejects metaphysical realism because there is never any hook-up, or correspondence, between my words and a particular batch of mind-independent entities. ‘Objects’ do not exist independently of conceptual schemes. ‘We cut up the world into objects when we introduce one sign or another. Since the objects *and* the signs are alike *internal* to the scheme of description, it is possible to say what matches what.’ (p. 52).

Putnam reports another difference between metaphysical and internal realism. The internalist says truth is optimal adequacy of theory. The externalist says that truth is, well, truth.

*Internalist:* If we had a complete theory of everything in the universe of interest to us, and the theory was thoroughly adequate by current standards of warranted assertability, rationality, or whatever, then that theory would, by definition, be true.

*Externalist:* Such a theory would very probably be true. But it is conceivable that the adequacy is a matter of luck or demonology. The theory might work for us, and yet still be a false theory about the universe.

### Queries about metaphysical realism

Putnam’s internalist can make no sense of a complete theory of the interesting universe which is entirely adequate but still false. I’m an externalist, and can make no sense of it either, *but for a different reason.* I cannot understand the idea of a complete theory of our interesting universe. *A fortiori*, I don’t understand the idea that such a theory might be adequate but false, for the idea of such a theory is itself incoherent. I can contemplate a complete theory for those wretched so-called possible worlds envisaged by logicians, but for our world? Balderdash.

Four articles were advertised on a flyer for the April 1979 *Scientific American*: How the bare hand strikes a karate blow; An enzyme clock; The evolution of disc galaxies; Oracle bones of the Shang and Chow dynasties. How could there be a complete theory of even those four topics, let alone a complete and unified theory of *everything* (including these four topics)?

How indeed could there be a complete account of even one thing or one person? P.F. Strawson remarks in his book *Individuals*, ‘The idea of an “exhaustive description” is in fact quite meaningless in general’ (p. 120). Strawson was then writing about Leibniz. Leibniz might be the best candidate for a metaphysical realist. He did think that there is a body of truth external to our own beliefs. He probably did think that there is one best, divine, description of the universe. He did think there is one totality of basic objects, namely monads. I don’t suppose he thought they are ‘mind-independent’ since monads are minds, more or less. But Leibniz did not hold a correspondence theory of truth. Even Leibniz does not fill Putnam’s bill. Was any serious thinker a metaphysical realist?

Maybe it does not matter. Putnam was describing a certain perspective, rather than a definitive theory of reality. We well recognize that externalist perspective. But here we must be careful. There could be some instances of that perspective – some kinds of external realism – which are immune to Putnam’s objections, because his objections are directed at metaphysical realism as *he* defined it.

For example, take his phrase in the definition: ‘fixed totality of mind-independent objects’. Why fixed? Why one totality? Consider only the banal example of Eddington’s – there are two tables, namely the wooden table at which I am writing, and a certain bundle of atoms. A realist about entities can well hold (a) there are mind-independent tables, (b) there are mind-independent atoms, and (c) no set of atoms is identical with this table at this instant. Atoms and tables have to do with different ways of carving up the world. There is no one fixed totality of objects. A Rubik’s cube may be a totality of 27 smaller cubes, but it need not be the case that each of these is a totality of atoms which taken together are the totality of the Rubik’s cube.

Do I not then grant Putnam’s assertion, quoted above? We cut up the world into objects when we introduce one or another scheme of

description. Yes, I grant that, metaphorically speaking. I do not grant the preceding sentence, “‘Objects’ do not exist independently of conceptual schemes.’ There are both atoms and Rubik’s cubes. To take another trite example, Inuit are said to distinguish ever so many kinds of snow that look pretty much the same to us. They cut up the frozen North by introducing a scheme of description. It in no way follows that there are not 22 distinct mind-independent kinds of snow, precisely those distinguished by the Inuit. For all I know, the powder snow, corn snow, or Sierra cement spoken of by some skiers neither contain nor are contained in any Inuit class of snow. The Inuit do not ski, and may never have wanted that category. I expect that there is still powder snow *and* all the Inuit kinds of snow, all real mind-independent distinctions in a real world.

These remarks do not prove that there is powder snow, whether anyone thinks of it or not. They merely observe that the fact that we cut up the world into various possibly incommensurable categories does not in itself imply that all such categories are mind-dependent.

Let us then be wary of the way in which Putnam runs a number of different theses together as if there were some logical connection between them.

### **Metaphysical fieldwork**

Putnam, I said, was a scientific realist who has become something of an anti-realist. Did he change sides? No. To use a gruesome analogy, he changed wars.

Scientific realism, opposed to anti-realism about science, is a *colonial war*. The scientific realist says that mesons and muons are just as much ‘ours’ as monkeys and meatballs. All of those things exist. We know it. We know some truths about each kind of thing and can find out more. The anti-realist disagrees. In the positivist tradition from Comte to van Fraassen, the phenomenal behaviour of meatballs and monkeys may be known, but talk about muons is at most an intellectual construct for prediction and control. Anti-realists about muons are realists about meatballs. I call this a colonial war because one side is trying to colonize new realms and call them reality, while the other side opposes such fanciful imperialism.

Then there is *civil* war, between say Locke and Berkeley. The realist (Locke) says that many familiar entities have an existence independent of any mental goings on: there would be monkeys even if there were no human thoughts. The idealist (Berkeley) says everything is mental. I call this a civil war because it is fought on the familiar ground of everyday experience.

Civil wars need not be fought only on home territory. Berkeley fought a colonial war too. He detested the corpuscular and mechanical philosophy of Robert Boyle. It said, in the extreme, that matter consists of bouncy springlike corpuscles (molecules, atoms, and particles, as we would say). Berkeley fought a colonial war partly because he thought that if he won, the imperialist home government of realism/materialism would collapse. Matter would be vanquished by mind.

Finally there is *total* war, chiefly a product of more recent times. Maybe Kant began it. He rejects the assumptions of civil war. Material events occur with as much certainty as mental ones. There is indeed a difference between them. Material events happen in space and time, and are ‘outer’, while mental events happen in time but not in space, and are ‘inner’. But I can know that the meatball on my plate is mush exactly as well as I know that my emotions are confusing. In general I no more infer the mushiness from my sense data than I infer that I am mixed up from my behaviour (though I could do either, on occasion).

Putnam once argued for scientific realism in a colonial war. He now argues for a position, which he says is like Kant’s, in a total war. Let us grasp Kant’s position in more detail before tackling Putnam’s.

### Kant

Kant watched his predecessors engaged in civil war. On one side there was Locke’s thesis. Kant calls it *transcendental realism*: there are objects really out there, and we infer their existence and their properties from our sense experience. Then there was Berkeley’s antithesis. Kant calls it *empirical idealism*. Matter itself does not exist; all that exists is mental.

Kant invented a synthesis to turn all this upside down. He literally reverses the labels. He calls himself an *empirical realist* and a *transcendental idealist*.

He did not go directly to his final position, but approached it by another duality. Is space merely a relative notion, as Leibniz urged and Einstein is supposed to have established? Or is it absolute, as in the Newtonian scheme? Newton had a thesis, that space and time are real. Objects occupy positions in a predetermined space and time. Leibniz voiced an antithesis, that space and time are not real. They are ideal, that is, constructs out of the relational properties of objects. Kant shilly-shallied between the two for most of his life, and then created a synthesis. Space and time are preconditions for the perception of something as an object. It is not an empirical fact that objects exist in space and time although we may experimentally determine the spatio-temporal relationships of objects within the framework of space and time. This is an *empirical realism* that grants 'the objective validity, of space, in respect of whatever can be presented to us outwardly as object'. At the same time it is a *transcendental idealism* which asserts that space 'is nothing at all . . . once we withdraw . . . its limitation to possible experience and so look upon it as something that underlies things in themselves' (p. 72).<sup>2</sup> It took Kant another decade to make this approach fit the whole range of philosophically problematic concepts. Berkeley the immaterialist had denied the existence of matter and the externality of external objects. There is nothing but mind and mental events. Kant's response: 'Matter is . . . only a species of representations (intuition) which are called external, not as standing in relation to objects in themselves external, but because they relate perceptions to the space in which all things are external to one another, while yet the space itself is within us'. Thus space itself is ideal, 'within us', and matter is properly called external because it exists as part of a system of representation within this ideal space. In order to arrive at the reality of outer objects I have just as little need to resort to inference as I have in regard to the reality of the objects of my inner sense, that is, in regard to the reality of my thoughts. For in both cases alike the objects are nothing but representations, the immediate perception (consciousness) of which is at the same time a sufficient proof of their reality.

The transcendental idealist is, therefore, an empirical realist. It is essential to Kant's point of view that what we call objects are

<sup>2</sup> All quotations from Kant are from the N. Kemp Smith translation of *The Critique of Pure Reason*, London, 1923.

constituted within a scheme, and that all our knowledge can pertain only to objects thus constituted. Our knowledge is of phenomena, and our objects lie in a phenomenal world. There are also noumena, or things in themselves, but we can have no knowledge of these. Our concepts and categories do not even apply to things-in-themselves. Philosophers from Hegel on have usually dismissed Kant's things-in-themselves. Putnam, warming to Kant, expresses gentle sympathy for the idea.

### Truth

According to Putnam, 'although Kant never quite says that this is what he is doing, Kant is best read as proposing for the first time what I have called the "internalist" or "internal realist" view of truth' (p. 60). Like so many modern philosophers, Putnam builds much of his philosophy around the idea of truth. Of Kant he says that 'there is no correspondence theory of truth in his philosophy'. Not surprising: There is *no* theory of truth in Kant's philosophy! Kant's concerns were not Putnam's. So far as affects realisms, he had two main problems:

- Are space and time real or ideal, Newtonian or Leibnizian?
- Are external objects mind-independent and Lockeian, or is everything mental and Berkeleyan?

His empirical realism and transcendental idealism is a synthesis of these oppositions and has little to do with truth. Yet Putnam's injection of a theory of truth into Kant is not strictly wrong. Putnam attributes to Kant the following ideas:

Kant does not believe that we have objective knowledge.  
 The use of the term 'knowledge' and 'objective' amount to the assertion that *there is still a notion of truth*.  
 A piece of knowledge (i.e. a 'true statement') is a statement that a rational being would accept on sufficient experience of the kind that it is possible for beings with our nature to have.  
*Truth is ultimate goodness of fit.* (p. 64).

Perhaps Putnam hit the nail on the head, particularly since he himself tends towards the pragmatist idea that truth is whatever a rational community would in due course find coherent and agree to. Kant wrote:

The holding of a thing to be true is an occurrence in our understanding which, though it may rest on objective grounds, also requires subjective

causes in the mind of the individual who makes the judgement. If the judgement is valid for everyone, provided only that he is in possession of reason, its ground is objectively sufficient. . . . Truth depends upon agreement with the object, and in respect of it the judgements of each and every understanding must therefore be in agreement with each other. . . . The touchstone, whereby we decide whether our holding a thing to be true is objective, is the possibility of communicating it and of finding it to be valid for all human reason. For there is then at least a presumption that the ground of the agreement of all judgements with each other, notwithstanding the differing characters of individuals, rests upon the common ground, namely, upon the object, and that it is for this reason that they are all in agreement with the object – the truth of the judgement being thereby proved (p. 645).

To what extent does this make Putnam mesh with Kant? I leave that to the reader. Putnam thinks warranted rational assertability and truth go hand in hand. Kant also wrote, ‘I cannot *assert* anything, that is, declare it to be a judgement necessarily valid for everyone, save as it gives rise to’ universal agreement among reasoning people (p. 646).

### Theoretical entities and things-in-themselves

Scholars do not agree about Kant’s noumenal world of things in themselves. Putnam reads Kant as saying that not only can we not describe things-in-themselves, but also ‘there is not even a one-to-one correspondence between things-for-us and things-in-themselves’. There is no horse-in-itself corresponding to the horse in the field. There is only the *noumenal world* which as a whole somehow ‘gives rise to’ our system of representation.

