Critical Realism

An Introduction to
Roy Bhaskar's Philosophy

◆

A N D R E W   C O L L I E R

V
VERSO

In memory of
Annie Pritchard

Contents

Acknowledgements

Abbreviations

Preface

Roy Bhaskar's books are referred to by their initials, thus:

PIF — *Philosophy and the Idea of Freedom*
PN — *The Possibility of Naturalism*
RR — *Reclaiming Reality*
RTS — *A Realist Theory of Science*
SRHE — *Scientific Realism and Human Emancipation*

Full publication details of all works cited in the text are given in the Bibliography.

The term 'critical realism' has come to be applied to the work of a number of philosophers of whom Roy Bhaskar is undoubtedly the most original and influential. His work has long seemed to me the most exciting development in Anglophone philosophy in this half-century. That is a big claim for a body of work which initially had the modest aim of solving some problems in the philosophy of science — a specialized area of philosophy, not even regarded by most philosophers as a central one. Certainly, *A Realist Theory of Science* deserves its place as a landmark in that area, alongside the works of Popper, Lakatos, Kuhn, Harré and others. But there is more to it than that.

Bhaskar's work offers us the possibility of a new beginning. This is so, in the first place, because it avoids the alternatives of irrationalism and a positivistic conception of rationality, which dilemma has beset modern philosophy. On the one hand, it is committed to unfettered reasoning, to a belief that science can give us real insights into the nature of things, and to an interest in the potential of reason and science for human emancipation. In this sense critical realism is an heir to the Enlightenment: on the first page of *Reclaiming Reality* Bhaskar quotes Kant's motto of the Enlightenment *'Sapere aude!* Have courage to use your own reason', and aspires to

> the dawning of a new enlightenment, a socialist enlightenment which will stand to some future order of things, as the eighteenth-century bourgeois enlightenment stood to the American Declaration of Independence, the French Revolution and the overthrow of colonial slavery for which it helped to prepare the cultural ground.

On the other hand, Bhaskar avoids the 'foundationalism' of most of the thought stemming from the Enlightenment, the belief that reason and/or sense-experience could provide out of

their own resources, abstracted from any historical and social context, foundations for the edifice of knowledge — and indubitable foundations at that. Hence the accusation against the Enlightenment that it lacked historical self-awareness does not apply here. He also avoids both the reduction of rationality to the mathematical calculation of optimal means to extraneous ends, as in 'economic rationality', and the equation of objectivity with value-neutrality. In sailing thus between Scylla and Charybdis, irrationalism and narrowly calculative rationalism, he resumes the great dialectical tradition in modern philosophy, the tradition of Hegel and Marx. His most recent book, *Dialectic: The Pulse of Freedom*, makes this affiliation explicit.

Bhaskar's premisses are generally found in the work of the sciences, but his conclusions belong for the most part to *ontology*, the study of being, its varieties and their articulation. His theory of knowledge is a corollary of this, and so too are the implications of his work for ethics and politics.

Like Spinoza, he writes in the indicative yet this cognitive inquiry is powered by the concern for human emancipation. As in Spinoza's case too, this expresses the view that 'the truth shall make you free', though for Bhaskar certainly, as for Spinoza arguably, this does not mean that knowledge *by itself* is sufficient for emancipation.

Bhaskar's work has been widely influential in the human and the borderline natural/human sciences. To my knowledge, workers in the fields of sociology, economics, psychoanalysis, linguistics, history, geography, biology, ecology and feminist theory have put critical realist ideas to good use. But the difficulty of some of Roy Bhaskar's texts is an obstacle to this process. The aim of this book is to make critical realist ideas more accessible to those without a degree in philosophy — though some difficulty must remain, for a technical vocabulary is unavoidable.

The difficulty of Bhaskar's writing is very varied. Some of the essays in *Reclaiming Reality* present no problems at all; much of *A Realist Theory of Science* and *The Possibility of Naturalism*, while they do contain technical terms, some of them new coinages, is written in a lucid, if rather condensed, philosophical prose. The tendency to condense complex thought into brief formulae gets

more pronounced in *Scientific Realism and Human Emancipation*, where it is combined with a large crop of unfamiliar expressions, acronyms and semi-formalized arguments (not to speak of typographical errors and sometimes obscure syntax). As a result, that book has to be read at about a quarter the speed of an average philosophical text. It rewards the effort; not only does it extend and deepen many of the findings of the earlier books, it is also the *only* place where there is a full account of 'explanatory critiques' — one of Bhaskar's most original and far-reaching notions. It really needs not just an introductory text such as the present one but a paragraph-by-paragraph commentary. In general I have not attempted to provide one, though I pay fairly close attention to the passages on explanatory critiques.

I have divided this book into two parts: on *transcendental realism* (which refers to the general ontology which Bhaskar derives from his analysis of scientific practices), and on *critical naturalism* (which refers to his development of the possible implications of transcendental realism for the human sciences). The term 'critical realism' arose out of these two phrases by elision. It may be a curious thing to say in a preface to a book called *Critical Realism*, but I have no great liking for the phrase, partly because of this origin, partly because 'critical' is something of a term of approval in philosophical contexts (as contrasting with 'naïve' or 'dogmatic'), and hence should not figure in the title of a philosophical position, on the principle 'let not him that girdeth on his harness boast himself as he that putteth it off'; and partly because the title 'critical realism' has already been used by Lukács for his version of Marxist aesthetics. However, the term is now used by most of the people working on these ideas, and Roy Bhaskar, though not the author of the phrase, has accepted it retrospectively, so it would be pedantic not to follow suit. And *critical realism*, like *transcendental realism*, does suggest a realist inversion of Kant's philosophy, which (provided one takes account of the structural transformation which such an inversion must involve) is not a bad characterization of what critical realism is.[1]

I have aimed to introduce and expound Bhaskar's ideas (and a few uses made of them by others). Occasionally, this could not be done without including some criticism, but for the most part I have segregated my own criticisms from the main text. In the

case of transcendental realism, this is made easy by the fact that I have very few disagreements with Roy Bhaskar. In the case of critical naturalism I have more, and the final chapter is concerned with what I see as difficulties or unresolved problems in that theory.

Notes

1. As Bhaskar puts it:

   I had called my general philosophy of science 'transcendental realism' and my special philosophy of the human sciences 'critical naturalism'. Gradually people started to elide the two and refer to the hybrid as 'critical realism'. It struck me that there were good reasons not to demur at the mongrel. For a start, Kant had styled his transcendental idealism the 'critical philosophy'. Transcendental realism had as much right to the title of critical realism. (RR, p. 190)

PART I

Transcendental Realism

Why Realism? Why Transcendental?

One afternoon some years ago I was sitting in a café in Bangor High Street drinking a cup of tea and reading David-Hillel Ruben's book *Marxism and Materialism*. It was pub closing time, and a man came and sat opposite me, obviously the worse for drink, and started trying to read my book upside down. He made out the word 'materialism', and asked me if it meant trying to get more and more money. I explained that in the present context it meant the idea that the material world really does exist, independently of our thought about it. Unsurprisingly, his response was outrage that something so obvious had to be said. I have no doubt that nearly all those fifty million or so of our fellow citizens who are not arts or social science graduates would agree with him.

Don't worry, I am not going to argue for realism along the lines of 'ten million Sun readers can't be wrong' — or the well-known bit of graffiti that parodies such statements. The point is rather that this apparent obviousness presents a problem for realists. Two opposite problems, in fact: it might be thought that realism is too obviously true to be worth saying; or it might be thought that anything so obvious to commonsense is probably false, like the ideas that the sun rises, that pigs sweat, that men are more rational and women more emotional, and so on. Oddly, these two objections are often combined: realism is *both* dismissed as obvious, *and* replaced by a non-realist account which is supposedly less 'naïve'.

But the following considerations suggest that the ordinary person's realism is not necessarily more naïve or likely to be wrong than the non-realism of some academics. Let us look at what might be called *regional non-realisms*. By this phrase I mean

views that some particular group of phenomena or of natural or social forces, which are generally taken to exist, do not. An example would be the 'Christian Science' view that pain and illness are unreal. We will generally find that they are held by people who have no practical dealings with the region concerned. I am not, of course, claiming that *all* regional non-realisms are false. But a great many that *are* false sustain themselves by practical disengagement from the aspect of the world about which they are non-realist. I doubt whether any surgeons have been converted to Christian Science. Now academics, at least in the arts, are mainly engaged in meta-discourse — that is, talking about talking — and do not, in their professional capacity, interact much with extra-linguistic realities. They are therefore prone to non-realism about such things.

For instance, I once attended a conference of literary critics at which one speaker was talking about Jean-Paul Sartre's account of his own childhood. A deconstructionist asked her, in a pained and patronizing tone, whether she was claiming that there really had existed such a person as Jean-Paul Sartre, independently of what we might say of him. When she said yes, she was, she at once lost the attention of the deconstructionist contingent. Now had I been rude enough to suggest that, while I agreed that Jean-Paul Sartre had most likely existed, I was not at all sure that there *was* anything that the deconstructionists were saying, I suspect they would have been upset. They were naïve realists about their own discourse, naïve idealists about dead French philosophers.

However, the realism which I shall be examining in this book, while it has accepted commonsense realism as far as it goes, certainly doesn't stop there. It is not like the rather mindless realism of G.E. Moore — the philosopher who thought that 'good' is the name of a simple property like yellow, and that oranges are yellow (see his *Principia Ethica*, chapter 1, section B). This is not a flippant remark. The theory of meaning according to which concepts name simple properties is brought into question by fact that Moore thought that oranges are yellow, while most of his countrymen hold them to be orange. It draws attention to the fact that meanings are by contrast, not binary relations of concept to property. Bits of language refer to the world as parts of a structure to parts of another structure, not as

simple concept to simple object. It is a mapping relation, not a pictorial one. Hence the grasp of thought on reality is a looser one than Moore allowed for, and it takes hard work to be a good realist in any given region. Moore thought that, if language was to have meaning at all, there must be some words that simply *referred* to some entity or quality, physical or moral. This has been called the 'Fido-Fido' theory of meaning, i.e. the theory that the meaning of the word 'Fido' is the dog Fido. This is not the place to discuss why this theory is now universally rejected by philosophers and linguists, but it is. Words as such don't refer to anything: *uses* of words refer to things, and one picks out the things they refer to by knowing the rules governing the *boundaries* between the correct use of one word and another, the *distinctions* we can make with them. So if my dog's name is Fido and my cat's name is Rutterkin, I can't say 'Rutterkin' meaning the dog, and expect to be understood. The word 'yellow' can be used to refer to the colour of yellow things because we understand it to contrast with 'orange' and all the other English colour-words. And 'good' would mean nothing if we could not contrast it, in various contexts, with 'bad', 'evil', 'best', 'rotten', 'nice', 'fair', 'invalid', 'righteous'.[1]

I have dwelt on this matter because one still hears it said that realism stands or falls with the Fido-Fido theory. Such a claim may be made for either of two reasons, one silly, the other serious. The silly one is the idea that since words get meaning from their relations (of contrast) with other words, they can't refer to anything outside language at all. But of course they can refer to things outside language *by virtue of* their relations with other words, just as a symbol on a map refers to a landmark by virtue of its relation to other symbols. The serious one is the claim that, since different languages make different contrasts between things, things have a different nature according to the language that you speak. Thus the Welsh colour-words 'glas' and 'llwyd' are often said to map their part of the spectrum differently from any English words, in that 'glas' covers certain greens, blue and certain greys, while 'llwyd' covers other greys and brown. But even if this is so, it does not follow that language gives reality its nature, since in the first place there has to be a *gradation* of differences before we can make contrasts, however arbitrary; in the second place, while *some* differences are

gradations, others are clearly marked in nature, e.g. between animal species or chemical elements; and in the third place, because it is not after all 'language' which makes distinctions, but *uses* of language, and we can, if we take the trouble, use language to make infinitely refined distinctions in context.[2]

Stronger and Weaker Realisms?