There have been quite different traditions of interpretation. One holds that theoretical entities are Kant’s things-in-themselves. I first find this in J.-M. Ampère (1775–1836), founder of the theory of electromagnetism. Deeply influenced by Kant, he could not tolerate the anti-realist impulses set loose on the world. He insisted that we can postulate noumena, and laws between them, to be tested in experience. This postulational and hypothetico-deductive method, said Ampère, is an intelligent investigation of the noumenal world. In our day the philosopher Wilfred Sellars holds a similar view.

There may even be an important connection, in the development of Kant’s own thought, between noumena and theoretical entities. In 1755, when he was young, Kant wrote a small physics tract called *Monadology*. This is a remarkable anticipation of our modern

theory of fields and forces. Two years later Boscovič elaborated it with far greater mathematical skill and launched field theory on the world. In Kant's early physics the world is made up of point particles – monads – separated by finite distances and exercising force fields in their environment. The properties of matter were explained by the resulting mathematical structure. *In 1755 these theoretical point particles of Kant's were his noumena.* Much later he revised this idea, and realized that there was a formal inconsistency in his theories. It could be resolved only by eliminating the things, the point particles, leaving nothing but fields of force. As a result, in the underlying structure of the universe, *there are no things, no noumena.* Then came the usual Kantian synthesis of these conflicting propositions: there are no *knowable* noumena.

Thus it is tempting to suggest that Kant's doctrine about things in themselves arose as much from his physics as his metaphysics. Kant was of little value as a scientist, but he would have been a wonderful member of a panel for a national science foundation, disbursing research money among widely different projects. He picked winners. There is what we now call the Kant–Laplace hypothesis, about the formation of the solar system. He was from the start on the side of evolutionary hypotheses about the origins of the species and the human race. He picked field theories over atomistic approaches. Now, the state of knowledge appropriate to his century was one which would downplay the significance of theoretical entities as things in themselves. There were, indeed, hypothetical stuffs of various sorts, such as the electric fluids of Franklin and many others, or the magnetic poles of Coulomb. There was an immense amount of talk about Newtonian particles and forces, but it was only at the time of Kant's death, just after the beginning of the nineteenth century, that these really got going again. Kant's attitude to the thing in itself is a quasi-scientific reaction to the modifications in his 1755 programme. Ampère, the first to preach that after all there are knowable noumena, namely the theoretical entities of the new physics and chemistry, reflects the transformation in physics. He began his career as a chemist, and was preaching knowable noumena almost as soon as he had mastered the new conjectures about the atomic structure of the elements.

What position ought Kant to have taken about theoretical entities that really do some work in science? What would he have done

when, in the twentieth century, we learned how to manipulate and even spray electrons and positrons? His own realism/idealism was directed at familiar observable objects. He denied that we infer them for our sense-data. Theoretical entities are in contrast inferred from data. Would Kant have been an empirical realist about chairs, that need no inferring, while staying an empirical anti-realist about electrons? That seems to be a possible position.

### Reference

Putnam's most original contribution concerns reference more than truth. His meaning of 'meaning' described in the previous chapter contains the seeds of its own decay. They are plain to see, for they are none other than what I called the 'dots of extension'. The meaning of a natural kind term is a sequence of elements ending in the extension, but you can't write that down.

Putnam first thought that unlike Fregeian senses, reference was unproblematic. The reference of 'glyptodon' could be indicated by pointing at a skeleton and some features in the stereotype. If glyptodons form a natural kind, nature would do the rest, and determine the extension. Theoretical entities could not be pointed at, but were to be handled by an historical story about the introduction of the terms that denote them, plus some charitable principles of the benefit of the doubt.

Putnam became sceptical. The malaise about meanings and Fregeian senses owes much to W.V.O. Quine's doctrine of the indeterminacy of translation. Quine had a parallel thesis about reference: the inscrutability of reference. To put the idea crudely: you can never tell what someone else is talking about, nor does it matter much. Quine asserted this with modest examples: where I speak of rabbits you might hear me as talking about spatio-temporal slices of rabbithood. Putnam adds real inscrutability. Whenever you talk of cats and mats, you might be referring to what I refer to when I speak of cherries and trees – yet the difference in reference would not come out, because anything I am confident of (some cat is on some mat) is expressed by a sentence which under your interpretation is something in which you have equal confidence (some cherry is on some tree).

This is indeed extraordinary. We are under two difficulties. We need to have this bizarre claim made plausible to us, and we need to

understand its place in the argument against external or metaphysical realism. Thus we need to have a local argument for the cat/cherry conclusion, and we need to have a global argument, showing how that leads to anti-metaphysical position.

### Cats and cherries

*No view which only fixes the truth-values of whole sentences can fix reference, even if it specifies truth values for sentences in every possible world.*

That is Putnam's theorem (p. 33), which we shall explain. Its cash value is presented in terms of cats and cherries. Every time you speak of cherries, you could be referring to what I call cats, and vice versa. Were I seriously to say that a cat is on a mat, you would assent, because you took me to be saying a cherry is on a tree. We can reach total agreement on the facts of the world – that is, on the sentences we hold to be true – and yet it might never appear that when I am talking about cats, you are talking about what I call cherries. Moreover your system of reference could systematically so differ from mine that the difference between us could not come out, no matter what is true about cats and cherries.

This striking conclusion follows a well-known result in mathematical logic, called the Löwenheim–Skolem theorem. The basic idea is the result of work by L. Löwenheim in 1915 and developed by Th. Skolem in 1920. In that era it seemed plausible to try to characterize mathematical objects, such as sets, by means of postulated axioms. An intended object, such as a set, would be something that fitted some postulates, and so the postulates would define the class of intended objects. Moreover we hoped to do this in the only well-understood branch of logic, called first-order logic – the logic of the sentential connectives ('and', 'not', 'or', or whatnot) and first-order quantifiers ('all', 'some').

It was thought by logicians of the day that some kind of theory of sets could serve as the foundations for many or all branches of mathematics. Georg Cantor proved a famous result. He first had clarified the idea of some infinite sets being bigger than others. Then he showed that the set of subsets of natural numbers is bigger than the set of natural numbers. In another formulation, he showed that the set of all real numbers, or of all numbers expressed as

decimal numbers, is larger than the set of natural numbers. Once this fact had been digested and accepted by classical logicians, Löwenheim and Skolem proved something that at first seemed paradoxical.

You write down some postulates that you hope capture the very essence of sets built up from sets of natural numbers. Within these postulates you prove Cantor's theorem, which says that the set of subsets of natural numbers is not denumerable, that is, it cannot be paired off with the natural numbers and so is bigger than the set of natural numbers itself. So far so good. In the way in which you intend your postulates to be understood, you are talking of Cantorian sets. Löwenheim and Skolem proved, however, that any theory, expressed in first-order logic, which is true of some domain of objects, is also true of a denumerable domain. Thus you intended your postulates to be true of Cantorian sets. Cantor's theorem at once convinces us that there are more Cantorian sets than there are natural numbers. But those very same postulates can be reinterpreted so as to be true of a much smaller domain. Suppose  $P$  is the sign which, in your theory, denotes the set of all subsets of the set of natural numbers. That is bigger than the set of natural numbers. Your theory can be reinterpreted so that  $P$  denotes something surely different, a set no bigger than the set of natural numbers.

The Löwenheim–Skolem theorem once seemed paradoxical, but it has now been digested. Most students of logic find it rather obvious, natural, and inevitable. They say things like, 'in a first-order formulation, there have to be nonstandard models'.

Putnam returns the theorem to seeming paradox. He makes a correct generalization. It applies to any domain of individuals, say cats and cherries. Take as the axioms all truths about these – all truths that I shall ever utter, or that people will ever utter, or simply all the genuine truths expressible in the first-order language. Whatever you choose, there will be unintended interpretations: moreover when we pick two kinds of objects, cats and cherries, and use a short list of truths, we can get the intended interpretation about cats to map on to the unintended interpretation about cherries. Putnam provides the details both for the short example and for the full theorem.

### The implications for scientific realism

Putnam supposes that these technical results are bad for scientific realism. Why? Largely because he thinks that scientific realism is in the end a copy or correspondence theory of truth. Our theories are true because they represent the world, and they latch on to the world by referring to objects – a reference which Putnam now thinks makes sense only within a system of beliefs.

Much of this position is well known. It is a longstanding criticism of correspondence theories that the sentences are supposed to correspond to facts, but there is no way to distinguish the facts except in terms of the sentences to which they correspond. G.E. Moore is not notable for his anti-realism, but here is how he expressed the idea 80 years ago, in an article on ‘Truth’ published in Baldwin’s *Dictionary of Philosophy*:

It is commonly supposed that the truth of a proposition consists in some relation which it bears to reality; and falsehood in the absence of this relation. The relation in question is generally called a ‘correspondence’ or ‘agreement’, and it seems to be generally conceived as one of partial similarity; but it is to be noted that only propositions can be said to be true in virtue of their partial similarity to something else, and hence that it is essential to the theory that a truth should differ in some specific way from the reality, in relation to which its truth is to consist, in every case except that in which the reality is itself a proposition. It is the impossibility of finding any such difference between a truth and the reality to which it is supposed to correspond which refutes the theory.

It has been argued, for example by J.L. Austin, that correspondence theories do have merit, because, contrary to Moore, there is an independent way to pick out facts. There are, first of all, independent ways to pick out things and qualities we are talking about – by pointing, for example. Then we make assertions by connecting referring expressions and names for properties and relations. A proposition is true just if the property named is possessed by the object referred to. Putnam must suppose that his use of the Löwenheim–Skolem vitiates this Austinian move, by showing once again that there is no way to make independent reference. But all he has shown is that you cannot succeed in reference by stating a set of truths expressed in first-order logic. When we look more closely at the Löwenheim–Skolem theorem, we

recall that it has premises. There are ways of evading these premises and thus casting doubt on Putnam's conclusions.

### Premises

1 The Löwenheim–Skolem theorem is about sentences in first-order logic. No one has ever shown that the commonplace language of physicists can ever be squeezed into a first-order format. So the argument is not known to be relevant to, say, quantum electrodynamics, and hence not to scientific realism.

2 There is a weighty school of thought, deriving impetus from the late Richard Montague, that ordinary English primarily deploys second-order quantifiers. In no direct way does the Löwenheim–Skolem theorem extend to such languages, so the applicability of Putnam's work to plain prescientific English is controversial.

3 Much common speech involves what are called indexicals. These are words whose reference depends on the context of utterance: this, that, you, me, here, now (not to mention our tensed verbs). As I walk out this fine morning I overhear: 'Hey you, stop picking my cherries, come here this instant!' Only dogma could insist that this ordinary sentence is expressible in first-order logic.

4 Introduction of indexicals goes only part of the way. Indexicals are pointers, but they are still linguistic. Language is embedded in a wide range of doings in the world. Putnam oddly refers to Wittgenstein during his discussion, recalling Wittgenstein's argument that meanings cannot be exhaustively given by rules. That did not mean, for Wittgenstein, that there was something intrinsically indeterminate and open to reinterpretation in our linguistic *practice*. It meant that language is more than talking. This is no place to expound a version of his insights, but cherries are for eating, cats, perhaps, for stroking. Once speech becomes embedded in action, talk of Löwenheim and Skolem seems scholastic. They were entirely right in what they said about a certain view of mathematical objects. They wisely refrained from discussing cats. We can do nothing with very large numbers except talk about them. With cats we relate in other ways than speech.

5 Putnam says that whatever theory we propound about reference and denotation, words such as 'denote' and 'refer' can themselves be reinterpreted. Suppose I say that 'cat' denotes

animals like those on my lap. He asks: How do I know that ‘denotes’ denotes denoting? But of course I never normally use words such as ‘denote’ in explaining the usage of words. That function may be served by ‘That is a glyptodon skeleton’, used to explain what a glyptodon is. I do not need a theory of reference in order to refer, and it is at least arguable, on grounds possibly learned from Wittgenstein, that there could be no general theory of reference.