I have introduced the notion of regional realisms and non-realisms to throw doubt on the credentials of non-realisms, in that people are only ever non-realist about matters with which they are not practically engaged. But it might be asked whether this balkanization of realism does not make *general* contrasts of realism with non-realism vacuous. And indeed, Roy Bhaskar has said that every philosophy is some kind of realism, but realism about what? Berkeley is a realist about sensations, Plato about the forms, Bradley about the Absolute. The question then is, in the first place: which realism? And in the second, whether some realisms are more realist than others — for if not, the term 'realism' loses its polemical bite.

The word 'real', in many contexts, draws its content from its contrast with 'apparent'. A theory is realist in a stronger sense than others if it makes the following claims for knowledge:

1. *Objectivity*, in the sense that what is known would be real whether or not it were known: something may be real without appearing at all.
2. *Fallibility*: for insofar as claims are being made, not about some supposedly infallible or incorrigible data of appearance, but about something that goes beyond them, the claims are always open to refutation by further information.
3. *Transphenomenality*, going beyond appearances: knowledge may be not only of what appears, but of underlying structures, which endure longer than those appearances, and generate them or make them possible. We may have knowledge, not just of actions but of characters; not just of historical events but of social systems; not just of family likenesses but of the molecular structure of DNA.

4. *Counter-phenomenality*: knowledge of the deep structure of something may not just go beyond, and not just explain, but also *contradict* appearances. It is well known that Marx thought that it was precisely the capacity of science for counter-phenomenality which made it necessary: without the contradiction between appearance and reality, science would be redundant, and we could go by appearances. Later, I shall defend the weaker claim that its capacity for counter-phenomenality is what makes science a force for human emancipation.

But my next task is to show the contentfulness of the sort of realism that I am defending, by contrasting it, as depth realism, first with *actualism* — the commonest form of realism in empiricist cultures; and second with a variety of non-realism which often claims radical credentials.

The Inadequacy of Actualism

Roy Bhaskar has used the term 'actualism' for the view which, while asserting the reality of things and/or events and/or states of affairs, denies the existence of underlying structures which determine how the things come to have their events, and instead locates the succession of cause and effect at the level of *events*: every time A happens, B happens. (Such regular succession is in fact rare, as we shall see, except when produced by human agency.) In the philosophy of science, attempts to demote so-called 'theoretical entities' to mere explanatory constructs would be an instance of actualism; Mrs Thatcher's statement that there is no such thing as society is presumably another one, since she would doubtless not wish us to draw the conclusion that there are no such things as nations. Gilbert Ryle is said to have been rashly asked what he thought the ultimate constituents of the universe were, and to have given the actualistic answer 'things and chaps'. Even if we charitably assume that in his idiolect 'chaps' includes women, this is not adequate, since it does not even account for the fact that chaps can do all sorts of things that things can't. As against these actualisms, depth realism asserts that various kinds of entity —

molecules, trees, people, societies — have just those powers that they do and not others, by virtue of their respective inner structures. Hence these powers can often be ascribed, on the basis of knowledge of the structures, whether or not the powers are exercised. We may know that the spacecraft will work before we launch it — and if we don't, it is utterly irresponsible to do so.

Consistent actualism denies the existence of unexercised powers. Take Nietzsche, for example:

> A quantum of force is equivalent to a quantum of drive, will, effect — more, it is nothing other than precisely this very driving, willing, effecting, and only owing to the seduction of language (and of the fundamental errors of reason that are petrified in it) which conceives and misconceives all effects as conditioned by something that causes effects, by a 'subject', can it appear otherwise. For just as the popular mind separates the lightning from its flash and takes the latter for an *action*, for the operation of a subject called lightning, so popular morality also separates strength from expressions of strength, as if there were a neutral substratum behind the strong man, which was *free* to express strength or not to do so. But there is no such substratum; there is no 'being' behind doing, effecting, becoming; 'the doer' is merely a fiction added to the deed — the deed is everything. The popular mind in fact doubles the deed; when it sees the lightning flash, it is the deed of a deed: it posits the same event first as cause and then a second time as its effect. Scientists do no better when they say 'force moves', 'force causes', and the like — its coolness, its freedom from emotion notwithstanding, our entire science still lies under the misleading influence of language and has not disposed of that little changeling, the 'subject' (the atom, for example, is such a changeling . . .).
> (*Genealogy of Morals*, p. 45)

First, a word is perhaps required about Nietzsche's rejection of 'that little changeling, the subject'. This is not just a denial that there is any such thing as the *self*, after the manner of Hume's, Wittgenstein's, Sartre's or Lacan's denial of the self. It is a denial that there are 'powerful particulars' — beings, whether people, beasts, comets, Venus fly-traps or atoms which have the power, by virtue of their various internal structures, to bring about various kinds of events.

What are we to make of this passage? In the example Nietzsche gives it is of course true that lightning does not first exist and then exercise its power to flash — has the 'popular mind' really ever thought otherwise? But this is of no great ontological purport. No event or action exists before it occurs or is done, but its agent and/or patient always does. A battle does not first exist and then be fought, but the armies do first exist and then fight. As Roy Bhaskar notes, while it is true that 'let there be light' does not *mean* 'let something shine', there is in fact light only when something does shine.

Further, as to the case for which lightning is an analogy, the strong individual is certainly not always exercising his or her strength, nor is human desire a drive existing only in its satisfaction. Stuart Hampshire is nearer the truth, in his interesting paper 'Disposition and Memory', in claiming that thought, and the distinctively human character of our desires, originates in the manner in which we learn to restrain immediate satisfaction and express the desire in language. Perhaps this is just what Nietzsche dislikes, and contrasts with a de-sublimated superman. But Nietzsche's preference for such a superman is no argument for his existence, let alone for the non-existence of unexercised powers right across the ontological board. And indeed, he would not be much of a superman: he would not even be out of nappies.

Unexercised powers are in no way mysterious, as empiricists have sometimes thought, nor are they a mere shadow cast by language, as Nietzsche suggests. Wine cheereth the heart of God and man, according to the Good Book — but not so long as it remains tightly corked in its bottle. That is an unexercised power. Wine taken in combination with sufficient quantities of beer, cider, whisky, brandy, etc. may lead only to a sick headache; in this case, the power is exercised but unrealized, that is, it has its effects, but not the effects it would have had by itself.

I suspect that the empiricist[3] notion that there is something mysterious about powers, which must be exorcised from a down-to-earth world-view, stems from the lack of obvious connection, in English, between the noun 'power' and the verb 'to be able'. To say that there are unexercised powers is only to say that 'can' does not equal 'does'.

Yet this elementary distinction, whereby language does not cast its shadow on the world, but registers a widespread feature of the world, is laden with consequences. Politics provides a striking example. If history is just 'one damned thing after another', then all the politics we need is a resolve to do better damned things than were done before. If, on the other hand, societies and their institutions have inner structures which generate and by the same token constrain their powers, then we can ask, first of all, what sort of thing can be done given existing structures and what cannot; second, what different sort of things could be done given different structures; and third, how one sort of structures can be transformed into another.

As Roy Bhaskar puts it:

> All social structures — for instance the economy, the state, the family, language — depend upon or presuppose social relations — which may include the social relations between capital and labour, ministers and civil servants, parents and children. The relations into which people enter pre-exist the individuals who enter into them, and whose activity reproduces or transforms them; so they are themselves structures. And it is to these structures of social relations that realism directs our attention — both as the explanatory key to understanding social events and trends and as the focus of social activity aimed at the self-emancipation of the exploited and oppressed. (RR, p. 4)

This enables him to make the contrast, which the actualist cannot make, between the transformation of structures and the amelioration of states of affairs, as political goals.

Two words of caution are needed here. First, there is of course no disparagement of the amelioration of states of affairs. Indeed, if it is desirable to transform structures, that is so that states of affairs can be ameliorated. The point is that certain states of affairs cannot be ameliorated within existing structures. It is, for instance, inconceivable that permanent full employment or the vital degree of care for the environment could be achieved in a free market economy.

Second, it should not be denied that some things can be made better without changing the main social structures, and the question which can and which can't is ultimately an empirical one — though this does not mean that it can be answered only by making the attempt. Depth realism does not by itself tell you where to draw the line between improvements which can be made without transforming the structure of the state and the economy, and those which can't. But the point is that 'shallow realism' — the actualism that holds, or tacitly assumes, that there are no structures, only states of affairs — can't make the distinction between the two sorts of reform. At the theoretical level, this leads to such ideas as Popper's, that in a parliamentary democracy any change is possible once the majority supports it — an idea which has recently been used in legitimating parliamentary democracy in Eastern Europe. But of course this is not and could not be so. Parliamentary democracies, like any other kind of state, have structures that determine what can be done within the system and what cannot, and all attempts to transgress these limits, however popular and democratic, must fail. At the practical level, this actualist assumption leads to the paring down of programmes of reform to small-scale tinkering such as can be achieved without structural change, as if those things that are precluded by the existing structure of society were precluded by laws of nature.

'Shallow realism' or actualism, then, is less realist than depth realism in that it either denies the transphenomenality of the objects of knowledge, or reduces it to the relatively trivial case of the unbeheld tree in the quad, denying the reality of inner structures and consequent latent powers. This usually means that counterphenomenality is also denied. For to make sense of, for example, Marx's claim that the exchange of labour-power at its value is exploitation of the worker, one must agree that there is a deep structure to capitalist economic relations, which is exploitive, and which explains the surface-structure of the labour-market within which the labour contract appears as fair.

However, it should be noted that there can be philosophical positions which deny transphenomenality while accepting counter-phenomenality. Some such view seems to be that of Heidegger:

> 'Behind' the phenomena of phenomenology there is essentially nothing else; on the other hand, what is to become a phenomenon can be hidden. And just because the phenomena are in the first

place and for the most part *not* given, there is need for phenomenology. Covered-up-ness is the counter-concept to 'phenomenon'. (*Sein und Zeit*, p. 36)

Whereas many kinds of knowledge can be to some degree *empowering* — and here I may say that I would like to see the slogan 'knowledge is power' restored to its optimistic place on the banners of trade unions, far from the sinister connotations it has acquired in the covens of poststructuralism — the cases where one would say that knowledge is *emancipating* are special cases. Knowledge is here contrasted not with innocent ignorance but with false and enslaving appearances. In the words of the American cynic, it ain't that folks are ignorant, it's that they know so damn much that ain't so. And the false appearances here are not isolated or accidental mistakes, but the *motivated* false appearances which, at the social level, Marx has called 'ideology', and, at the personal level, Freud has called 'defence-mechanisms'. If the domains of knowledge opened up by Marx and Freud are liberating ones, that is because of the enslaving nature of these appearances, and the counter-phenomenal nature of these knowledges.

The Pitfalls of Non-Realism

So far I have been contrasting depth realism with shallow realism. What about non-realism? Even though non-realists may in the end turn out to be realists about something, they have a characteristic position, in that they deny that there is anything knowable that is independent of mind. If we take 'mind' in a wide enough sense, to include 'objective mind', discourse, etc. as well as subjective impressions, this formula covers not only Berkeley, Kant and Schopenhauer but Hegel, and also modern tendencies such as poststructuralism.

All these positions can be said to deny objectivity in the sense defined above; in the case of Berkeley and sense-datum theorists generally, it is uncontentious that this denial of objectivity is also a denial of fallibility, and is motivated by the quest for certainty — Descartes' quest for knowledge that could

not be doubted, transplanted into empiricist soil. Not that these philosophers claim infallibility for human knowledge in general, or indeed for their own theories; but they claim to place human knowledge on the foundation of an infallible stratum: sense-data, which, since they are purely appearances in the minds of individuals, have nothing to be mistaken about, and cannot themselves have mistakes made about them. Thus it is supposed that I may be mistaken if I say that I saw the dog chasing the cat, but not if I say that I saw a canoid patch of colour crossing my visual field in hot pursuit of a feloid patch of colour, to the accompaniment of miaowing and woofing sounds in my ears. I shall not reiterate the often-repeated objections to this view.

What is perhaps less obvious is that modern, discursive idealism also, if I may so express it, fails to be fallibilist. It tries. It takes its starting point from the fact of scientific change, and hence the recognition that just as past scientific theories are now abandoned, our scientific theories are likely to be abandoned in due course.

Indeed, modern non-realists often accuse realists of dogmatism because of our defence of objectivity. They accuse us of arrogance in claiming truth for our theories, and preen themselves on their modesty in proffering only tentative theories. It is surprising how often even those who are well aware that science is inherently fallibilistic in its practice will say that claims for scientific status on the part of this or that theory are a ploy to make the theory immune from criticism. This line of attack is often used against Marx, despite his, Engels's and Lenin's explicit espousal of fallibilism at times when it was less generally accepted than today. Furthermore, radical feminists have sometimes alleged that objectivity is just male subjectivity; and there is a school of family therapists who hold that belief in objective truth causes most of the problems of the world — and on these grounds finds it acceptable to lie to their clients. In all these cases, claims for objectivity are decried as arrogant, whereas disavowal of it is billed as open-minded.

But the boot is entirely on the other foot. To claim objective truth for one's statements is to lay one's cards on the table, to expose oneself to the possibility of refutation. It is to make it clear that one is talking *about* something, and saying that that

'something' is thus and not so; this makes it possible for others to point out features of that something which are not as claimed, and hence disprove your opinion. All claims to objective knowledge are vulnerable in this way. There are of course other discourses that are not. And importantly, there are many forms of words that may in one context be making claims about how the world is, while in others they are not. When asked 'how are you?', some people reply with a brief medical autobiography, but most say 'very well, thank you', meaning only to acknowledge a polite greeting, not to report their state of health. The words of the Nicene Creed may be intended to state the truth when said in church by an orthodox believer,[4] but not when sung in a concert hall as part of a performance of Beethoven's Mass in C. The statement 'there were an Englishman, an Irishman and a Scotsman' is usually the opening line of a joke, though it could be the opening line of a lecture on the British empiricists. Depending on the interpretation of these statements, it would be appropriate or inappropriate to question their veracity. In the cases in which no objective claims were being made, it would be absurd to criticize on grounds of falsehood — though one might have other grounds: one might have a Quakerly aversion to formal greetings, or no ear for classical music, or a moral objection to national stereotypes in humour. But the difference in these cases is clear to anyone.

Non-realism assimilates cognitive forms of discourse which only make sense on the assumption that they do make claims to objectivity, to the other sorts. This renders them invulnerable to any criticisms based on the claim that the facts are different. Non-realism as a meta-theory licenses any and every form of dogmatism at the first-order level. It enables the theorist to say 'since I am not claiming objective truth for my theories, I can go on saying what I like, and your counter-examples have no relevance for me' — and then to go on saying things that have no point at all unless they are making claims about how the world is. I am not just saying that this might be expected to happen; wherever non-realism gains credence, uncritical dogmatism flourishes, from the epidemic of superstition and bigotry under the lee of Ockham's philosophy in the late Middle Ages,[5] to the oracular style and indifference to evidence favoured by Lacan,

in contrast to the realist Freud's agonized questioning of his own theories.

It might be said: non-realism may make cognitive discourse invulnerable to cognitive assessment, but it lays it open to the other kinds of criticism. It enables us to criticize a scientific theory, for example, as being aesthetically or morally or politically nasty. It subordinates theoretical criteria to practical criteria in assessing theories. And this is sometimes defended, in the name of the 'primacy of practice'. This position has to be met on its own terms: is it *practically* desirable that theories should be accepted or rejected on practical rather than theoretical grounds?

This kind of 'primacy of practice' undermines the possibility of subjecting a practical orientation to a certain kind of critique. It prevents us from saying that a given practice rests on certain false or contradictory beliefs. Practical attitudes become immune to theoretical critique, and, by the same token, are reduced to *mere* attitudes, which may certainly clash with other such attitudes, but not be argued about rationally.

There are a number of ways in which practice does have primacy over theory, but primacy in some respects does not imply primacy in every respect. The practical importance of theory is that a theory can transform a practice. Theory is the growing point of a practice, and to abolish its autonomy is to nip off that growing point. Furthermore, a theory transforms a practice, in the best case, by exposing and correcting cognitive errors implicit in that practice. Those who demand that theory, for instance in politics, be judged by practical criteria rather than by its adequacy to reality are generally saying that the criteria of some *existing* practice should judge the theory. They are in the business of conserving existing political practices, protecting them from rational criticism. The only kind of criticism which can be accommodated by this view is the sort which is based on an arbitrary, irrational leap.

Depth realism, on the other hand, is in four ways transformative and potentially emancipatory:

1. in that, because it allows that knowledge may be counter-phenomenal, it makes a place for our liberation from enslaving appearances;

2. in that, because it calls for theories to be judged by objective criteria, it promotes theories that can transform, rather than merely rationalize, existing practices;
3. in that, because it recognizes that states of affairs are brought about by the working of relatively enduring structures, it directs the attention of people who want to make the world a better place to the task of transforming those structures;
4. in that, because it recognizes that theories must make claims about what the world is like independently of those theories, it treats all theories as fallible, and open to transformation.

Why Philosophy?

Suppose it be granted that our account of our knowledge must be a realist one. It is still open to doubt whether we need a realist *philosophy*, as distinct from simply being realists in the practice of science or of everyday knowledge. Historically, the turn from idealism to realism has often also been a turning away from philosophy altogether, towards the particular sciences, which are considered to provide all the theoretical knowledge we need or can have (this is *one* of the meanings of 'positivism'). Connected with this doubt is another comment I have heard made about Bhaskar's project: realism is fine, but why 'go transcendental'? I shall try to answer these points.

A good part of the answer to the question 'why philosophy?' is that the alternative to philosophy is not *no* philosophy, but *bad* philosophy. The 'unphilosophical' person has an unconscious philosophy, which they apply in their practice — whether of science or politics or daily life. As Gramsci puts it:

> Having first shown that everyone is a philosopher, though in his own way and unconsciously, since even in the slightest manifestation of any intellectual activity whatever, in 'language', there is contained a specific conception of the world, one then moves on to the second level, which is that of awareness and criticism. That is to say, one proceeds to the question — is it better to 'think', without having a critical awareness, in a disjointed and episodic way? In other words, is it better to take part in a conception

> of the world mechanically imposed by the external environment, i.e., by one of the many social groups in which everyone is automatically involved from the moment of his entry into the conscious world. . . . Or, on the other hand, is it better to work out consciously and critically one's own conception of the world and thus, in connection with the labours of one's own brain, choose one's sphere of activity, take an active part in the creation of the history of the world, be one's own guide, refusing to accept passively and supinely from outside the moulding of one's personality. (*Prison Notebooks*, pp. 323–4)

On the one hand, this means that philosophy works by making explicit knowledge that is already implicit in some practice or other. Thus Bhaskar can cite Kant approvingly to the effect that it is 'the function of philosophy to analyse concepts which are "already given" but "confused"' (RTS, p. 24). The practice in which the concepts are implicit may itself be either a cognitive one (like science) or a non-cognitive one (politics, personal relations, the work-world, art). No doubt most of the knowledge we have got is implicit in our non-cognitive practices, and much philosophy is concerned with explicating that knowledge — for example Heidegger's existential analytic as an explication of our work-world, or many texts on ethics, before the British linguistic philosophers corrupted and debased that discipline. But when the practice to which philosophy turns is itself a cognitive practice, that in no way alters its relation to philosophy; it is not the *cognitive results* of science which interest the philosopher *qua* philosopher. It is the set of concepts implicit in the practice of the science, and which the scientists *qua* scientists do not need to make explicit, and may not even suspect that they use. Bhaskar does not derive his conclusions about the structure of the world from, for example, the theory of relativity, or quantum theory, or the theory of evolution. Attempts so to do are always blind alleys. But as we shall see, he is able to derive very far-reaching ontological conclusions from the practice of scientific experiment itself.

Philosophy, then, as a critical rational activity, presupposes that we have 'philosophies' implicit in our practices, which can be made explicit. But this is not simply a matter of *spelling out* our implicit knowledge, to satisfy our curiosity or our desire for self-knowledge. The work of philosophy can perform two

polemical functions in relation to the practices it lights up: a critical one, when it exposes internal contradictions in the beliefs implicit in the practice; and a defensive one, when it shows how the practice does what some (absolutely or relatively) a priori theory claims cannot be done. This defensive function is often a matter of one philosophy defending some non-philosophical cognitive practice against the objections of another philosophy. But it is best done not by a general refutation of the objection but by *pointing* to the practice doing what had been said to be impossible, and spelling out the steps by which it does it. We shall see later how Roy Bhaskar shows us social science validly deriving values from facts, despite the massed ranks of post-Humean philosophers claiming that this is impossible (see chapter 6).

But this is very far from obscurantism, after the manner of those Wittgensteinians who say 'these language games are played', and rule out objections to any ongoing practice, insofar as those objections allege incoherency. For to illuminate a practice may also be to criticize it. The criticism will indeed be *internal*, but only in the sense that the contradictions discovered are internal to the practice, rather than contradictions between the practice and assumptions brought from outside it. This does *not* mean that only a practitioner can understand, or criticize, a practice. Those who deny the right of a non-South African to criticize apartheid, or of the unanalysed to criticize psycho-analysis, or of men to criticize feminism, or of unbelievers to criticize theology, are merely exposing their own bad intellectual conscience. If these practices (cognitive or otherwise) can be understood, they can be criticized, and criticized internally, by one who is outside them.