6 Putnam is writing about unscientific anti-realism, so it is right to discuss cherries and cats. Might we not grant him his point for the theoretical entities of natural science? Is not the dubbing of entities with names entirely at the level of language? No, often it is not. Look at the 1936 paper of Anderson and Neddermeyer, mentioned in the last chapter. That is the one with the data on the basis of which the physics community dubbed the mesotron or meson – later muon. The paper is full of photographs. Not snapshots of muons, but tracks. It measures angles between the tracks caused by the collisions of this and that. We do use indexicals as brief as ‘this’ and ‘that’ to point to the most theoretical of entities – not by pointing at them, but by pointing at their traces. Not that we stop there. As is clear from my previous chapter, people at first were pretty unsure about those things that came to be called muons. But now for example we know that the mass of the muon is 206.768 times that of the electron.

This last sentence will seem grist to Putnam’s mill. For that is just the sort of truth we can put in as an axiom in an account of muons. Can we not then expose it to Löwenheim–Skolem reinterpretation? I do not think so, for how did we get this fine number to three places of decimals? It is a rather complicated computation in which we determine a whole bunch of quantities, such as the magnetic moment of the free electron, the Bohr magneton, and other fancy stuff, and in particular, relationships between a number of constants of nature. Now if these were just a bunch of sentences, and we could do all the mathematical physics in terms of first-order logic, the Löwenheim–Skolem theorem would apply. But in every case the numbers and ratios are intimately connected to specific experimental determinations. These in turn are all connected up with people, places, and, above all, doings. (Typical example: the University of Washington–Lawrence Radiation Laboratory group, i.e. K.M. Crowe, J.F. Hague, J.E. Rotherberg, A. Schenck, D.L.

Williams, R.W. Williams and K.K. Young, *Phys. Rev. D.* 2145 (1972).) Nor is it just one such set of doings, but lots of independent but not totally dissimilar *doings* all over the world.

7 Putnam does address the question of whether humans could ever use his unintended interpretation of the word 'cat'. He notes a symmetry between intended and unintended interpretations – everything we explain in terms of cats, others can explain in terms of cherries. He reiterates a discussion that derives from Nelson Goodman's book *Fact, Fiction and Forecast*. There is an important fact that he ignores. The Löwenheim–Skolem theorem is non-constructive. That is, there is in principle no humanly accessible way to generate an unintended interpretation.

8 Nor do we need technical examples to begin to query Putnam's confidence. Putnam cites his colleague Robert Nozick, as suggesting that (in Putnam's view) all women might mean cats when they speak of cherries, whereas 'we' men mean cherries. But there are for example nominal adjectives, illustrated by Bing cherries and Persian cats. Nominal adjectives such as 'Bing' are not ordinary modifiers like 'sweet', for sweet Bing cherries are sweet fruit, but they are not 'Bing fruit'. How is the Putnam/Nozick reinterpretation continued? Do their fantasy women mean Persian cats when they speak of Queen Anne cherries? That is, does one kind of cherry map on to one kind of cat? That won't do, for the number of kinds of cherries is different from the number of kinds of cats, so no such mapping will preserve the structure of nominal adjectives. More importantly, Queen Anne cherries are for preserving or for pies, while Bing cherries are for eating ripe from the tree. How are these facts to show up in the structure of facts about cats?

Putnam perhaps commits one of the gravest errors of philosophy. He has an abstract theorem. Then he explains its content in terms of one sentence that no one before him has ever uttered, nor would commonly have any point outside logic in uttering: 'Some cherry is on some tree'. Then he passes to the assertion that just as you can reinterpret 'cherry' you can reinterpret 'denote'. All the flourishing ordinary world of making a pie out of Queen Anne cherries, of determining the ratio of the masses of muons and electrons – all that is left out.

I shall not continue. I wanted only to emphasize that (a) assuring reference is not primarily a matter of uttering truths, but of

interacting with the world, and that (b) even at the level of language there is vastly more structure than Putnam discusses, be it deep questions about the language of mathematical physics, or trivial observations about Bing cherries.

### Nominalism

The above reflections do not mean you need dissent from Putnam's underlying philosophy. They mean only that what looks like a spiffy argument needs more polishing than it has yet received. What is the underlying point of view? I have followed Putnam in comparing his ideas to Kant, but there is a significant difference. Kant called himself a transcendental idealist. I would call Putnam a transcendental nominalist. Both are kinds of anti-realism. Before Kant, realism usually meant anti-nominalism. After Kant, it usually meant anti-idealism.

Idealism is a thesis about *existence*. In its extreme form it says that all that exists is mental, a production of the human spirit.

Nominalism is about *classification*. It says that only our modes of thinking make us sort grass from straw, flesh from foliage. The world does not have to be sorted that way; it does not come wrapped up in 'natural kinds'. In contrast the Aristotelian realist (the anti-nominalist) says that the world just comes in certain kinds. That is nature's way, not man's.

The idealist need have no opinion about classification. He may hold that there is indeed a real distinction between grass and straw. He says only that there is no stuff, grass and straw; there are only ideas, mental entities. But the ideas could well have real essences.

Conversely the nominalist does not deny that there is real stuff, existing independent of the mind. He denies only that it is naturally and intrinsically sorted in any particular way, independent of how we think about it.

In fact nominalism and idealism tend to be part of the same cast of mind. That is one reason that the word 'realism' has been used to denote opposition to either doctrine. But the two are logically distinct.

I read Kant in a possibly extreme way. He thought that space and time are ideal. They literally do not exist. Although there are empirical relations determinable within space and time, those relations, being spatio-temporal, have no existence beyond the

mind. Kant was indeed a transcendental *idealist*. Putnam is instead a transcendental *nominalist*.

Putnam's internal realism comes to this: Within my system of thought I refer to various objects, and say things about those objects, some true, some false. However, I can never get outside my system of thought, and maintain some basis for reference which is not part of my own system of classification and naming. That is precisely empirical realism and transcendental nominalism.

### **Revolutionary nominalism**

T.S. Kuhn has also been read as an idealist. I think he too is better understood as a transcendental nominalist – one who got there before Putnam. But whereas Putnam's reflections are based on an *a priori* theorem and alleged implications for language, Kuhn has more of a real-life basis for his position.

A scientific revolution, in Kuhn's opinion, produces a new way of addressing some aspect of nature. It provides models, conjectured laws, classes of entities, causal powers which did not enter into the predecessor science. In a completely uncontroversial sense we may now live in a different world from the nineteenth-century age of steam – a world in which aeroplanes are everywhere and railways are going bankrupt. More philosophically (perhaps) it is a different world, in that it is categorized in new ways, thought of as filled with new potentialities, new causes, new effects. But this novelty is not the production of new entities in the mind. It is the imposition of a new system of categories upon phenomena, including newly created phenomena. That is why I call it a kind of nominalism. Here is a recent formulation of Kuhn's own:

What characterizes revolutions is, thus, change in several of the taxonomic categories prerequisite to scientific descriptions and generalizations. That change, furthermore, is an adjustment not only of criteria relevant to categorization, but also of the way in which given objects and situations are distributed among pre-existing categories. Since such redistribution always involves more than one category and since those categories are interdefined, this sort of alteration is necessarily holistic.<sup>3</sup>

Kuhn is no old fashioned nominalist. That would be someone who thought that all our classifications were a product of the human

<sup>3</sup> T.S. Kuhn, 'What are scientific revolutions?' Center for Cognitive Science Occasional Paper 18, Massachusetts Institute for Technology, 1981, p. 25.

mind, not the world, and that those classifications were all the same absolutely stable features of our minds. He can disagree with that nominalist on both counts. Obviously he favours the possibility of revolutionary change, and he furnishes us with examples of it. He can equally assert that many of our prescientific categories *are* natural kinds: people and grass, flesh and horseflesh. The world simply does have horses and grass in it, no matter what we think, and any conceptual scheme will acknowledge that. There is no reason that the history of science should deny that the world sorts itself in these ways. Nor is there much reason, in the comparative study of cultures, to suppose that any other people fail to sort in similar ways. Kuhn's nominalism, in so far as it is founded upon his historical studies, could teach only that some of our scientific categories can be dislodged. Time-honoured categories, such as substance and force, may go under. Time and space may even take a beating. Kuhn does teach a certain relativism, that there is no uniquely right categorization of any aspect of nature. Indeed the idea of an aspect of nature, of comprising just such and such affairs, is itself a variable. The Greeks, we say, had no concept of electricity, Franklin no concept of electricity-and-magnetism. Even such 'aspects of nature' emerge, weave in and out, during our history. The revolutionary nominalist infers that we have not reached the end of the road. Nor is the notion of an end of the road, of a final science, a truly comprehensible one.

The old-fashioned nominalist of times gone by held that our systems of classification are products of the human mind. But he did not suppose that they could be radically altered. Kuhn has changed all that. The categories have been altered and may be altered again. We can hardly avoid approaching nature with our present categories, problems, systems of analysis, methods of technology and of learning. We are in fact empirical realists: we think as if we are using natural kinds, real principles of sorting. Yet in the course of historical reflection we realize that the inquiries most dear to us may be replaced.

To sum up the idea: we do investigate nature as sorted into the natural kinds delivered by our present sciences, but at the same time hold that these very schemes constitute only an historical event. Moreover, there is no concept of *the* right, final representation of the world.

Putnam's remarks might incline one in the same direction, but there is one sense in which his present rendition is rather Kantian. Putnam has become conservative. For Kant there was no way out of our conceptual scheme. Putnam gives no reason to suppose there is any way either. Kuhn details ways in which there have been profound alterations. Thus his is a revolutionary transcendental nomenalism, whereas Putnam's is more conservative.

### Rationality

There is another stand in Putnam's present position, reminiscent of Peirce. He holds that what is true is whatever we come to agree on by rational means, and he acknowledges that there may be at least evolution as we develop more and more styles of reasoning. I find it natural to explain this not in terms of Putnam's philosophy, but rather in terms of that of Imre Lakatos.

## 8 A surrogate for truth

‘Mob psychology’ – that is how Imre Lakatos (1922–74) caricatured Kuhn’s account of science. ‘Scientific method (or “logic of discovery”), conceived as the discipline of rational appraisal of scientific theories – and of criteria of *progress* – vanishes. We may of course still try to explain changes in “paradigms” in terms of social psychology. This is . . . Kuhn’s way’ (I, p. 31).<sup>1</sup> Lakatos utterly opposed what he claimed to be Kuhn’s reduction of the philosophy of science to sociology. He thought that it left no place for the sacrosanct scientific values of truth, objectivity, rationality and reason.

Although this is a travesty of Kuhn the resulting ideas are important. The two current issues of philosophy of science are epistemological (rationality) and metaphysical (truth and reality). Lakatos *seems* to be talking about the former. Indeed he is universally held to present a new theory of method and reason, and he is admired by some and criticized by others on that score. If that is what Lakatos is up to, his theory of rationality is bizarre. It does not help us at all in deciding what it is reasonable to believe or do now. It is entirely backward-looking. It can tell us what decisions in past science were rational, but cannot help us with the future. In so far as Lakatos’s essays bear on the future they are a bustling blend of platitudes and prejudices. Yet the essays remain compelling. Hence I urge that they are about something other than method and rationality. He is important precisely because he is addressing, not an epistemological issue, but a metaphysical one. He is concerned with truth or its absence. He thought science is our model of objectivity. We might try to explain that, by holding that a scientific proposition must say how things are. It must correspond to the truth. That is what makes science objective. Lakatos, educated in Hungary in an Hegelian and Marxist tradition, took for granted the

<sup>1</sup> All references to Imre Lakatos in this chapter are to his *Philosophical Papers*, 2 Volumes (J. Worrall and G. Currie, eds.), Cambridge, 1978.

post-Kantian, Hegelian, demolition of correspondence theories. He was thus like Peirce, also formed in an Hegelian matrix, and who, with other pragmatists, had no use for what William James called the copy theory of truth.