For example, a pseudo-science may be exposed by showing that, while claiming to uncover some secrets of nature, no uncovering is being done. In showing this, no external demands are being made on the practice. There are many practices which do no uncovering of nature, and make no claim to. The contradiction occurs when the claim to be uncovering something is essential to the practice, yet something quite different is actually done by it. In cartomancy, for example, the result is supposed to be the result of an inquiry into nature, but the process by which it is obtained is not inquiry but shuffling, etc.

What then of the generalizing function of philosophy? Does the dependence of each bit of philosophy on some other bit of practice mean that it is lost? Can we not transfer knowledge derived from one practice to the practice of lighting up another? To this last question, I think the answer is: except under special circumstances, we can do so only with the knowledge of possibilities, not with that of impossibilities. Marxist theory, for example, may say to aesthetics 'there are such things as dialectical contradictions', but not 'there is no such thing as classless art'. The special circumstances, connected with stratification and emergence, will become clear in the light of chapter 4. They concern the constraints placed by the nature of an entity at one level (e.g. a human body) on its powers at another level (e.g. a human agent).

The above characterization of the practice of philosophy and its relation to other practices is intended to apply to any critically rational philosophy which neither disclaims any cognitive work (after the manner of positivism), nor claims access to any special, esoterically philosophical knowledge. In characterizing Roy Bhaskar's practice of such philosophy, four further points must be made, three brief, one long.

1. While he holds that philosophy may work on other subject matter than the sciences, his own work is almost entirely based on the practices of sciences, both natural and social.
2. He holds the main work of such philosophy to be an *underlabouring* one. He aims to remove the idols (Bacon), obstacles (Locke) or ideologies (Marx) that stand in the way of, or distort the understanding of, new knowledge to be produced by the sciences.
3. In addition to this underlabouring role (first so dubbed by Locke), Bhaskar sometimes talks of an 'occasional' role for philosophy as the *midwife* of new sciences. Philosophy has often had this role historically, though not always with happy results — consider some of the little monsters delivered by positivism, making up a large part of that unhappy family, 'the human sciences' (particularly in psychology). I take it that a good deal of the motivation of his work is to replace this positivist brood by something both more scientific and more conducive to human emancipation.

At times he seems to suggest that these better human sciences are already overdue for birth, and merely await the arrival of a suitably qualified midwife. (I must confess to being more pessimistic about this matter.)

4. Finally, in the following section, let us consider the features of Bhaskar's thought which warrant the term 'transcendental'.

Transcendental Arguments

One specific way in which philosophy can turn the light on a practice is by what have been known, since Kant, as transcendental arguments. In such arguments, we ask 'what must be true in order for x to be possible?', where 'x' usually refers to some feature of human activity. For example, Sartre's existentialism starts from the question 'what must be true of human consciousness in order for it to be able to ask (genuinely open) questions about the world? How is it that, when we go into a café to look for Pierre, our consciousness is not just filled with the sights and sounds of the café), but with the absence — or possible presence — of Pierre?' (*Being and Nothingness*, pp. 9ff) In order to account for our capacity to question the immediate input of experience, Sartre is led to the view that consciousness is radically free from prior or external causes. I shall not discuss whether this argument is valid (I don't think it is), but merely note its form: from something that is actual, to a more fundamental 'something' that grounds its possibility. In general, such arguments will be from a phenomenon that occurs to a structure that endures, though this is a relative distinction.

The place of transcendental arguments in Roy Bhaskar's work can be brought out by showing the parallels with, and differences from, Kant's transcendental arguments. Kant's initial question is 'How is synthetic *a priori* knowledge possible?'[6] — knowledge, that is to say, which tells us something about the world (unlike statements which are true by definition), yet can be known independently of any experience. The substance of his work, however, is on the question 'how is empirical knowledge possible?' The answer to this question

answers the first one too, since empirical knowledge presupposes that the known world has a knowable structure — i.e. that it is ordered in space and time, behaves in a regular manner, consists of things and their properties, which can be measured, and which change only in accordance with causal laws, and so on. The synthetic a priori knowledge turns out to be knowledge of what these features of the world are — the features that *must* be the case if the world is to be known. However, Kant didn't think that the world in itself had these properties; rather, he thought that our mind imposed this knowable form on it.

Bhaskar's fundamental question is closely parallel to Kant's, and the parallel and differences are summed up in his appropriation of the term 'transcendental realism' — also used by Kant, in a slightly different sense, as we shall see. The chief differences are as follows.

First, while Kant clearly had the science of his time in mind, he talks about knowledge in general; Bhaskar neither regards science as the only source of knowledge, nor treats it as smoothly continuous with other forms. It is a special kind of knowledge, presupposing other kinds, yet also able to contradict and correct pre-scientific ideas. So Bhaskar's initial inquiry — though it is possible to broaden it afterwards — is into the possibility of scientific experiments.

Experiments, like the other practices from which we derive knowledge, are not purely 'mental' activities. They involve intentional causal interaction with the world about us — interaction which is possible only because we are embodied beings, subject to the same laws that govern our material environment. That we have got hands and eyes and ears is as necessary for knowledge as that we have got reason and imagination and memory. And this suggests that there may be other transcendental arguments too, which throw light on practices other than cognitive ones.

In fact it might be a useful digression here — if only to dispel some of the mystery that, for many people, attaches to the word 'transcendental' — to ask what happens if we ask seemingly quite frivolous 'how is x possible?' questions. For example, how are chickens possible? At least such a question will demand an answer of a rather different order than 'chickens come from

eggs which come from chickens which come from eggs. . .' that standard case of a futile explanatory regress. We shall want to know not about chicken-and-egg but about chicken-and ecological-niche. In the case of domestic fowl, of course, the ecological niche will include human agricultural and culinary practices. But as with any other explanandum, an explanation of this sort will tell us about wider structures of the world, not about the particular antecedents of particular chickens (eggs). Perhaps this illustrates Spinoza's distinction between two kinds of explanation (I may say that it was in trying to teach Spinoza's philosophy that I first got on to the track of these transcendental fowl). But in this case the two kinds are equally empirical. Chicken—ecological-niche questions belong to a different scientific sub-discipline than chicken—egg questions, but neither are philosophical questions, though the difference between them may be philosophically interesting. Roy Bhaskar says that transcendental arguments are a species of *retroductive* argument, i.e. arguments 'from a description of some phenomenon to a description of something which produces it or is a condition for it' (SRHE, p. 11). He does not say what the differentia is, but perhaps the term 'transcendental' should be reserved for arguments at a philosophical meta-level — and hence not about chickens. However, there is no reason for restricting either genus or species to a privileged set of 'big issues'.

Second, while Kant's arguments lead to a theory about the power of mind to impose a structure on the world, Bhaskar's lead to conclusions not only about the mind or about ourselves, but also about what the world must be like. Bhaskar's philosophy can therefore dispense with the unknowable 'noumena' or things-in-themselves which haunt Kant's philosophy. However, it does not dispense with them in the same way as Kant's idealist successors did — by denying that there is a world independent of the knowledge minds may have of it. The nature of the *work* we must do in order to find out about the world shows us both that the world is not transparent to us but needs to be discovered, and that it can be made to yield up its secrets.

Kant had grounded the a priori nature of his synthetic a priori — its independence of experiences — on the idea that *we* impose

it: 'what reason produces entirely out of itself cannot be concealed, but is brought to light by reason itself immediately the common principle has been discovered' (*Critique of Pure Reason*, p. 14). For Bhaskar, those features of the world which make knowledge possible are not necessarily a priori; they are real features of the world, which could have been otherwise. There is nothing impossible about an unknowable (and hence necessarily unpeopled) world, but granted that we do exist knowingly in the world, we can construct transcendental arguments from this fact to illuminate some structures of that world.

Finally, Kant's conception of his theory of knowledge is timeless, both in the sense that (despite his foregrounding of Newtonian mechanics) it is supposed to apply to human knowledge in general, irrespective of its historical forms; and in the sense that he believed himself to have discovered the key to certain eternal truths. Bhaskar makes no such claims. Scientific experiment in the relevant sense is a relatively new phenomenon. While most of the facts discovered by science had been true but unknown before (the earth went round the sun, before Galileo proved it), the *form* of experimental science cannot be read back into pre-scientific forms of knowledge — or of course 'read forward' into forms of inquiry as yet undiscovered. And since the structure of the world is neither necessary nor transparent to reason, our knowledge of it is always fallible. A transcendental argument may account for the possibility of some phenomenon, but there may be rival transcendental arguments to explain the same thing, just as there are rival theories at the frontiers of science. One transcendental argument may explain more than others, and so be the best available account. But in philosophy as in science, while there can be justified beliefs and there can be progress, there can be no *final theory*, unsusceptible to revision and improvement.

In order to tie the foregoing description of transcendental arguments to Bhaskar's own account of them, I now want to look closely at two paragraphs from PN (pp. 5—6).

If philosophy is to be possible . . . then it must follow the Kantian road. But in doing so it must both avoid any commitment to the

content of specific theories and recognize the conditional nature of all its results.

Thus, Kant's tendency to take the most fundamental results of his contemporary sciences as true a priori and hence unrevisable must be avoided, by a twofold refrainment: (1) as I have already suggested, it is not the *results* of science which philosophy takes as its premises, but the *practice* of science; hence, the revision of particular results (e.g. Newton's theories) does not as such necessitate philosophical revision. (2) Even so, philosophy should make no claims to be unrevisable; a radically new *kind* of knowledge of nature might require a new philosophy — and even without such new forms of knowledge, any philosophical theory is always liable to be replaced by a better one based on the same premises in scientific practice.

> Moreover it must reject two presuppositions which were central to Kant's own philosophical project, viz. that in any inquiry of the form 'what must be the case for ϕ to be possible?' the conclusion, X, would be a fact about us and that ϕ must invariably stand for some universal operation of mind. That is to say, it must reject the idealist and individualist cast into which Kant pressed his own inquiries.

The first (anti-idealist) point is clear; a transcendental argument may tell us how the world must be structured, not how the mind must. The point about individualism is less transparent. I take it that what is in question is Kant's focusing on what every human mind must do, rather than on historically specific social practices.

The subject-matter of philosophical investigation, Bhaskar goes on to suggest, is 'the necessary conditions for social activities as conceptualized in experience' — a wider field than Kant's, since not only universal but also 'historically transient' activities and conceptualizations may be investigated, and because 'the activity may depend on powers that people possess as material things rather than just as thinkers or perceivers'.