At the beginning of the twentieth century philosophers in England and then in America denounced Hegel and revived correspondence theories of truth and referential accounts of meaning. These are still central topics of Anglophone philosophy. Hilary Putnam is instructive here. In *Reason, Truth and History* he makes his own attempt to terminate correspondence theories. Putnam sees himself as entirely radical, and writes ‘what we have here is the demise of a theory that lasted for over two thousand years’ (p. 74). Lakatos and Peirce thought the death in the family occurred about two hundred years earlier. Yet both men wanted an account of the objective values of Western science. So they tried to find a substitute for truth. In the Hegelian tradition, they said it lies in process, in the nature of the growth of knowledge itself.

### **A history of methodologies**

Lakatos presented his philosophy of science as the upshot of an historical sequence of philosophies. This sequence will include the familiar facts about Popper, Carnap, Kuhn, about revolution and rationality, that I have already described in the Introduction. But it is broader in scope and far more stylized. I shall now run through this story. A good many of its peripheral assertions were fashionable among philosophers of science in 1965. These are simplistic opinions such as: there is no distinction in principle between statement of theory and reports of observation; there are no crucial experiments, for only with hindsight do we call an experiment crucial; you can always go on inventing plausible auxiliary hypotheses that will preserve a theory; it is never sensible to abandon a theory without a better theory to replace it. Lakatos never gives a good or even a detailed argument for any of these propositions. Most of them are a consequence of a theory-bound philosophy and they are best revised or refuted by serious reflection on experimentation. I assess them in Part B, on Intervening. On crucial experiments and auxiliary hypotheses, see Chapter 15. On the distinctions between observation and theory, see Chapter 10.

### **Euclidean model and inductivism**

In the beginning, says Lakatos, mathematical proof was the model of true science. Conclusions had to be demonstrated and made absolutely certain. Anything less than complete certainty was defective. Science was by definition infallible.

The seventeenth century and the experimental method of reasoning made this seem an impossible goal. Yet the tale is only modified as we pass from deduction to induction. If we cannot have secure knowledge let us at least have probable knowledge based on sure foundations. Observations rightly made shall serve as the basis. We shall generalize upon sound experiments, draw analogies, and build up to scientific conclusions. The greater the variety and quantity of observations that confirm a conclusion, the more probable it is. We may no longer have certainty, but we have high probability.

Here then are two stages on the high road to methodology: proof and probability. Hume, knowing the failure of the first, already cast doubts on the second by 1739. In no way can particular facts provide ‘good reason’ for more general statements or claims about the future. Popper agreed, and so in turn does Lakatos.

### **Falsificationism**

Lakatos truncates some history of methodology but expands others. He even had a Popper<sub>1</sub>, Popper<sub>2</sub>, and a Popper<sub>3</sub>, denoting increasingly sophisticated versions of what Lakatos had learned from Popper. All three emphasize the testing and falsifying of conjectures rather than verifying or confirming them. The simplest view would be, ‘people propose, nature disposes’. That is, we think up theories, and nature junks them if they are wrong. That implies a pretty sharp distinction between fallible theories and basic observations of nature. The latter, once checked out, are a final and indubitable court of appeal. A theory inconsistent with an observation must be rejected.

This story of conjecture and refutation makes us think of a pleasingly objective and honest science. But it won’t do: for one thing ‘all theories are born refuted’, or at least it is very common for a theory to be proposed even when it is known not to square with all

the known facts. That was Kuhn's point about puzzle-solving normal science. Secondly (according to Lakatos), there is no firm theory–observation distinction. Thirdly there is a claim made by the great French historian of science, Pierre Duhem. He remarked that theories are tested via auxiliary hypotheses. In his example, if an astronomer predicts that a heavenly body is to be found in a certain location, but it turns up somewhere else, he need not revise his astronomy. He could perhaps revise the theory of the telescope (or produce a suitable account of how phenomena differ from reality (Kepler), or invent a theory of astronomical aberration (G.G. Stokes), or suggest that the Doppler effect works differently in outer space). Hence a recalcitrant observation does not necessarily refute a theory. Duhem probably thought that it is a matter of choice or convention whether a theory or one of its auxiliary hypotheses is to be revised. Duhem was an outstanding anti-realist, so such a conclusion was attractive. It is repugnant to the staunch instincts for scientific realism found in Popper or Lakatos.

So the falsificationist adds two further props. First, no theory is rejected or abandoned unless there is a better rival theory in existence. Secondly, one theory is better than another if it makes more novel predictions. Traditionally theories had to be consistent with the evidence. The falsificationist, says Lakatos, demands not that the theory should be consistent with the evidence, but that it should actually outpace it.

Note that this last item has a long history of controversy. By and large inductivists think that evidence consistent with a theory supports it, no matter whether the theory preceded the evidence or the evidence preceded the theory. More rationalistic and deductively oriented thinkers will insist on what Lakatos calls '*the Leibniz–Whewell–Popper requirement that the – well planned – building of pigeon holes must proceed much faster than the recording of facts which are to be housed in them*' (I, p. 100).

### Research programmes

We might take advantage of the two spellings of the word, and use the American spelling 'research program' to denote what investigators normally call a research program, namely a specific attack on a problem using some well-defined combination of theoretical and

experimental ideas. A research program is a program of research which a person or group can undertake, seek funding for, obtain help with, and so on. What Lakatos spells as ‘research programme’ is not much like that. It is more abstract, more historical. It is a sequence of developing theories that might last for centuries, and which might sink into oblivion for 80 years and then be revived by an entirely fresh infusion of facts or ideas.

In particular cases it is often easy to recognize a continuum of developing theories. It is less easy to produce a general characterization. Lakatos introduces the word ‘heuristic’ to help. Now ‘heuristic’ is an adjective describing a method or process that guides discovery or investigation. From the very beginnings of Artificial Intelligence in the 1950s, people spoke of heuristic procedures that would help machines solve problems. In *How to solve it* and other wonderful books, Lakatos’s countryman and mentor, the mathematician Georg Polya, provided classic modern works on mathematical heuristics. Lakatos’s work on the philosophy of mathematics owed much to Polya. He then adapted the idea of heuristics as a key to identifying research programmes. He says a research programme is defined by its positive and negative heuristic. The negative heuristic says: Hands off – don’t meddle here. The positive heuristic says: Here is a set of problem areas ranked in order of importance – worry only about questions at the top of the list.

### **Hard cores and protective belts**

The negative heuristic is the ‘hard core’ of a programme, a body of central principles which are never to be challenged. They are regarded as irrefutable. Thus in the Newtonian programme, we have at the core the three laws of dynamics and the law of gravitation. If planets misbehave, a Newtonian will not revise the gravitational law, but try to explain the anomaly by postulating a possibly invisible planet, a planet which, if need be, can be detected only by its perturbations on the solar system.

The positive heuristic is an agenda determining which problems are to be worked on. Lakatos imagines a healthy research programme positively wallowing in a sea of anomalies, but being none the less exuberant. According to him Kuhn’s vision of normal science makes it almost a chance affair which anomalies are made

the object of puzzle-solving activity. Lakatos says on the contrary that there is a ranking of problems. A few are systematically chosen for research. This choice generates a 'protective belt' around the theory, for one attends only to a set of problems ordained in advance. Other seeming refutations are simply ignored. Lakatos uses this to explain, why, *pace* Popper, verification seems so important in science. People choose a few problems to work on, and feel vindicated by a solution; refutations, on the other hand, may be of no interest.

### Progress and degeneration

What makes a research programme good or bad? The good ones are progressive, the bad ones are degenerating. A programme will be a sequence of theories  $T_1, T_2, T_3, \dots$ . Each theory must be at least as consistent with known facts as its predecessor. The sequence is theoretically progressive if each theory in turn predicts some novel facts not foreseen by its predecessors. It is empirically progressive if some of these predictions pan out. A programme is simply *progressive*, if it is both theoretically and empirically progressive. Otherwise it is *degenerating*.

The degenerating programme is one that gradually becomes closed in on itself. Here is an example.<sup>2</sup> One of the famous success stories is that of Pasteur, whose work on microbes enabled him to save the French beer, wine and silk industries that were threatened by various small hostile organisms. Later we began to pasteurize milk. Pasteur also identified the micro-organisms that enabled him to vaccinate against anthrax and rabies. There evolved a research programme whose hard core held that every hitherto organic harm not explicable in terms of parasites or injured organs was to be explained in terms of micro-organisms. When many diseases failed to be caused by bacteria, the positive heuristic directed a search for something smaller, the virus. This progressive research programme had degenerating subprogrammes. Such was the enthusiasm for microbes that what we now call deficiency diseases *had* to be caused by bugs. In the early years of this century the leading professor of tropical disease, Patrick Manson, insisted that beriberi and some other deficiency diseases are caused by bacterial contagion. An

<sup>2</sup> K. Codell Carter, 'The germ theory, Beri-beri, and the deficiency theory of disease', *Medical History* 21 (1977), pp. 119–36.

epidemic of beriberi was in fact caused by the new processes of steam-polishing rice, processes imported from Europe which killed off millions of Chinese and Indonesians whose staple food was rice. Vitamin B<sub>1</sub> in the hull of the rice was destroyed by polishing. Thanks largely to dietary experiments in the Japanese Navy, people gradually came to realize that not presence of microbes, but absence of something in polished rice was the problem. When all else failed, Manson insisted that there are bacteria that live and die in the polished but not in the unpolished rice, and they are the cause of the new scourge. This move was theoretically degenerating because each modification in Manson's theory came only after some novel observations, not before, and it was empirically degenerating because no polished-rice-organisms are to be found.

### Hindsight

We cannot tell whether a research programme is progressive until after the fact. Consider the splendid problem shift of the Pasteur programme, in which viruses replace bacteria as the roots of most evils that persist in the developed world. In the 1960s arose the speculation that cancers – carcinomas and lymphomas – are caused by viruses. A few extremely rare successes have been recorded. For example, a strange and horrible tropical lymphoma (Burkitt's lymphoma) that causes grotesque swellings in the limbs of people who live above 5000 feet near the equator, has almost certainly been traced to a virus. But what of the general cancer-virus programme? Lakatos tells us, 'We must take budding programmes leniently; programmes may take decades before they get off the ground and become empirically progressive' (I, p. 6). Very well, but even if they have been progressive in the past – what more so than Pasteur's programme – that tells us exactly nothing except 'Be open-minded, and embark on numerous different kinds of research if you are stymied.' It does not merely fail to help choose new programmes with no track record. We know of few more progressive programmes than that of Pasteur, even if some of its failures have been hived off, for example into the theory of deficiency diseases. Is the attempt to find cancer viruses progressive or degenerating? We shall know only later. If we were trying to decide what proportion of the 'War on Cancer' to spend on molecular biology and what on viruses (not

necessarily mutually exclusive, of course) Lakatos could tell us nothing.

### **Objectivity and subjectivism**

What then was Lakatos doing? My guess is indicated by the title of this chapter. He wanted to find a substitute for the idea of truth. This is a little like Putnam's subsequent suggestion, that the correspondence theory of truth is mistaken, and truth is whatever it is rational to believe. But Lakatos is more radical than Putnam. Lakatos is no born-again pragmatist. He is down on truth, not just a particular theory of truth. He does not want a replacement for the correspondence theory, but a replacement for truth itself. Putnam has to fight himself away from a correspondence theory of truth because, in English-speaking philosophy, correspondence theories, despite the pragmatist assault of long ago, are still popular. Lakatos, growing up in an Hegelian tradition, almost never gives the correspondence theory a thought. However, like Peirce, he values an objectivity in science that plays little role in Hegelian discourse. Putnam honours this value by hoping, like Peirce, that there is a scientific method upon which we shall come to agree, and which in turn will lead us all to agreement, to rational, warranted, belief. Putnam is a simple Peircian, even if he is less confident than Peirce that we are already on the final track. Rationality looks forward. Lakatos went one step further. There is no forward-looking rationality, but we can comprehend the objectivity of our present beliefs by reconstructing the way we got here. Where do we start? With the growth of knowledge itself.