He concludes that philosophy 'operates by the use of pure reason. But not by the use of pure reason alone.' At first reading, this may sound paradoxical, since 'pure reason' and 'reason alone' mean much the same, and 'pure reason alone' looks like a pleonastic equivalent. But the following sentence makes clear what is meant. 'For it always exercises that reason on the basis of prior conceptualizations of historical practice, of some more or less determinate social form.' So philosophy's *manner of work* is pure reason, but its raw materials are not. Hence the transparency claimed by Kant for his subject of inquiry in the last quote from him *cannot* be expected. In the following two paragraphs (PN, pp. 5–7), Bhaskar makes two main points about philosophy to distinguish it, on the one hand, from strictly a priori reasoning with its claim to unrevisable certainty, and, on the other, from science. Unlike the search for 'necessary truths', the transcendental arguments on which this conception of philosophy is founded have two aspects: they show how the conclusion ('the world must be thus') accounts for the possibility of the activity (or whatever) that forms the premiss; and they show the incoherency or implausibility of alternative accounts. Hence they are always situated polemically in relation to those accounts, and there can be no guarantee that a better rival account will not come forward. In practice, of course, Kant also argued in this way: consider his critique of Newtonian and Leibnizian theories of space. But in principle, Kant's arguments are supposed to establish not just that his account is the best runner but that it is necessarily true.

In its fallibility, philosophy resembles science. But while it is about the same world as the sciences, it does not compete with them. It 'can tell us that it is a condition of the possibility of scientific activities ϕ and ψ that the world is differentiated X and Y. But it cannot tell us *what* structures the world contains or *how* they differ. These are entirely matters for substantive scientific investigation.' Just what this means in concrete terms will become clear in the next chapter, after I have considered Bhaskar's central transcendental argument from the possibility of scientific experiment.

The Term 'Transcendental Realism' in Kant and Bhaskar

When Roy Bhaskar uses the term 'transcendental', he is obviously using it in a sense close to Kant's, and not, for instance, to 'Transcendental Meditation', or Dr Jekyll's

'transcendental medicine'. For Kant, transcendental questions are questions about the boundaries of possible experience (knowledge), and as such contrast not only with empirical questions, concerning the content of knowledge, but also with *transcendent* questions, which overstep those boundaries. And Kant not only uses the term in the phrase 'transcendental idealism', to designate his own philosophy, but also in the phrase 'transcendental realism', to designate his opponent on the issue of space and time, the one who 'regards time and space as something given in themselves, independently of our sensibility' (*Critique of Pure Reason*, p. 346). Possibly, he would also apply the term transcendental realist to someone who attributed like objectivity to causality, substance, etc. — the categories that, according to Kant, our understanding imposes on the phenomena. Is this the sense in which Bhaskar is a transcendental realist?

The answer, I think, is that while Kant would no doubt regard Bhaskar as a transcendental realist in his (Kant's) sense — and rightly, so far as it goes — the phrase functions rather differently in Kant and in Bhaskar. This is brought out by the ways transcendental realism is contrasted with empirical realism: this is not the same difference in the two philosophers. For Kant, empirical realism is realism about the concrete contents of experience — an empirical realist believes that chickens and blizzards and magnetic fields exist independently of the observer. Kant claims to be one. But whereas for Kant the *form* of experience (here, space and time) is not real but contributed by the mind, the transcendental realist is, for Kant, the realist about form as well as content.

While Bhaskar is, no doubt, a transcendental realist in this sense, the use to which he puts the concept of transcendental realism is a little different. First, this is so in that the contrasting term, empirical realism, is used not for anyone who holds concrete objects to be real — and hence Kant and Bhaskar, as well as most empiricists — but for one who also *denies* the reality of *underlying* mechanisms, structures etc., which don't appear in experience, but cause phenomena that do. A transcendental realist, by contrast, is one who claims that such mechanisms can be shown to be real by means of transcendental arguments (the details of which I shall discuss in the next chapter).

The reality of underlying mechanisms is not the same issue as the reality of conditions of the possibility of experience, but they are closely connected. Underlying mechanisms are discovered by the particular sciences. But transcendental arguments taking the practice of these sciences as their premises can establish that there must be *some* such mechanisms[7] if those sciences are to be possible. And that is a proposition very like some of Kant's, about the conditions of the possibility of knowledge. For Kant believes that his transcendental arguments show that (as a necessary condition of there being knowledge) every alteration must have a cause — though it is up to the empirical sciences to discover particular causal laws. We are concerned with transcendentally established conditions of the possibility of knowledge in both cases. The differences are (1) Bhaskar's premisses are special ways of acquiring empirical knowledge, not empirical knowledge in general (though in principle *any* such ways could be taken as a premiss of such an argument); (2) Bhaskar arrives at the richer hypothesis that there must be a *multiplicity* of causal mechanisms; (3) Bhaskar's conclusions are a priori only in the relative sense: that in order for *any* of the kind of knowledge described in the premiss to be possible, the conclusion must be true; (4) if transcendental arguments tell us what must be the case, Bhaskar's 'must' is different from Kant's, in two ways: (i) it does not mean 'the phenomena are compelled to conform to these conditions — our minds have the power to impose them' but 'we are compelled to assume that these conditions really hold — otherwise our knowledge would be impossible, and it is actual' — hence 'realism'; (ii) it is not a claim to necessary truth; it is open to refuting arguments. It is more like an explanatory 'must' than a 'must' of logical necessity.

I mentioned earlier a very Kantian list of conditions of the possibility of (our kind of) knowledge: the world must be ordered in space and time, behave in a regular manner, consist of things and their properties, which can be measured, and which only change in accordance with causal laws. Can this sort of conditions really be established outside of an idealist, synthetic a priori framework? While Bhaskar's list may be slightly different (to include the structuredness of things, the stratification of nature, etc.), I think the answer is yes. I would

like at this point to suggest that it is quite plausible that we have got implicit knowledge of such synthetic *relatively* a priori propositions, by means of an example (which certainly does not amount to a transcendental argument).

It would be most apposite if the example could be concerned with space or time, since these are the issues on which Kant explicitly contrasts his own view to transcendental realism. But this question has lost the thematic unity that it had for Kant, whose theory was meant to cover both the common experience of space and time, and their place in science. Scientific notions of space and time have been so transformed in the wake of Einstein that they have been irretrievably prised apart from experiential concepts of space and time; moreover, I would suggest that Heidegger has shown that the attempt to assimilate lived space and time to *any* scientific concept of space and time was always an error. So let us consider another example, where it is at least arguable (though also debatable) that this falling apart of science and experience has not occurred: the concepts of substance and causality, taken together — i.e. the idea that things don't change unless a causal power of something (whether an internal tendency or the force of something outside) changes them.

Let us ask why we do not accept the following, extremely well verified hypothesis, which has great predictive power: 'from time to time, household objects (books, earrings, cutlery, spare parts of sewing machines, gramophone records) cease to exist without trace.' On the basis of this 'Law of Disappearing Household Objects', I can predict (and it is more certain than any meteorological or economic prediction) that in six months' time some valued household object now in my possession, and which I have not destroyed, sold, lent or transported, and which I am certain no friend or burglar has taken, will no longer be in my possession.

Our rejection of such a law is prior to and independent of any scientific laws with which it might be incompatible. From the age of about three, we confidently ask of any missing object 'what has happened to it?' — even though at that age such disappearances are frequent enough to make any genuinely empiricist child wonder rather by what strange mechanism things sometimes turn up in their right place.

I suggest that this indicates that we are not really empiricists. We implicitly take some of the Kantian categories for granted as presuppositions of all empirical questions about the world, as things without which those questions would make no sense. And it is reasonable to do so. A transcendental refutation of the Law of Disappearing Household Objects would merely spell out what every normal three-year-old knows — even though normal three-year-olds are also quite capable of understanding, and making up, stories about worlds in which such laws do hold. It is not that our minds shut out such possibilities; it is that we go about getting to know the world in ways that we could not if they were actualities.

So far I have been pointing out ways in which transcendental realism differs from some other positions (empiricism, relativism, Kant) and suggesting reasons why we need a philosophy which differs from them in such ways. In the following chapter I shall look closely at the *arguments* for the main transcendental realist claims.

Notes

1. Consider Nietzsche's contrast between the opposites 'good/evil' and 'good/bad'; the proverb 'the good is the enemy of the best'; good and rotten apples; *The Nice and the Good* (novel by Iris Murdoch); school reports graded 'good, fair, medium, poor, bad'; pass cards 'not good unless signed'; St Paul's 'For scarcely for a righteous man will one die: yet peradventure for a good man some *would* even dare to die' (Romans, 5.7).

2. Welsh speakers tell me that in fact 'glas' and 'llwyd' do correspond to 'blue' and 'grey', though there are anomalous usages such as 'cae glas' for a lush green field, 'papur llwyd' for brown paper — much as in English we call white horses grey, red hunting gear pink, greyish pink people white and honey-coloured people black. But the loss of the example does not alter the point — except to draw attention to the fact that it is uses of words, not words themselves, that have meaning.

3. While suspicion of talk about powers is characteristically empiricist, not *all* empiricists avoid such talk. Powers have a central place in Locke's philosophy, for instance.

4. Someone may say that the orthodox believer is not making truth-claims either, since all religious language is symbolic. Certainly, the orthodox believer is not committed to the idea that, for instance, God the Father has a right hand. But language can be used non-literally to make truth-claims. If I say 'Aunt Maud would have kittens if she heard that news', and Aunt Maud is actually unperturbed by the news, then I was mistaken. If saying the creed in church is not saying *something* about what exists and has happened, then it does not differ

essentially from singing it *with feeling for its meaning* in a concert hall — and no believer could assent to that.

5. The sense in which Ockham's philosophy contrasts with realism may be thought unrelated to the present sense of realism, since some (e.g. Engels) have claimed Ockham as a forerunner of 'materialism' (i.e. a version of realism in the modern sense). But I would argue (though this is not the place to do so) that transcendental realism sides with Thomas Aquinas and John Wyclif against William of Ockham in the medieval controversy over universals. This is assuming that realism about universals takes an Aristotelian, not a Platonist, form. Plato's realism about the Forms is one of the two *philosophical* senses of 'realism' in which transcendental realism is *not* realist. Perhaps this is the place to mention the other.

Heidegger, towards the end of the existential analytic, argues forcefully that non-realism is a non-starter, as it presupposes a worldless subject, and we are essentially Being-in-the-World. He admits that this puts him on the side of realism 'doxographically, as it were' — i.e. in terms of alignment on one side or the other of a historic dispute. But he goes on to say that 'compared with realism, *idealism*, no matter how contrary and untenable it may be in its results, has an advantage in principle' since it 'expresses an understanding of the fact that Being cannot be explained through entities' (*Being and Time*, p. 251). He seems to be taking realism (perhaps on etymological grounds) to be the doctrine that 'things', in the narrow sense of physical objects (in his terms, that which is *vorhanden*), are the only reality. In other words, realism = what English-speaking non-Marxist philosophers call 'materialism', and Marxists call 'vulgar materialism'. Transcendental realism, with its theory of emergence (see chapter 4), and of the efficacy of reasons (see chapters 4, 5 and 6), is not of this kind.

I had thought this usage was an aberration of Heidegger's, but C.S. Lewis, who was of course familiar with the Oxford philosophical scene, and careful in his choice of words, uses 'realism' in much the same sense. He tells us, in effect, that he abandoned 'realism' since it undermines itself by treating thought as governed not by its own norms but by physical laws. (*Surprised by Joy*, pp. 167–8).