### **The growth of knowledge**

The one fixed point in Lakatos's endeavour is the simple fact that knowledge does grow. Upon this he tries to build his philosophy without representation, starting from the fact that one can see that knowledge grows whatever we think about 'truth' or 'reality'. Three related aspects of this fact are to be noticed.

First, one can see by direct inspection that knowledge has grown. This is not a lesson to be taught by general philosophy or history but by detailed reading of specific sequences of texts. There is no doubt that more is known now than was grasped by past genius. To take an

example of his own, it is manifest that after the work of Rutherford and Soddy and the discovery of isotopes, vastly more was known about atomic weights than had been dreamt of by a century of toilers after Prout had hypothesized in 1815 that hydrogen is the stuff of the universe, and that atomic weights are integral multiples of that of hydrogen. I state this to remind ourselves that Lakatos starts from a profound but elementary point. The point is not that there is knowledge but that there is growth; we know more about atomic weights than we once did, even if future times plunge us into quite new, expanded, reconceptualizations of those domains.

Secondly, there is no *arguing* that some historical events do exhibit the growth of knowledge. What is needed is an *analysis* that will say in what this growth consists, and tell us what is the growth that we call science and what is not. Perhaps there are fools who think that the discovery of isotopes is no growth in real knowledge. Lakatos's attitude is that they are not to be contested – they are likely idle and have never read the texts or engaged in the experimental results of such growth. We should not argue with such ignoramuses. When they have learned how to use isotopes or simply read the texts, they will find out that knowledge does grow.

This thought leads to the third point. The growth of scientific knowledge, given an intelligent analysis, might provide a demarcation between rational activity and irrationalism. Although Lakatos expressed matters in that way, it is not the right form of words to use. Nothing has grown more consistently and persistently over the years than the commentaries on the Talmud. Is that a rational activity? We see at once how hollow is that word 'rational' if used for positive evaluation. The commentaries are the most reasoned great bodies of texts that we know, vastly more reasoned than the scientific literature. Philosophers often pose the tedious question of why twentieth-century Western astrology, such as it is, is no science. That is not where the thorny issues of demarcation lie. Popper took on more serious game in challenging the right of psychoanalysis or Marxist historiography to the claim of 'science'. The machinery of research programmes, hard cores and protective belts, progress and degeneration, must, if it is of worth, effect a distinction not between the rational and reasoning, and the irrational and unreasoning, but between those reasonings which lead to what Popper and Lakatos call objective knowledge and those

which pursue different aims and have different intellectual trajectories.

### Appraising scientific theories

Hence Lakatos provides no forward-looking assessments of present competing scientific theories. He can at best look back and say why, on his criteria, this research programme was progressive, why another was not. As for the future, there are few pointers to be derived from his ‘methodology’. He says that we should be modest in our hopes for our own projects because rival programmes may turn out to have the last word. There is a place for pig-headedness when one’s programme is going through a bad patch. The mottos are to be proliferation of theories, leniency in evaluation, and honest ‘score-keeping’ to see which programme is producing results and meeting new challenges. These are not so much real methodology as a list of the supposed values of a science allegedly free of ideology.

If Lakatos were in the business of theory appraisal, then I should have to agree with his most colourful critic, Paul Feyerabend. The main thrust of the often perceptive assaults on Lakatos to be found in Chapter 17 of *Against Method* is that Lakatos’s ‘methodology’ is not a good device for advising on current scientific work. I agree, but suppose that was never the point of the analysis which, I claim, has a more radical object. Lakatos had a sharp tongue, strong opinions and little diffidence. He made many entertaining observations about this or that current research project, but these acerbic asides were incidental to and independent of the philosophy I attribute to him.

Is it a defect in Lakatos’s methodology that it is only retroactive? I think not. There are no significant general laws about what, in a current bit of research, bodes well for the future. There are only truisms. A group of workers who have just had a good idea often spends at least a few more years fruitfully applying it. Such groups properly get lots of money from corporations, governments, and foundations. There are other mild sociological inductions, for example that when a group is increasingly concerned to defend itself against criticism, and won’t dare go out on a new limb, then it seldom produces interesting new research. Perhaps the chief practical problem is quite ignored by philosophers of rationality. How do you stop funding a program you have supported for five or

fifteen years – a program to which many young people have dedicated their careers – and which is finding out very little? That real-life crisis has little to do with philosophy.

There is a current vogue among some philosophers of science, that Lakatos might have called ‘the new justificationism’. It produces whole books trying to show that a system of appraising theories can be built up out of rules of thumb. It is even suggested that governments should fund work in the philosophy of science, in order to learn how to fund projects in real science. We should not confuse such creatures of bureaucracy with Lakatos’s attempt to understand the content of objective judgement.

### **Internal and external history**

Lakatos’s tool for understanding objectivity was something he called history. Historians of science, even those given to considerable flights of speculative imagination, find in Lakatos only ‘an historical parody that makes one’s hair stand on end’. That is Gerald Holton’s characterization in *The Scientific Imagination* (p. 106); many colleagues agree.

Lakatos begins with an ‘unorthodox, new demarcation between “internal” and “external” history’ (I, p. 102), but is not very clear what is going on. External history commonly deals in economic, social and technological factors that are not directly involved in the content of a science, but which are deemed to influence or explain some events in the history of knowledge. External history might include an event like the first Soviet satellite to orbit the earth – Sputnik – which was followed by the instant investment of vast sums of American money in science education. Internal history is usually the history of ideas germane to the science, and attends to the motivations of research workers, their patterns of communication and lines of intellectual filiation – who learned what from whom.

Lakatos’s internal history is to be one extreme on this spectrum. It is to exclude anything in the subjective or personal domain. What people believed is irrelevant: it is to be a history of some sort of abstraction. It is, in short, to be a history of Hegelian alienated knowledge, the history of anonymous and autonomous research programmes.

This idea about the growth of knowledge into something

objective and non-human was foreshadowed in his first major philosophical work, *Proofs and Refutations*. On p. 146 of this wonderful dialogue on the nature of mathematics, we find:

Mathematical activity is human activity. Certain aspects of this activity – as of any human activity – can be studied by psychology, others by history. Heuristic is not primarily interested in these aspects. But mathematical activity produces mathematics. Mathematics, this product of human activity, ‘alienates itself’ from the human activity which has been producing it. It becomes a living growing organism that acquires a certain autonomy from the activity which has produced it.

Here then are the seeds of Lakatos’s redefinition of ‘internal history’, the doctrine underlying his ‘rational reconstructions’. One of the lessons of *Proofs and Refutations* is that mathematics might be both the product of human activity and autonomous, with its own internal characterization of objectivity which can be analysed in terms of how mathematical knowledge has grown. Popper has suggested that such objective knowledge could be a ‘third world’ of reality, and Lakatos toyed with this idea.

Popper’s metaphor of a third world is puzzling. In Lakatos’s definition, ‘the “first world” is the physical world; the “second world” is the world of consciousness, of mental states and, in particular, of beliefs; the “third world” is the Platonic world of objective spirit, the world of ideas’ (II, p. 108). I myself prefer those texts of Popper’s where he says that the third world is a world of books and journals stored in libraries, of diagrams, tables and computer memories. Those extra-human things, uttered sentences, are more real than any talk of Plato would suggest.

Stated as a list of three worlds we have a mystery. Stated as a sequence of three emerging kinds of entity with corresponding laws it is less baffling. First there was the physical world. Then when sentient and reflective beings emerged out of that physical world there was also a second world whose descriptions could not be in any general way reduced to physical world descriptions. Popper’s third world is more conjectural. His idea is that there is a domain of human knowledge (sentences, print-outs, tapes) which is subject to its own descriptions and laws and which cannot be reduced to second-world events (type by type) any more than second-world events can be reduced to first-world ones. Lakatos persists in the metaphorical expression of this idea: ‘The *products* of human

knowledge; propositions, theories, systems of theories, problems, problemshifts, research programmes live and grow in the “third world”; the producers of knowledge live in the first and second worlds’ (II, p. 108). One need not be so metaphorical. It is a difficult but straightforward question whether there is an extensive and coherent body of description of ‘alienated’ and autonomous human knowledge that cannot be reduced to histories and psychologies of subjective beliefs. A substantiated version of a ‘third world’ theory can provide just the domain for the content of mathematics. It admits that mathematics is a product of the human mind, and yet is also autonomous of anything peculiar to psychology. An extension of this theme is provided by Lakatos’s conception of ‘unpsychological’ internal history.

Internal history will be a rational construction of what actually happened, one which displays why what happened in many of the best incidents of the history of science are worthy of designations such as ‘rational’ and ‘objective’. Lakatos had a fine sounding maxim, a parody of one of Kant’s noble turns of phrase: ‘Philosophy of science without history of science is empty; history of science without philosophy of science is blind.’ That sounds good, but Kant had been speaking of something else. All we need to say about rather unreflective history of science was said straightforwardly by Kant himself in his lectures on *Logic*: ‘Mere polyhistory is a cyclopean erudition that lacks one eye, the eye of philosophy.’ Lakatos wants to rewrite the history of science so that the ‘best’ incidents in the history of science are cases of progressive research programmes.

### Rational reconstruction

Lakatos has a problem, to characterize the growth of knowledge internally by analysing examples of growth. There is a conjecture, that the unit of growth is the research programme (defined by hard core, protective belt, heuristic) and that research programmes are progressive or degenerating and, finally, that knowledge grows by the triumph of progressive programmes over degenerating ones. To test this supposition we select an example which must *prima facie* illustrate something that scientists have found out. Hence the example should be currently admired by scientists, or people who think about the appropriate branch of knowledge, not because we

kow-tow to orthodoxy, but because workers in a given domain tend to have a better sense of what matters than laymen. Feyerabend calls this attitude elitism. Is it? The next Lakatosian injunction is for all of us to read all the texts we can lay hands on, covering a complete epoch spanned by the research programme, and the entire array of practitioners. Yes, that is elitism because few can afford the time to read. But it has an anti-elite intellectual premise (as opposed to an elite economic premise) that if texts are available, anyone is able to read them.

Within what we read we must select the class of sentences that express what the workers of the day were trying to find out, and how they were trying to find it out. Discard what people felt about it, the moments of creative hype, even their motivation or their role models. Having settled on such an 'internal' part of the data we can now attempt to organize the result into a story of Lakatosian research programmes.

As in most inquiries, an immediate fit between conjecture and articulated data is not to be expected. Three kinds of revision may improve the mesh between conjecture and selected data. First, we may fiddle with the data analysis, secondly, we may revise the conjecture, and thirdly, we may conclude that our chosen case study does not, after all, exemplify the growth of knowledge. I shall discuss these three kinds of revision in order.

By improving the analysis of data I do not mean lying. Lakatos made a couple of silly remarks in his 'falsification' paper, where he asserts something as historical fact in the text, but retracts it in the footnotes, urging that we take his text with tons of salt (I, p. 55). The historical reader is properly irritated by having his nose tweaked in this way. No point was being served. Lakatos's little joke was not made in the course of a rational reconstruction despite the fact that he said it was. Just as in any other inquiry, there is nothing wrong with trying to re-analyse the data. That does not mean lying. It may mean simply reconsidering or selecting and arranging the facts, or it may be a case of imposing a new research programme on the known historical facts.