6. I am taking slight, and defensible, liberties with Kant's phraseology. He introduces his fundamental question as 'How are *a priori* synthetic judgments [not knowledge] possible?' (*Critique of Pure Reason*, p. 55) — but he is not concerned with the unprovable synthetic a priori judgements that any fool could make up all day long, and some do. And I take Kant to mean by 'experience' something more like empirical knowledge than like some 'stream of consciousness', or whatever people mean by 'experience' in this subjectivistic age.

7. This formulation may be at risk of making these transcendental arguments look trivial: given that science discovers underlying mechanisms, science would not be possible if there were no underlying mechanisms — rather as if one were to present a 'transcendental argument' from the actuality of the activity of gathering blackberries to the reality of blackberries. The contentious issue, however, is whether what sciences do really is discover underlying mechanisms.

Experiment and Depth Realism

> Scientifically significant generality does not lie on the face of the world, but in the hidden essences of things.
> (RTS, p. 227)

How are Experiments Possible?

Most of the leading ideas of transcendental realism are rooted in a single transcendental argument which answers the question 'how are experiments possible?' This is not only an extremely fertile question for the philosophy of nature and of our knowledge of it; it is also a key strategic question in the polemical situation of modern philosophy. For experiment is the defining activity of 'the experimental sciences', and their prestige as our foremost means of discovery of nature depends on it. Both empiricism and transcendental idealism, in all their forms, stand or fall with their capacity to account for the success of these sciences. And if Bhaskar's argument is right, they fall.

I have formulated the question, in Kantian style, 'how are experiments possible?' The first part of the question, though, is 'why are experiments necessary?' Since we would hardly go to the bother of setting up experimental situations if we could get the same information without them, they are possible only if necessary. If we could, as Aldous Huxley says of D.H. Lawrence, taste the hydrogen and oxygen in water, we would not need to separate them by electrolysis. Knowledge which we in fact have only by virtue of scientific experiment (water = H_2O) could then have been acquired in the same way as we discover that grass is green and lemons are sour. Most of our knowledge, after all, is not acquired experimentally, and is not scientific. We do not look for it, we happen upon it while doing

something else. We could imagine a possible world in which everything there was to be known could be discovered in this way. But if, in *our* world, we restricted ourselves to such sources of knowledge, we would never have got out of the Middle Ages.

Empiricism and transcendental idealism presuppose a 'modern' world outlook, i.e. one transformed by science. Indeed it is arguable (though this is a different argument from the present one) that they *misrepresent* the world of our everyday experience by assimilating it to that discovered by science. Yet both these philosophies offer pictures of our knowledge of the world that, were they true, would make experimental science redundant. Empiricism in particular tends to flatten the distinctions between different sources of knowledge: knowledge comes from 'experience'; experience consists in nature's impressing its image on the blank 'wax tablet' of our minds. How, then, could one experience (say, testing the boiling point of a liquid) be more significant for knowledge than another (say, staring at the blue sky)?

Let us consider a scientific experiment. I have chosen the example for its simplicity and elegance. It might be considered a disadvantage of the example that the point the experiment 'proves' would no longer be accepted by science. However, as to the available theories between which it arbitrated, it refuted one and supported another; and in so doing made way for the emergence of a new concept with lasting importance: the magnetic field.

The experiment (recounted by Harré in his *Great Scientific Experiments*, pp. 49–56), was made by Robert Norman, and published by him in *The Newe Attractive* (1581), though it was left to William Gilbert to place a more fruitful interpretation on it. It had previously been assumed that a compass needle was so to speak pulled to the north as a 'point attractive'. The conclusion to be drawn by Gilbert was that 'the direction is not produced by attraction but by a disposing and conversory power existing in the earth as a whole' (quoted by Harré, p. 53).

The experiment consisted in magnetizing a piece of wire that had been thrust through a cork such that, when placed in a glass of water, the wire was suspended a little below the surface. Once magnetized and replaced in the glass, the wire will lie north and south, dipping towards the north, but without

descending to the bottom or moving to the north end of the glass.

How does this experiment tell us more than we know already from using a compass? A fixed compass needle can only point, it can't move from the centre of the dial, since another force than the earth's magnetism is restraining it. By removing that other force, we allow the earth's magnetism to operate on it unimpeded. It does what the mechanism being tested (the earth's magnetism) makes it do, and not what anything else makes it do. Hence we can discover *what* that mechanism makes it do. Under non-experimental conditions, we can see only what that mechanism *in conjunction with other factors* makes it do. And in conjunction with the fixity of its swivel, *either* of the two postulated mechanisms would make it do what it does, i.e. point north without travelling from its position.

What the experiment does, in short, is to isolate one mechanism of nature from the effects of others, to see what that mechanism does on its own. Of course, that mechanism is not *literally* 'isolated'. There cannot be needles or magnetic fields without a lot of other things as well. But we can know (fallibly, of course, like all knowledge) that other mechanisms are not interfering; we can *neutralize* the effect of other mechanisms, either by the way the experiment is set up (as in the present example), or, where a known mechanism other than the one to be tested is unavoidably present, we may be able to determine in what way and how much it is affecting the outcome, and make allowances. Bhaskar sums up the nature of experiment as

an attempt to trigger or unleash a single kind of mechanism or process in relative isolation, free from the interfering flux of the open world, so as to observe its detailed workings or record its characteristic mode of effect and/or to test some hypothesis about them. (SRHE, p. 35).

Where an experiment has been so set up that one mechanism alone operates, we have a *closed system*. In fact, no system in our universe is ever perfectly closed, but experiments can approximate close enough to closure for the purposes of science. It is a characteristic of closed systems that in them a given causal stimulus will always produce the same effect: experiments are

repeatable. Where a genuine causal mechanism has been isolated as a closed system, we can say 'every time A occurs, B follows', as in Humean causality. But in open systems — i.e. almost everywhere outside carefully set up experimental conditions — nothing of the kind occurs. The events that we can ordinarily observe are not invariably preceded or followed by any other constantly conjoined event. Red sky at night is not always followed by a fine day, or deflationary budgets by reductions of inflation, or burglars entering by dogs barking, or spots on the sun by war, or sexual intercourse by conception. If we level off the distinction between ordinary observation and experiment, and retain a Humean definition of causation as regular succession, we will discover no causal laws outside astronomy, where the incapacity of other mechanisms to deflect heavenly bodies from their courses approximates to a (unique) natural closure.

Experiments, then, are necessary because closure does not in general occur naturally. We need to produce 'unnatural' sequences of events in order to discover the mechanisms at work in natural ones. This is the point of Bacon's reference to experiments, not only as questions put to nature, but as 'putting nature to the question'; this metaphor refers to judicial torture, and some moderns have objected that this expresses an attitude of cruelty, and, moreover, since Bacon like many others refers to nature in the feminine, of misogyny. But of course, nature is *not* a woman, or a goddess, or a man, or an animal. It has no feelings, intentions or desires. So the concept of cruelty is inapplicable here; the metaphor of torture cannot be extended beyond its precise function: to indicate that it is not possible to discover the laws of nature by passive observation, one must intervene actively and make nature do what it would not do spontaneously. When R.D. Laing protests against the Baconian project of science by asking 'whether torture is the best way to get to know a lady?' (*The Voice of Experience*, p. 21n), he is extending the metaphor inapplicably, like one who asks whether the Marxian superstructure is safe from lightning, or whether magnetic fields are grazed by rabbits. But while the moral pathos of the question is misplaced, a serious point remains.

The point is this: how can experiments inform us about nature when they are very special processes produced by us, in which things happen differently from the way they do in the open systems of the world outside the laboratory? What if experimental results only tell us what happens under experimental conditions? If they don't tell us how things happen in the open systems of nature at all, then they lack all epistemic value, and are no more than interesting tricks. I have heard an eminent scientist argue that this is just how the ancient Greeks would have regarded them — as telling us no more about the real tendencies of things than the tricks of a circus animal tell us about the real tendencies of its species. This cannot be the whole truth about the Greeks, since they did after all make some experiments. But it helps us to understand the nature of experiments if we can recover a sense of the strangeness of this idea, now long familiar to us, that active interference in the course of nature is more informative about just that course than observation of it is. This idea is at least as paradoxical as Freud's view (which I would also defend) that the study of the pathological workings of the mind is the clue to the understanding of its normal working.

But a paradox is not a contradiction. The whole purpose of experiment is to *isolate* some mechanism which normally operates alongside others. In its normal operation, it has effects: it makes different things happen from what would have happened in its absence. But since what happens in an open system is the effect of a conjunction of forces, it is not what one would have predicted from any of those forces taken in isolation.

Once we can isolate mechanisms and test their effects in closed systems, we can sometimes use the knowledge thus obtained to predict the effects (other things being equal) of their conjoint operation with other known mechanisms in open systems. Our success in doing so shows that experiments are no mere tricks: we make experiments in order to find out what goes on when we are not making experiments, and we do find it out.

This point is sometimes overlooked; I think that some of Paul Davies's arguments, if I understand them correctly, overlook it. He tells us that 'reality triggered only by observation . . . must apparently be accepted on the experimental evidence' (*God and*

the New Physics, pp. 106–7). I wonder what interest such 'reality' could possibly have from the experimental point of view, any more than 'spots before the eyes', which are undoubtedly a sort of reality triggered only by observation. If any mechanism involved in an experiment *were* found to be present only in experimental situations, it would at once cease to be scientifically significant — or at least, its significance would only be as a possible source of experimental error, like a dirty slide or a closed mind. It would not tell us what we want to know — i.e. how things work when we are not experimenting on them. Norman's experiment identifies the mechanism which explains why fixed compass needles point north. Knowledge gained from boiling distilled water at sea-level can tell us about the quite different behaviour of sulphurous water in the hot springs of up-country Rotorua.

Because the mechanisms discovered by experiment, while they affect outcomes in open systems, don't get it all their own way, we need to distinguish various fault-lines between what we experience and how nature is really structured, which break up the unity implicit in the phrase 'the empirical world'. In the first place, at the most minimal level of realism, Berkeley is wrong: things can exist and events can occur unperceived by us. So much would be granted by empirical realists. But the above account of experiment suggests two other fault-lines in 'the empirical world'. For that phrase suggests, on the one hand, a world *defined* by its relation to our experience, and, on the other, the only world there is (or at least the only one accessible to us). Encapsulated in the phrase 'the empirical world' is the licence to reduce questions about what there is (ontological questions) to questions about what we can know (epistemic questions). I shall have more to say about this *epistemic fallacy* in the following chapter. But to return to the fault-lines that vitiate the concept of the empirical world: in open systems, mechanisms operate and have effects other than those they would have in experimental situations, due to the codetermination of these systems by other mechanisms. That is just what makes such systems open, and experiment necessary. And finally, natural mechanisms may exist while they are not operating at all. An experiment (like a natural event) may make them operate when they were not operating before, but that is not the same as making them come

into being where they did not exist before, otherwise the experiment would be no discovery, but an invention. Hence things have unexercised powers, and powers that are exercised unrealized, and powers that are realized unperceived. Unperceived events are unproblematic to anyone but a Berkeleian, but it might be helpful to give examples of the other two: a batsman may have the power to hit the ball to the boundary; this power is unexercised while he is sitting in the pavilion; it is exercised unrealized when his fine cover drive is brilliantly fielded. These distinctions bear witness to the first kind of depth to be discussed in the next section. But first I shall consider some possible objections.