If the data and the Lakatosian conjecture cannot be reconciled, two options remain. First, the case history may itself be regarded as something other than the growth of knowledge. Such a gambit could easily become monster-barring, but that is where the

constraint of external history enters. Lakatos can always say that a particular incident in the history of science fails to fit his model because it is ‘irrational’, but he imposes on himself the demand that one should allow this only if one can say what the irrational element is. External elements may be political pressure, corrupted values or, perhaps, sheer stupidity. Lakatos’s histories are normative in that he can conclude that a given chunk of research ‘ought not to have’ gone the way it did, and that it went that way through the interference of external factors not germane to the programme. In concluding that a chosen case was not ‘rational’ it is permissible to go against current scientific wisdom. But although in principle Lakatos can countenance this, he is properly moved by respect for the implicit appraisals of working scientists. I cannot see Lakatos willingly conceding that Einstein, Bohr, Lavoisier or even Copernicus was participating in an irrational programme. ‘Too much of the actual history of science’ would then become ‘irrational’ (I, p. 172). We have no standards to appeal to, in Lakatos’s programme, other than the history of knowledge as it stands. To declare it to be globally irrational is to abandon rationality. We see why Feyerabend spoke of Lakatos’s elitism. Rationality will simply be defined by what a present community calls good, and nothing shall counterbalance the extraterrestrial weight of an Einstein.

Lakatos then defines objectivity and rationality in terms of progressive research programmes, and allows an incident in the history of science to be objective and rational if its internal history can be written as a sequence of progressive problem shifts.

### Cataclysms in reasoning

Peirce defined truth as what is reached by an ideal end to scientific inquiry. He thought that it is the task of methodology to characterize the principles of inquiry. There is an obvious problem: what if inquiry should not converge on anything? Peirce, who was as familiar in his day with talk of scientific revolutions as we are in ours, was determined that ‘cataclysms’ in knowledge (as he called them) have not been replaced by others, but this is all part of the self-correcting character of inquiry. Lakatos has an attitude similar to Peirce’s. He was determined to refute the doctrine that he attributed to Kuhn, that knowledge changes by irrational ‘conversions’ from one paradigm to another.

As I said in the Introduction, I do not think that a correct reading of Kuhn gives quite the apocalyptic air of cultural relativism that Lakatos found there. But there is a really deep worry underlying Lakatos's antipathy to Kuhn's work, and it must not be glossed over. It is connected with an important side remark of Feyerabend's, that Lakatos's accounts of scientific rationality at best fit the major achievements 'of the last couple of hundred years'.

A body of knowledge may break with the past in two distinguishable ways. By now we are all familiar with the possibility that new theories may completely replace the conceptual organization of their predecessors. Lakatos's story of progressive and degenerating programmes is a good stab at deciding when such replacements are 'rational'. But all of Lakatos's reasoning takes for granted what we may call the hypothetico-deductive model of reasoning. For all his revisions of Popper, he takes for granted that conjectures are made and tested against some problems chosen by the protective belt. A much more radical break in knowledge occurs when an entirely new style of reasoning surfaces. The force of Feyerabend's gibe about 'the last couple of hundred years' is that Lakatos's analysis is relevant not to timeless knowledge and timeless reason, but to a particular kind of knowledge produced by a particular style of reasoning. That knowledge and that style have specific beginnings. So the Peircian fear of cataclysm becomes: Might there not be further styles of reasoning which will produce yet a new kind of knowledge? Is not Lakatos's surrogate for truth a local and recent phenomenon?

I am stating a worry, not an argument. Feyerabend makes sensational but implausible claims about different modes of reasoning and even seeing in the archaic past. In a more pedestrian way my own book, *The Emergence of Probability* (1975), contends that part of our present conception of inductive evidence came into being only at the end of the Renaissance. In his book, *Styles of Scientific Thinking in the European Tradition* (1983), the historian A.C. Crombie, from whom I take the word 'style', writes of six distinguishable styles. I have elaborated Crombie's idea elsewhere. Now it does not follow that the emergence of a new style is a cataclysm. Indeed we may add style to style, with a cumulative body of conceptual tools. That is what Crombie teaches. Clearly both

Putnam and Laudan expect this to happen. But these are matters which are only recently broached, and are utterly ill-understood. They should make us chary of an account of reality and objectivity which starts from the growth of knowledge, when the kind of growth described turns out to concern chiefly a particular knowledge achieved by a particular style of reasoning.

To make matters worse, I suspect that a style of reasoning may determine the very nature of the knowledge that it produces. The postulational method of the Greeks gave a geometry which long served as the philosopher's model of knowledge. Lakatos inveighs against the domination of the Euclidean mode. What future Lakatos will inveigh against the hypothetico-deductive mode and the theory of research programmes to which it has given birth? One of the most specific features of this mode is the postulation of theoretical entities which occur in high-level laws, and yet which have experimental consequences. This feature of successful science becomes endemic only at the end of the eighteenth century. Is it even possible that the questions of objectivity, asked for our times by Kant, are precisely the questions posed by this new knowledge? If so, then it is entirely fitting that Lakatos should try to answer those questions in terms of the knowledge of the past two centuries. But it would be wrong to suppose that we can get from this specific kind of growth to a theory of truth and reality. To take seriously the title of a book that Lakatos proposed, but never lived to write, 'The changing logic of scientific discovery' is to take seriously the possibility that Lakatos has, like the Greeks, made the eternal verities depend on a mere episode in the history of human knowledge.

There remains an optimistic version of this worry. Lakatos was trying to characterize certain objective values of Western science without an appeal to copy theories of truth. Maybe those objective values are recent enough that his limitation to the past two or three centuries is exactly right. We are left with no external way to evaluate our own tradition, but why should we want that?

**BREAK**

## BREAK

# Reals and representations

Incommensurability, transcendental nominalism, surrogates for truth, and styles of reasoning are the jargon of philosophers. They arise from contemplating the connection between theory and the world. All lead to an idealist cul-de-sac. None invites a healthy sense of reality. Indeed much recent philosophy of science parallels seventeenth-century epistemology. By attending only to knowledge as representation of nature, we wonder how we can ever escape from representations and hook-up with the world. That way lies an idealism of which Berkeley is the spokesman. In our century John Dewey has spoken sardonically of a spectator theory of knowledge that has obsessed Western philosophy. If we are mere spectators at the theatre of life, how shall we ever know, on grounds internal to the passing show, what is mere representation by the actors, and what is the real thing? If there were a sharp distinction between theory and observation, then perhaps we could count on what is observed as real, while theories, which merely represent, are ideal. But when philosophers begin to teach that all observation is loaded with theory, we seem completely locked into representation, and hence into some version of idealism.

Pity poor Hilary Putnam, for example. Once the most realist of philosophers, he tried to get out of representation by tacking ‘reference’ on at the end of the list of elements that constitute the meaning of a word. It was as if some mighty referential sky-hook could enable our language to embed within it a bit of the very stuff to which it refers. Yet Putnam could not rest there, and ended up as an ‘internal realist’ only, beset by transcendental doubts, and given to some kind of idealism or nominalism.

I agree with Dewey. I follow him in rejecting the false dichotomy between acting and thinking from which such idealism arises. Perhaps all the philosophies of science that I have described are part of a larger spectator theory of knowledge. Yet I do not think that the idea of knowledge as representation of the world is in itself the

source of that evil. The harm comes from a single-minded obsession with representation and thinking and theory, at the expense of intervention and action and experiment. That is why in the next part of this book I study experimental science, and find in it the sure basis of an uncontentious realism. But before abandoning theory for experiment, let us think a little more about the very notions of representation and reality.

### The origin of ideas

What are the origins of these two ideas, *representation* and *reality*? Locke might have asked that question as part of a psychological inquiry, seeking to show how the human mind forms, frames, or constitutes its ideas. There is a legitimate science that studies the maturation of human intellectual abilities, but philosophers often play a different game when they examine the origin of ideas. They tell fables in order to teach philosophical lessons. Locke himself was fashioning a parable when he pretended to practice the natural history of the mind. Our modern psychologies have learned how to trick themselves out in more of the paraphernalia of empirical research, but they are less distant from fantastical Locke than they assume. Let us, as philosophers, welcome fantasies. There may be more truth in the average *a priori* fantasy about the human mind than in the supposedly disinterested observations and mathematical model-building of cognitive science.

### Philosophical anthropology

Imagine a philosophical text of about 1850: ‘Reality is as much an anthropomorphic creation as God Himself.’ This is not to be uttered in a solemn tone of voice that says, ‘God is dead and so is reality.’ It is to be a more specific and practical claim: *Reality is just a byproduct of an anthropological fact*. More modestly, the concept of reality is a byproduct of a fact about human beings.

By anthropology I do not mean ethnography or ethnology, the studies practised in present-day departments of anthropology, and which involve lots of field work. By anthropology I mean the bogus nineteenth-century science of ‘Man’. Kant once had three philosophical questions. What must be the case? What should we do? For what may we hope? Late in life he added a fourth question: *What is Man?* With this he inaugurated (*philosophische*) *Anthropologie* and

even wrote a book called *Anthropology*. Realism is not to be considered part of pure reason, nor judgement, nor the metaphysics of morals, nor even the metaphysics of natural science. If we are to give it classification according to the titles of Kant's great books, realism shall be studied as part of *Anthropologie* itself.

A Pure Science of Human Beings is a bit risky. When Aristotle proposed that Man is an animal that lives in cities, so that the *polis* is a part of Man's nature to which He strives, his pupil Alexander refuted him by re-inventing the Empire. We have been told that Man is a tool-maker, or a creature that has a thumb, or that stands erect. We have been told that these fortuitous features are noticed only by attending to half of the species wrongly called Man, and that tools, thumbs and erectness are scarcely what define the race. It is seldom clear what the grounds might be for any such statements, pro or con. Suppose one person defines humans as rational, and another person defines them as the makers of tools. Why on earth should we suppose that being a rational animal is co-extensive with making tools?

Speculations about the essential nature of humanity license more of the same. Philosophers since Descartes have been attracted by the conjecture that humans are speakers. It has been urged that rationality, of its very nature, demands language, so humans as rational animals, and humans as speakers are indeed co-extensive. That is a satisfactory main theorem for a subject as feeble as fanciful anthropology. Yet despite the manifest profundity of this conclusion, a conclusion that has fuelled mighty books, I propose another fancy. *Human beings are representers*. Not *homo faber*, I say, but *homo depictor*. People make representations.

### **Limiting the metaphor**

People make likenesses. They paint pictures, imitate the clucking of hens, mould clay, carve statues, and hammer brass. Those are the sorts of representations that begin to characterize human beings.

The word 'representation' has quite a philosophical past. It has been used to translate Kant's word *Vorstellung*, a placing before the mind, a word which includes images as well as more abstract thoughts. Kant needed a word to replace the 'idea' of the French and English empiricists. That is exactly what I do *not* mean by representation. Everything I call a representation is public. You

cannot touch a Lockeian idea, but only the museum guard can stop you touching some of the first representations made by our predecessors. I do not mean that all representations can be touched, but all are public. According to Kant, a judgement is a representation of a representation, a putting before the mind of a putting before the mind, doubly private. That is doubly not what I call a representation. But for me, some public verbal events can be representations. I think not of simple declarative sentences, which are surely not representations, but of complicated speculations which attempt to represent our world.

When I speak of representations I first of all mean physical objects: figurines, statues, pictures, engravings, objects that are themselves to be examined, regarded. We find these as far back as we find anything human. Occasionally some fortuitous event preserves even fragments of wood or straw that would otherwise have rotted. Representations are external and public, be they the simplest sketch on a wall, or, when I stretch the word 'representation', the most sophisticated theory about electromagnetic, strong, weak, or gravitational forces.

The ancient representations that are preserved are usually visual and tactile, but I do not mean to exclude anything publicly accessible to the other senses. Bird whistles and wind machines may make likenesses too, even though we usually call the sounds that they emit imitations. I claim that if a species as smart as human beings had been irrevocably blind, it would have got on fine with auditory and tactile representations, for to represent is part of our very nature. Since we have eyes, most of the first representations were visual, but representation is not of its essence visual.