It might be thought that I have chosen examples which support my case, but that there are other examples of scientific procedure which would lead to different conclusions, marginalizing the role that I have attributed to experiment. First of all, it might be said that experiment does sometimes alter the realities experimented on in ways which mislead science, and have to be corrected by evidence of a purely observational, non-experimental kind — by observing open systems, not setting up closed ones. Animal ethology might be cited as the classic example of this. If you study the behaviour of animals in a zoo, you discover only how they respond to unnatural conditions of life. Only when carefully concealed observers (people or cameras) began to record how animals behaved when unaffected by human intervention were misconceptions derived from zoo studies corrected. For the same sort of reason, it might be argued, a keenly observant realistic novelist will have far more to tell us about human behaviour than an experimental psychologist.

On the misleading nature of ethology based only on zoo studies, this objection is of course quite right. But this does not tell against the nature of experiment outlined above. For while the similarity of the laboratory to the zoo is obvious, it is also superficial. Both are artificial situations set up for the benefit of science. But it is not artificiality that makes an experiment, it is closure. If the artificial situation fails to establish an approximation to a closed system, it does not have the significance of an experiment. And the artificiality of the zoo, so far from eliminating irrelevant variables and allowing a single

natural mechanism to be actualized, *introduces* an irrelevant variable with devastating effect on the subject matter, i.e. on the behaviour of the animals. To establish the closest possible approximation to closure in this case, the first necessity is to eliminate precisely this factor: the interference of humans into the animals' world. The sort of artifice which makes for a good experiment in animal ethology will therefore be just such things as the careful concealment of observers of animals in their natural habitat. On the issue between novelists and experimental psychologists, I shall for the time being hold my peace.

There is another criticism of the crucial role I have given to experiments. The work of scientific discovery in developed science, it may be said, is primarily a work of *theory*. The researcher (in modern physics, for example) spends much more time working on mathematical formulae than inspecting experimental equipment. The raw material of this theoretical work is already existing knowledge, including doubtless the results of past experiments, but these are made to yield knowledge that they did not when they were devised, since the concepts needed for their (new) interpretation were not then available. And sometimes, the product of such purely theoretical work can gain the acceptance of the scientific community — and rationally so — even without new experimental verification. When experiments *are* required, the reason why they are can only be given by the theory itself; they are *internal* to the theoretical practice of the scientists. Thus Althusser argues that the difference between Priestley and Lavoisier or Ricardo and Marx was not in their investigation of their subject-matter, but in their mode of theoretical work; in each pair, the latter made advances over the former, not by devising new experiments, but by inventing new concepts (see *Reading Capital*, pp. 149–55).

It is not often noticed that this 'theoreticist' account of science makes science out to be more like ordinary pre-scientific knowledge than the experiment-oriented account does. Experiments — or, more generally, practices specifically designed for the acquisition and testing of knowledge — play a relatively marginal role in everyday, pre-scientific knowledge. We acquire everyday knowledge largely in the course of

activities whose aim is not knowledge. Like Picasso, we don't seek, we find. But the experience acquired in such practically oriented interaction with the world around us is then sifted, criticized, ordered, explained, redescribed. I am not talking about any unusual or consciously 'philosophical' self-examination. I am talking about virtually everything that we normally call 'thinking'. Without such thinking, experience teaches us very little. But thinking works on already acquired experience.

Spinoza, whose work is *one* of the sources of Althusser's earlier philosophy (which in places, if not consistently, exemplifies the theoreticism that I am discussing), is primarily concerned with the knowledge involved in morality — i.e. not scientific knowledge at all, but knowledge of oneself and others and our common life-world. About such knowledge, this account is surely correct. This is 'rationalism', not in the sense of constructing models of the world independently of experience, but of recognizing that 'random experience' and the 'association of ideas' based on it are as much the source of error as of truth, and that the critical work of reason must winnow them before we can get reliable ideas. In Bacon's words:

> The Empirics are like ants; they gather and consume. The Rationalists are spiders spinning webs out of themselves. But the bee combines both functions. It gathers its material from flowers of garden and field, and digests and transforms them by a faculty of its own. This is the type of true philosophy. ('Thoughts and Conclusions', in Farrington's *The Philosophy of Francis Bacon*, p. 97)

Though Bacon is often regarded as the king of the ants, and Spinoza of the spiders, both alike are on the side of the bees.

But now we must ask to what extent this model of gathering first, transforming by thought afterwards, applies to *scientific* knowledge. In the first place, scientific knowledge, like some but not most pre-scientific knowledge, is the result of practices designed specifically for the purpose of producing knowledge. It seeks, and finds only because it seeks. It seeks by means of 'putting questions to nature', i.e. of so setting up a sequence of interaction with nature that the outcome of the sequence will be X if nature is one way, Y if nature is another. But the sequence of

events that yields such an answer does not *have to* occur *after* the asking of the question. It is just as good if it has already happened, whether as an experiment intended to ask another question, or as an action or event with no cognitive purpose. A new theory can, so to speak, retroactively confer the status of experiment on a past event, by asking a new question through it. If we are lucky enough to have an already documented experiment to test our theory, that is in no way epistemically inferior to making a new one. But in many cases we haven't. Ordinary interaction with nature takes place in open systems, and therefore can rarely be made to answer *one* question; it would usually yield a disjunction of possible answers, due to the multiplicity of the processes involved, variations in any of which might affect the outcome. And experiments designed to put *other* questions to nature will answer a new question only if the two questions are closely related, or by accident. So sometimes, at least, the classic sequence of experimental science is necessary: first we construct a theory, then we design an experiment to test it, then we receive nature's answer to our question.

Does this classic experimental sequence have any precedence over observation without intervention, on the one hand, and 'theoretical practice', on the other? I think it does, in that it brings out what is crucial in the other two cases, but less obvious in them. It shows us by analogy just what it is that is so important about watching animals in the wild: the elimination of the irrelevant variable, captivity. It also shows us what is going on in a really knowledge-bearing theoretical practice — for obviously enough, anyone can produce elegant formulae of no cognitive value. Without the experimental paradigm, and the analysis of observation and theoretical practice showing their common features with it, it would be easy to infer empiricist, antlike conclusions from instances of observation, and spiderishly theoreticist ones from theoretical practice.

The analysis of experiment, then, has a crucial illustrative role, throwing light also on areas of science where experiments are rare, inconclusive, or even impossible. Roy Bhaskar is at least as interested in those areas — which include all the human sciences — as in experimental science. It is probably true to say that his work has been more influential on people engaged in the human

sciences than in any other disciplines. But without the results of his analysis of experimental activity, that part of his work could not have got started.

This can be illustrated by an example that has features of both the observational and the theoreticist models of the production of knowledge. I refer to Marx's method in *Capital*.

In the preface to *Capital Vol. 1*, Marx tells us that 'in the analysis of economic forms neither microscopes nor chemical reagents are of assistance. The power of abstraction must replace both' (p. 90). This may appear theoreticist, but of course the 'power of abstraction' works on a mass of empirical material: some of it information that anyone who has lived in a capitalist society will be familiar with, some familiar to those running capitalist concerns (Engels's experience as a factory manager and jobber on the Manchester Exchange was doubtless of value here), some familiar to workers struggling against its tendencies (consider the chapter on the working day); much was derived from the famous reports of the factory inspectors whom Marx praised so highly, and from many other historical records. Years of research had gone into accumulating the empirical raw material of *Capital*. *None* of it is experimental. By virtue of the work of abstraction, Marx is able to put questions to this mass of empirical data in terms of precisely defined concepts, and to make it answer them, in some cases, as if they were experimental in nature. Thus Marx's entirely theoretical argument against 'Senior's "Last Hour"' (pp. 333–8) is tested by the Ten Hours Act.

In addition to this, just as a physicist, when it is impossible to make 'experiments under conditions which ensure that the process will occur in its pure state' (p. 90), will observe the processes 'where they occur in their most significant form, and are least affected by disturbing influences' (p. 90), so Marx gives special significance to 'Branches of English Industry without Legal Limits to Exploitation' (section heading, p. 353) — observing capitalism, so to speak, in the wild.

Methodologically, the strength of Marx's work is that he had learnt something about the nature of science (including the human sciences) from the role of experiments in the natural sciences, but did not imagine that experiments could be made (i.e. closure established) in the human world.

Three Kinds of Depth

I have referred to the type of realism defended by Bhaskar as 'depth realism'. This now requires a fuller account. There are in fact three related ways in which Bhaskar's transcendental arguments establish a depth to reality.

Three domains

The first concerns the distinction I have already discussed between powers and their exercise, and the consequent fault-lines in 'the empirical world'. Berkeleian empiricism, once Hume has exorcized 'spirits' from it, has a one-level notion of what there is: there are experiences (interpreted as impressions, sense-data). Berkeley abolishes the things experienced, Hume the subjects experiencing them. But even such everyday notions as losing something, finding something, or wondering whether one was mistaken about something are very hard to sustain on this account, and most empiricists assume another level, tacitly or explicitly distinguishable from experiences: a level of things and/or events, with events usually foregrounded as the terms of (empiricist theories of) causality. Events can occur unexperienced, can be inferred from their effects, and so on.

But if events are caused by the powers of things — powers that exist even when they are not causing events — then we need to recognize a third level of reality. In commonsense terms, we recognize that, for instance, a motor bike may have the power to travel at 100 m.p.h., even if its careful and law-abiding rider will never make it do so. We may know about this power from what the bike has done in the past, or what other bikes like it have done, though it has not; but we may know enough about the structure in which its working parts are organized, etc. to know that it can do a ton without testing it. This kind of knowledge, predicting powers from structures, is of a more advanced kind, and often presupposes a high level of science (though one can think of simple instances, as when a child can sometimes foresee the uses or dangers of an object at first inspection, without prior experience of similar objects). A good deal of technological research is aimed at knowing how something will work before it is made. The ideal of applied science is that the proof of the pudding should be in the dietitian's analysis. This ideal may never be fully realized — practical tests may often reveal the unexpected. But if we had not travelled way beyond the stage of 'suck it and see', there would be many times more aeronautical and pharmaceutical disasters than there are.

Things have the powers that they do because of their structures, then, and we can investigate the structures that generate the powers, and to an extent predict the powers from the structures. Structures cause powers to be exercised given some input, some 'efficient cause', e.g. the match lights when you strike it. In asking about the structure generating some power of some entity, we are asking about a *mechanism* generating an *event*. A mechanism in this sense is not necessarily mechanical in the sense of Newtonian mechanics. It could be an animal instinct, an economic tendency, a syntactic structure, a Freudian 'defence-mechanism'. The term 'mechanism' has a useful disambiguating function in the philosophy of science. When we talk about 'scientific laws', 'laws of nature', 'laws of history', etc. we may be referring to formulations in words or symbols, which constitute part of the discourse of a science; or we may be referring to that feature of nature which makes such a formulation true. (This is an instance of what C.S. Lewis calls 'the methodological idiom' [*Studies in Words*, p. 20], whereby the name of an intellectual discipline comes to be used for that which it is about: hence 'Freud's psychology' can refer either to his theories or to his personal traits. A *mechanism*, in Bhaskar's sense, is that to which a law refers.