Representations are intended to be more or less public likenesses. I exclude Kant's *Vorstellungen* and Lockeian internal ideas that represent the external world in the mind's eye. I also exclude ordinary public sentences. William James jeered at what he called the copy theory of truth, which bears the more dignified label of correspondence theory of truth. The copy theory says that true propositions are copies of whatever in the world makes them true. Wittgenstein's *Tractatus* has a picture theory of truth, according to which a true sentence is one which correctly pictures the facts. Wittgenstein was wrong. Simple sentences are not pictures, copies, or representations. Doubtless philosophical talk of representation

invites memories of Wittgenstein's *Sätze*. Forget them. The sentence, 'the cat is on the mat', is no representation of reality. As Wittgenstein later taught us, it is a sentence that can be used for all sorts of purposes, none of which is to portray what the world is like. On the other hand, Maxwell's electromagnetic theories were intended to represent the world, to say what it is like. Theories, not individual sentences, are representations.

Some philosophers, realizing that sentences are not representations, conclude that the very idea of a representation is worthless for philosophy. That is a mistake. We can use complicated sentences collectively in order to represent. So much is ordinary English idiom. A lawyer can represent the client, and can also represent that the police collaborated improperly in preparing their reports. A *single* sentence will in general not represent. A representation can be verbal, but a verbal representation will use a good many verbs.

### **Humans as speakers**

The first proposition of my philosophical anthropology is that human beings are depicators. Should the ethnographer tell me of a race that makes no image (not because that is tabu but because no one has thought of representing anything) then I would have to say that those are not people, not *homo depitor*. If we are persuaded that humankind (and not its predecessors) lived in Olduvai gorge three million years ago, and yet we find nothing much except old skulls and footprints, I would rather postulate that the representations made by those African forbears have been erased by sand, rather than that people had not yet begun to represent.

How does my *a priori* paleolithic fantasy mesh with the ancient idea that humans are essentially rational and that rationality is essentially linguistic? Must I claim that depiction needs language or that humanity need not be rational? If language has to be tucked into rationality, I would cheerfully conclude that humans may *become* rational animals. That is, *homo depitor* did not always deserve Aristotle's accolade of rationality, but only earned it as we smartened up and began to talk. Let us imagine, for a moment, pictorial people making likenesses before they learn to talk.

### The beginnings of language

Speculation on the origin of language tends to be unimaginative and condescending. Language, we hear, must have been invented to help with practical matters such as hunting and farming. ‘How useful,’ goes the refrain, ‘to be able to talk. How much more efficient people would have been if they could talk. Speech makes it much more likely that hunters and farmers will survive.’

Scholars who favour such rubbish have evidently never ploughed a field nor stalked game, where silence is the order of the day, not jabber. People out in the fields weeding do not usually talk. They talk only when they rest. In the plains of East Africa the hunter with the best kill rate is the wild dog, yet middle-aged professors short of wind and agreeing never to talk nor signal are much better at catching the beeste and the gazelle than any wild dog. The lion that roars and the dogs that bark will starve to death if enough silent humans are hunting with their bare hands.

Language is not for practical affairs. Jonathan Bennett tells a story about language beginning when one ‘tribesman’ warns another that a coconut is about to fall on the second native’s head.<sup>1</sup> Native One does this first by an overacted mime of bonking on the head, and later on does this by uttering a warning and thereby starting language. I bet that no coconut ever fell on any tribesman’s head except in racist comic strips, so I doubt this fantasy. I prefer a suggestion about language attributed to the Leakey family who excavate Olduvai gorge. The idea is that people invented language out of boredom. Once we had fire, we had nothing to do to pass away the long evenings, so we started telling jokes. This fancy about the origin of language has the great merit of regarding speech as something human. It fixates not on tribesmen in the tropics but on people.

Imagine *homo depictor* beginning to use sounds that we might translate as ‘real’, or, ‘that’s how it is’, said of a clay figurine or a daub on the wall. Let discourse continue as ‘this real, then that real’, or, more idiomatically, ‘if this is how it is, then that is how it is too’. Since people are argumentative, other sounds soon express, ‘no, not that, but this here is real instead’.

<sup>1</sup> J. Bennett, ‘The meaning-nominalist strategy’, *Foundations of Language* 10 (1973), pp. 141–68.

In such a fantasy we do not first come to the names and descriptions, or the sense and reference of which philosophers are so fond. Instead we start with the indexicals, logical constants, and games of seeking and finding. Descriptive language comes later, not as a surrogate for depiction but as other uses for speaking are invented.

Language then starts with ‘this real’, said of a representation. Such a story has to its credit the fact that ‘this real’ is not at all like ‘You Tarzan, Me Jane’, for it stands for a complicated, that is, characteristically human, thought, namely that this wooden carving shows something real about what it represents.

This imagined life is intended as an antidote to the deflating character of the quotation with which I began: Reality is an anthropomorphic creation. Reality may be a human creation, but it is no toy; on the contrary it is the second of human creations. The first peculiarly human invention is representation. Once there is a practice of representing, a second-order concept follows in train. This is the concept of reality, a concept which has content only when there are first-order representations.

It will be protested that reality, or the world, was there before any representation or human language. Of course. But conceptualizing it as reality is secondary. First there is this human thing, the making of representations. Then there was the judging of representations as real or unreal, true or false, faithful or unfaithful. Finally comes the world, not first but second, third or fourth.

In saying that reality is parasitic upon representation, I do not join forces with those who, like Nelson Goodman or Richard Rorty, exclaim, ‘the world well lost!’ The world has an excellent place, even if not a first one. It was found by conceptualizing the real as an attribute of representations.

Is there the slightest empirical evidence for my tale about the origin of language? No. There are only straws in the wind. I say that representing is curiously human. Call it species specific. We need only run up the evolutionary tree to see that there is some truth in this. Drug a baboon and paint its face, then show it a mirror. It notices nothing out of the ordinary. Do the same to a chimpanzee. It is terribly upset, sees there is paint on its face and tries to get it off. People, in turn, like mirrors to study their make-up. Baboons will never draw pictures. The student of language, David Premack, has

taught chimpanzees a sort of language using pictorial representation. *Homo depictor* was better than that, right from the start. We still are.

### Likeness

Representations are first of all likenesses. Saying so flies in the face of philosophical truisms. There is, we all know, no representation without style. Even the most untutored of cultures must have a system of representation if it is to represent at all. So it may be argued that there cannot in the beginning have been simply representation, a creating of likeness. There must have been a style of representing before there was representing.

I need not disagree with this doctrine, so long as it be admitted that styles do not precede representation. They grow with representation as materials are worked, and craftspeople produce artifacts that affect the sensibility of their customers.

A more philosophical conundrum lurks hereabouts. Things are alike, it is said, in some respect or another, and cannot be simply like. There must be some concept used to express that in which likeness consists. Two people have the same walk, or the same bearing, or the same nose, or the same parents or the same character. But two people cannot simply be ‘like’ each other. I agree with this too, but tentatively hold that it does not preclude simple likeness.

I am too brainwashed by philosophy to hold that things *in general* can be simply, or unqualifiedly alike. They must be like or unlike in this or that respect. However a particular kind of thing, namely a human-made representation, can unqualifiedly be like what it is intended to represent. Our generalized notion of likeness is, like our idea of reality, parasitic upon our practices of representation. There may be some initial way in which representations are like what they represent. There is no doubt that some human artifacts of very foreign and very ancient peoples are immediately recognized as likenesses, even when we do not quite know what they are likenesses of. Those pictures, carvings, gold inlay, worked copper, clay faces, mammoth rock carvings, pocket-sized canoes for burial purposes – all the artistic detritus that we find where people once lived – are likenesses. I may not know what they are likenesses of nor what they are for. I ill-understand the systems of representation but I know these are representations all the same. At Delphi I see an archaic

ivory carving of a person, perhaps a god, in what we call formal or lifeless style. I see the gold leggings and cloak in which the ivory was dressed. It is engraved in the most minute and ‘realistic’ detail with scenes of bull and lion. The archaic and the realistic objects in different media are made in what the archaeologists say is the same period. I do not know what either is for. I do know that both are likenesses. I see the archaic bronze charioteer with its compelling human deep-set eyes of semi-precious stone. How, I ask, could craftspeople so keen on what we call lifeless forms work with others who breathed life into their creations? Because different crafts using different media evolve at different rates? Because of a forgotten combination of unknown purposes? Such subtle questions are posed against a background of what we take for granted. We know at least this: these artifacts are representations.

We know likeness and representation even when we cannot answer, likeness to what? Think of the strange little clay figures on which are painted a sketch of garments, but which have, instead of heads, little saucer shaped depressions, perhaps for oil. These finger-high objects litter Mycenae. I doubt that they represent anything in particular. They most remind me of the angel-impressions children make by lying in the snow and waving their arms and legs to and fro to create the image of little wings and skirt. Children make these angels for pleasure. We do not quite know what the citizens of Cnossus did with their figurines. But we know that both are in some way likenesses. The wings and skirt are like wings and skirt, although the angel depicted is like nothing on earth.

Representations are not in general intended to say how it is. They can be portrayals or delights. After our recent obsession with words it is well to reflect on pictures and carvings. Philosophers of language seldom resist the urge to say that the first use of language must be to tell the truth. There should be no such compulsion with pictures. To argue of two bison sketches, ‘If this is how it is, then that is how it is too’, is to do something utterly unusual. Pictures are seldom, and statues are almost never used to say how things are. At the same time there is a core to representation that enables archaeologists millenia later to pick out certain objects in the debris of an ancient site, and to see them as likenesses. Doubtless ‘likeness’ is the wrong word, because the ‘art’ objects will surely include products of the imagination, pretties and uglies made for their own

sake, for the sake of revenge, wealth, understanding, courtship or terror. But within them all there is a notion of representation that harks back to likeness. Likeness stands alone. It is not a relation. It creates the terms in a relation. There is first of all likeness, and then likeness to something or other. First there is representation, and then there is ‘real’. First there is a representation and much later there is a creating of concepts in terms of which we can describe this or that respect in which we have similarity. But likeness can stand on its own without any need of some concepts  $x$ ,  $y$ , or  $z$ , so that one must always think, like in represent of  $z$ , but not of  $x$  or  $y$ . There is no absurdity in thinking that there is a raw and unrefined notion of likeness springing up with the making of representations, and which, as people become more skilful in working with materials, engenders all sorts of different ways of noticing what is like what.

### **Realism no problem**

If reality were just an attribute of representation, and we had not evolved alternative styles of representation, then realism would be a problem neither for philosophers nor for aesthetes. The problem arises because we have alternative systems of representation.

So much is the key to the present philosophical interest in scientific realism. Earlier ‘realistic’ crises commonly had their roots in science. The competition between Ptolemaic and Copernican systems begged for a shoot-out between instrumentalist and realistic cosmologies. Disputes about atomism at the end of the nineteenth century made people wonder if, or in what sense, atoms could be real. Our present debate about scientific realism is fuelled by no corresponding substantive issue in natural science. Where then does it come from? From the suggestions of Kuhn and others that with the growth of knowledge we may, from revolution to revolution, come to inhabit different worlds. New theories are new representations. They represent in different ways and so there are new kinds of reality. So much is simply a consequence of my account of reality as an attribute of representation.

When there were only undifferentiated representations then, in my fantasy story about the origin of language, ‘real’ was unequivocal. But as soon as representations begin to compete, we had to wonder what is real. Anti-realism makes no sense when only one kind of representation is around. Later it becomes possible. In our

time we have seen this as the consequence of Kuhn's *Structure of Scientific Revolutions*. It is, however, quite an old theme in philosophy, best illustrated by the first atomists.

### The Democritean dream

Once representation was with us, reality could not be far behind. It is an obvious notion for a clever species to cultivate. The prehistory of our culture is necessarily given by representations of various sorts, but all that are left us are tiny physical objects, painted pots, moulded cookware, inlay, ivory, wood, tiny burial tools, decorated walls, chipped boulders. *Anthropologie* gets past the phantasies I have constructed only when we have the remembered word, the epics, incantations, chronologies and speculations. The pre-Socratic fragments would be so much mumbo-jumbo were it not for their lineage down to the strategies we now calmly call 'science'. Today's scientific realist attends chiefly to what was once called the inner constitution of things, so I shall pull down only one thread from the pre-Socratic skein, the one that leads down to atomism. Despite Leucippus, and other forgotten predecessors, it is natural to associate this with Democritus, a man only a little older than Socrates. The best sciences of his day were astronomy and geometry. The atomists were bad at the first and weak in the second, but they had an extraordinary hunch. Things, they supposed, have an inner constitution, a constitution that can be thought about, perhaps even uncovered. At least they could guess at this: atoms and the void are all that exist, and what we see and touch and hear are only modifications of this.

Atomism is not essential to this dream of knowledge. What matters is an intelligible organization behind what we take in by the senses. Despite the central role of cosmology, Euclidean proof, medicine and metallurgy in the formation of Western culture, our current problems about scientific realism stem chiefly from the Democritean dream. It aims at a new kind of representation. Yet it still aims at likeness. This stone, I imagine a Democritus saying, is not as it looks to the eye. It is like this – and here he draws dots in the sand or on the tablet, itself thought of as a void. These dots are in continuous and uniform motion, he says, and begins to tell a tale of particles that his descendants turn into odd shapes, springs, forces, fields, all too small or big to be seen or felt or heard except in the

aggregate. But the aggregate, continues Democritus, is none other than this stone, this arm, this earth, this universe.

Familiar philosophical reflections ensue. Scepticism is inevitable, for if the atoms and the void comprise the real, how can we ever know that? As Plato records in the *Gorgias*, this scepticism is three-pronged. All scepticism had had three prongs, since Democritus formulated atomism. There is first of all the doubt that we could check out any particular version of the Democritean dream. If much later Lucretius adds hooks to the atoms, how can we know if he or another speculator is correct? Secondly, there is a fear that this dream is only a dream; there are no atoms, no void, just stones, about which we can, for various purposes, construct certain models whose only touchstone, whose only basis of comparison, whose only reality, is the stone itself. Thirdly, there is the doubt that, although we cannot possibly believe Democritus, the very possibility of his story shows that we cannot credit what we see for sure, and so perhaps we had better not aim at knowledge but at the contemplative ignorance of the tub.

Philosophy is the product of knowledge, no matter how sketchy be the picture of what is known. Scepticism of the sort ‘do I know this is a hand before me’ is called ‘naïve’ when it would be better described as degenerate. The serious scepticism which is associated with it is not, ‘is this a hand rather than a goat or an hallucination?’ but one that originates with the more challenging worry that the hand represented as flesh and bone is false, while the hand represented as atoms and the void is more correct. Scepticism is the product of atomism and other nascent knowledge. So is the philosophical split between appearance and reality. According to the Democritean dream, the atoms must be like the inner constitution of the stone. If ‘real’ is an attribute of depiction, then in asserting his doctrine, Democritus can only say that his picture of particles pictures reality. What then of the depiction of the stone as brown, encrusted, jagged, held in the hand? That, says the atomist, must be appearance.

Unlike its opposite, reality, ‘appearance’ is a thoroughly philosophical concept. It imposes itself on top of the initial two tiers of representation and reality. Much philosophy misorders this triad. Locke thought that we have appearance, then form mental representations and finally, seek reality. On the contrary, we make

public representations, form the concept of reality, and, as systems of representation multiply, we become sceptics and form the idea of mere appearance.

No one calls Democritus a scientific realist: ‘atomism’ and ‘materialism’ are the only ‘isms’ that fit. I take atomism as the natural step from the Stone Age to scientific realism, because it lays out the notion of an ‘inner constitution of things’. With this seventeenth-century phrase, we specify a constitution to be thought about and, hopefully, to be uncovered. But no one did find out about atoms for a long, long time. Democritus transmitted a dream, but no knowledge. Complicated concepts need criteria of application. That is what Democritus lacked. He did not know enough beyond his speculations to have criteria of whether his picture was of reality or not. His first move was to shout ‘real’ and slander the looks of things as mere appearance. Scientific realism or anti-realism do not become possible doctrines until there are criteria for judging whether the inner constitution of things is as represented.

### **The criteria of reality**

Democritus gave us one representation: the world is made up of atoms. Less occult observers give us another. They painted pebbles on the beach, sculpted humans and told tales. In my account, the word ‘real’ first meant just unqualified likeness. But then clever people acquired conjectured likenesses in manifold respects. ‘Real’ no longer was unequivocal. As soon as what we would now call speculative physics had given us alternative pictures of reality, metaphysics was in place. Metaphysics is about criteria of reality. Metaphysics is intended to sort good systems of representation from bad ones. Metaphysics is put in place to sort representations when the only criteria for representations are supposed to be internal to representation itself.

That is the history of old metaphysics and the creation of the problem of realism. The new era of science seemed to save us from all that. Despite some philosophical malcontents like Berkeley, the new science of the seventeenth century could supplant even organized religion and say that it was giving the true representation of the world. Occasionally one got things wrong, but the overthrow of false ideas was only setting us on what was finally the right path. Thus the chemical revolution of Lavoisier was seen as a real

revolution. Lavoisier got some things wrong: I have twice already used the example of his confidence that all acids have oxygen in them. So we sorted that out. In 1816 the new professor of chemistry at Harvard College relates the history of chemistry in an inaugural lecture to the teenagers then enrolled. He notes the revolutions of the recent past, and says we are now on the right road. From now on there will only be corrections. All of that was fine until it began to be realized that *there might be several ways to represent the same facts*.

I do not know when this idea emerged. It is evident in the important posthumous book of 1894, Heinrich Hertz's *Principles of Mechanics*. This is a remarkable work, often said to have led Wittgenstein towards his picture theory of meaning, the core of his 1918 *Tractatus Logico-Philosophicus*. Perhaps this book, or its 1899 English translation, first offers the explicit terminology of a scientific 'image' – now immortalized in the opening sentence of Kuhn's *Structure*, and, following Wilfred Sellars, used as the title of van Fraassen's anti-realist book. Hertz presents 'three images of mechanics' – three different ways to represent the then extant knowledge of the motions of bodies. Here, for perhaps the first time, we have three different systems of representation shown to us. Their merits are weighed, and Hertz favours one.

Hence even within the best understood natural science – mechanics – Hertz needed criteria for choosing between representations. It is not only the artists of the 1870s and 1880s who are giving us new systems of representation called post-impressionism or whatever. Science itself has to produce criteria of what is 'like', of what shall count as the right representation. Whereas art learns to live with alternative modes of representation, here is Hertz valiantly trying to find uniquely the right one for mechanics. None of the traditional values – values still hallowed in 1983 – values of prediction, explanation, simplicity, fertility, and so forth, quite do the job. The trouble is, as Hertz says, that all three ways of representing mechanics do a pretty good job, one better at this, one better at that. What then is the truth about the motions of bodies? Hertz invites the next generation of positivists, including Pierre Duhem, to say that there is no truth of the matter – there are only better or worse systems of representation, and there might well be inconsistent but equally good images of mechanics.

Hertz was published in 1894, and Duhem in 1906. Within that

span of years pretty well the whole of physics was turned upside down. Increasingly, people who knew no physics gossiped that everything is relative to your culture, but once again physicists were sure they were on the only path to truth. They had no doubt about the right representation of reality. We have only one measure of likeness: the hypothetico-deductive method. We propose hypotheses, deduce consequences and see if they are true. Hertz's warnings that there might be several representations of the same phenomena went unheeded. The logical positivists, the hypothetico-deductivists, Karl Popper's falsificationists – they were all deeply moved by the new science of 1905, and were scientific realists to a man, even when their philosophy ought to have made them somewhat anti-realist. Only at a time when physics was rather quiescent would Kuhn cast the whole story in doubt. Science is not hypothetico-deductive. It does have hypotheses, it does make deductions, it does test conjectures, but none of these determine the movement of theory. There are – in the extremes of reading Kuhn – no criteria for saying which representation of reality is the best. Representations get chosen by social pressures. What Hertz had held up as a possibility too scaring to discuss, Kuhn said was brute fact.

### **Anthropological summary**

People represent. That is part of what it is to be a person. In the beginning to represent was to make an object like something around us. Likeness was not problematic. Then different kinds of representation became possible. What was like, which real? Science and its philosophy had this problem from the very beginning, what with Democritus and his atoms. When science became the orthodoxy of the modern world we were able, for a while, to have the fantasy that there is one truth at which we aim. That is the correct representation of the world. But the seeds of alternative representations were there. Hertz laid that out, even before the new wave of revolutionary science which introduced our own century. Kuhn took revolution as the basis for his own implied anti-realism. We should learn this: When there is a final truth of the matter – say, the truth that my typewriter is on the table – then what we say is either true or false. It is not a matter of representation. Wittgenstein's *Tractatus* is exactly wrong. Ordinary simple atomic sentences are not representations

of anything. If Wittgenstein derived his picture account of meaning from Hertz he was wrong to do so. But Hertz was right about representation. In physics and much other interesting conversation we do make representations – pictures in words, if you like. In physics we do this by elaborate systems of modelling, structuring, theorizing, calculating, approximating. These are real, articulated, representations of how the world is. The representations of physics are entirely different from simple, non-representational assertions about the location of my typewriter. There is a truth of the matter about the typewriter. In physics there is no final truth of the matter, only a barrage of more or less instructive representations.

Here I have merely repeated at length one of the aphorisms of the turn-of-the-century Swiss-Italian ascetic, Danilo Domodosala: ‘When there is a final truth of the matter, then what we say is brief, and it is either true or false. It is not a matter of representation. When, as in physics, we provide representations of the world, there is no final truth of the matter.’ Absence of final truth in physics should be the very opposite of disturbing. A correct picture of lively inquiry is given by Hegel, in his preface to the *Phenomenology of Spirit*: ‘The True is thus the Bacchanalian revel in which no member is not drunk; yet because each member collapses as he drops out, the revel is just as much transparent and simple repose.’ Realism and anti-realism scurry about, trying to latch on to something in the nature of representation that will vanquish the other. There is nothing there. That is why I turn from representing to intervening.

### Doing

In a spirit of cheerful irony, let me introduce the experimental part of this book by quoting the most theory-oriented philosopher of recent times, namely Karl Popper:

I suppose that the most central usage of the term ‘real’ is its use to characterize material things of ordinary size – things which a baby can handle and (preferably) put into his mouth. From this, the usage of the term ‘real’ is extended, first, to bigger things – things which are too big for us to handle, like railway trains, houses, mountains, the earth and the stars, and also to smaller things – things like dust particles or mites. It is further extended, of course, to liquids and then also to air, to gases and to molecules and atoms.

What is the principle behind the extension? It is, I suggest, that the entities which we conjecture to be real should be able to exert a causal effect upon the *prima facie* real things; that is, upon material things of an ordinary size: that we can explain changes in the ordinary material world of things by the causal effects of entities conjectured to be real.<sup>2</sup>

That is Karl Popper's characterization of our usage of the word 'real'. Note the traditional Lockean fantasy beginnings. 'Real' is a concept we get from what we, as infants, could put in our mouths. That is a charming picture, not free from nuance. Its absurdity equals that of my own preposterous story of reals and representations. Yet Popper points in the right direction. Reality has to do with causation and our notions of reality are formed from our abilities to change the world.

Maybe there are two quite distinct mythical origins of the idea of 'reality'. One is the reality of representation, the other, the idea of what affects us and what we can affect. Scientific realism is commonly discussed under the heading of representation. Let us now discuss it under the heading of intervention. My conclusion is obvious, even trifling. We shall count as real what we can use to intervene in the world to affect something else, or what the world can use to affect us. Reality as intervention does not even begin to mesh with reality as representation until modern science. Natural science since the seventeenth century has been the adventure of the interlocking of representing and intervening. It is time that philosophy caught up to three centuries of our own past.

<sup>2</sup> Karl Popper and John Eccles, *The Self and its Brain*, Berlin, New York and London, 1977, p. 9.