Now it is well known that the laws of nature are, in one sense, very far from 'laws that never shall be broken'. Nothing ever behaves as a law of nature says it should, since in open systems other laws are operating as well. It does not bother us that Macavity 'breaks the law of gravity', since we know that live cats have other powers than those generated by their weight and the earth's gravity. For a law to be true, it must hold when the mechanism it designates works unimpeded — i.e. in a closed system. And for a law to be useful, it must contribute to explaining events in open systems in which that mechanism is operating alongside others.

In open systems, then, a multiplicity of mechanisms is operating, conjointly bringing about a series of events, which

Table 2.1

	Domain of Real	Domain of Actual	Domain of Empirical
Mechanism	✓		
Events	✓	✓	
Experiences	✓	✓	✓

Source: RTS, p. 13.

would not have been brought about by any proper subset of those mechanisms. The series of events that occurs can be called *the Actual* (though this usage is nearer to the continental sense of the term than to the more familiar English ones). But the mechanisms that codetermine it are just as *real*, even though none of them is ever perfectly exemplified by the events. Bhaskar sums this up in Table 2.1.

'The Empirical' (its claim to the *world* has been dropped) is comprised only of experiences; not all events are experienced; the Actual consists of events and experiences, but *mechanisms*, insofar as they are not realized, do not belong here; nevertheless they are real. Theories which relegate mechanisms to a lower ontological league, as 'theoretical entities', 'logical constructs', etc., are refusing to allow causal criteria for reality — i.e. they will only let something through the ontological customs office if it is a possible object of experience. Yet *within* the level of the Actual we are employing causal criteria all the time, and would never get out of the Empirical if we did not: when we find the garden muddy in the morning, we assume a real rainstorm, though we slept through it; a murder-victim implies a murderer, even though one might never be identified. Rainstorms and murderers are possible objects of experience, but their existence is in these cases asserted on causal criteria only, since they are not 'experienced' in the sense of perceived. Why should we not likewise allow that mechanisms are real, though unperceived? Furthermore, the barrier between causal and perceptual criteria cannot mark a frontier between different kinds of being, since (a) things that were once only 'theoretical entities' are sometimes later discovered perceptually, and (b) we can in various ways extend our sense organs (microscopes, etc.), and what we then *see*, we see only granted a causal account of the working of the microscope.

It might still be alleged that since mechanisms can *only* be evidenced by causal criteria, they should not be given the same reality-status as things that could in principle be perceived. This takes us back to the account of experimental closure. For just as we can — and do all the time — bring non-empirical denizens of the Actual into the domain of the Empirical (in short, we experience events), so we can (experimentally) *actualize mechanisms*. As Bhaskar puts it, '$D_r \geq D_a \geq D_e$' (i.e. the domain of the real is greater than or equal to the domain of the actual, which is greater than or equal to the domain of the empirical); and 'the special case $D_r = D_a = D_e$, assumed to be spontaneously satisfied by empirical realism, has in fact to be worked for in the social activity of science' (RTS, p. 229). That is to say, we *can* set up a situation in which the three domains coincide — in which a mechanism is actualized, i.e. is isolated from its usual codeterminants, so that it can operate as a closed system, and be manifested as an event exemplifying the law to which it corresponds. And such an experiment, of course, will be observed. So within this highly circumscribed situation, the three domains coincide. Experiments are windows on to the world of underlying mechanisms which usually operate unactualized.

Note that Bhaskar remarks that empirical realism assumes this coinciding to be spontaneously realized. That is to say, the empirical realist denies that there are any underlying mechanisms, yet also postulates laws of regular succession; these must then be thought to be justified by the pattern of events at the level of the Actual. Sequences that in fact normally only occur under conditions of experimental closure have to be supposed to occur spontaneously if causality is to be justified without recourse to underlying mechanisms. So the distinction of the three domains has to be assumed if the possibility and necessity of experiments is to be accounted for.

Multiple strata

One consequence of the argument from experiment is that there is a multiplicity of mechanisms in nature. If there were a single mechanism only, there would be a naturally closed system, and

passive observation would be enough to establish laws (or *the law*) of nature (except that in such a world, there could be no human observers). Only because nature is an open system are experiments necessary. But since they are also *possible*, mechanisms must be real and distinct, not just schemes imposed by us on a 'buzzing and booming confusion'; for the mechanisms (or some of them) can be *isolated* in experimentally established closed systems. Nature is neither a closed system nor just one damned thing after another, it is a multiplicity of mechanisms jointly producing the course of events. So the course of events is in principle explicable, but not in terms of any one science.

So far, I have referred to multiple mechanisms, but Bhaskar also refers to *strata*. That is to say, these mechanisms are, so to speak, *layers* of nature, and are *ordered*, not just jumbled up together. This will become clearer against the background of a discussion of some features of the 'scientific world-view'. This view, as I shall describe it, is a set of very general conclusions from the *results* of the sciences, not a philosophical argument from their practice, such as Bhaskar provides. But the familiarity of the world-view will make it easier to situate the philosophical argument.

It appears that the material universe existed before there was organic life, and that living organisms can only exist as composed of and surrounded by matter. In this sense, matter may be said to be more 'basic' than life; life in turn may be said to be more basic than rationality (in the sense that we are rational animals), and hence than human society and its history. This suggests that the sciences that explain a more basic layer may have some explanatory primacy over those explaining a less basic layer. Laws of physics and chemistry may *in some sense* explain the laws of biology. There are important disagreements about what sense. Some have speculated that a fully developed science of matter could explain *everything*, so that the laws of biology (and likewise of 'higher'-level sciences like economics or psychology) would be redundant. According to this view (which may be called *reductive materialism*), the less basic sciences exist only because of the undeveloped state of the more basic sciences; its ideal is a single science of matter. As against this view, others have argued that though the more basic sciences may explain something about the mechanisms of the less basic ones, they cannot explain them *away*. The laws of biology are irreducible to the laws of chemistry, even though chemistry may tell us why, for example, DNA molecules replicate themselves. Such views may be called *emergence theories*. As we shall see at length in chapter 4, Bhaskar's theory is an emergence theory, not a reductive materialist one.

One word of clarification is necessary before moving on to discuss Bhaskar's theory of stratification. It is tempting to think that the mineral kingdom is governed by the laws of physics and chemistry, the vegetable and animal kingdoms by the laws of botany and zoology, and so on. But in the first place, animals do not break the laws of physics and chemistry. They are after all composed of atoms, and those atoms obey the same laws whether or not they are parts of living organisms. So animals are necessarily governed by *both* kinds of law, physico-chemical and biological. Minerals, while not *governed* by biological laws, are nevertheless *affected* by them. To explain what happens to stones in the garden, one must know about the habits of ants; to explain the damage to the ozone layer, one must know about the laws of economics. In discussing the stratification of nature, one must keep it in mind that it is *mechanisms*, not things or events, that are stratified. As Bhaskar puts it: 'the predicates "natural", "social"', "human", "physical", "chemical", "aero-dynamical", "biological", "economic", etc. ought not to be regarded as differentiating distinct kinds of events, but as differentiating distinct kinds of *mechanisms*' (RTS, p. 119). And hence also, not as distinct kinds of thing. There is a common tendency, both in everyday discourse and in theory, to commit what has been called the fallacy of misplaced concreteness: to treat as if it were a kind of concrete thing or event or activity or institution what is in fact a kind of mechanism. Thus it is commonly thought that only certain kinds of substance are 'chemicals', and that there aren't any in natural foodstuffs; that certain human needs are 'biological' while others are 'social'; that certain social institutions are 'economic', others 'political' and others 'ideological'. If these last terms, for instance, are treated instead as applying to mechanisms, all of which may govern any particular institution and codetermine its activities, a lot of mistakes can be avoided.

If we say that the laws of chemistry explain the laws of biology, we are *not* saying that chemical mechanisms are somehow more causally effective, that they outweigh biological ones. The proportion in which different mechanisms contribute to the course of events will vary from case to case. It can only be discovered empirically, by examining the concrete conjuncture in each case; it can't be determined by any theory about the stratification of nature and the consequent ordering of the sciences.

So it is possible to distinguish *horizontal explanation* (the explanation of events by mechanisms and antecedent causes) and *vertical explanation* (the explanation of one mechanism by another, more basic one). I have argued elsewhere (in *Scientific Realism and Socialist Thought*) that the most fruitful interpretation of Marx's 'base/superstructure' model of society, with its hypothesis of the explanatory primacy of the economic over the political and ideological, is as a thesis about vertical explanation. Economic mechanisms explain political and ideological mechanisms — but economic mechanisms do not explain all historical events. Economic, political and ideological mechanisms all contribute, in no fixed proportion, to such explanation (as of course do mechanisms outside the social sphere altogether — geography, meteorology, etc.).

Now I come to Roy Bhaskar's argument for the stratification of nature, i.e. for an ordered series of generative mechanisms, in which the lower explain without replacing the higher. It is characteristic of science that the explanatory quest does not come to an end. When one mechanism has been identified and described, and shown to explain various phenomena, it becomes itself something to be explained. Bhaskar explains and exemplifies this clearly in the following passage:

> Thus the observable reactions of chemistry, which are represented in the textbooks by formula[e] such as 2Na + 2HCl = 2NaCl + H$_2$, are explained by reference to the atomic hypothesis and the theory of valency and chemical bonding. The patterns which constitute the explananda of the theory of valency are needless to say by no means superficially obvious or readily available. Both the concepts and the substances and conditions had and have to be worked for, produced in the social activity of science. The theory itself sets out to describe the causal mechanisms responsible for the overt behaviour of the substances. Once its reality has been established (which justifies our assuming that chemical bonding occurs and the laws of chemistry hold outside the laboratory) and the consequences of the theory have been fully explored, the next task consists in the discovery of the mechanisms responsible for chemical bonding and valency. This has been explained in terms of the electronic theory of atomic structure. Once the reality of this explanation has been established, science moves on to the discovery of the mechanisms responsible for what happens in the sub-atomic microcosm of electrons, protons, and neutrons; and we now have various theories of sub-atomic structure. The historical development of chemistry may thus be represented by the following schema:

Stratum I	$2Na + 2HCl = 2NaCl + H_2$	
	explained by	
Stratum II	theory of atomic number and valency	*Mechanism 1*
	explained by	
Stratum III	theory of electrons and atomic structure	*Mechanism 2*
	explained by	
Stratum IV	[competing theories of sub-atomic structure]	[*Mechanism 3*]

> It should be noted that the historical order of the development of our knowledge of strata is opposite to the causal order of their dependence in being. No end to this process of the successive discovery and description of ever new and deeper, and explanatorily more basic, strata can be envisaged. (RTS, pp. 168–9)

Thus the progress of science is a process of *deepening* our knowledge of nature. Underlying each mechanism there are others which explain it waiting to be discovered. The metaphor of 'digging deeper' suggests that we reach the upper layers first, and Bhaskar has said as much in the penultimate sentence quoted. This does not mean that, as a general fact about the history of the sciences, an upper stratum must have been opened up by science before a more basic one can be. That would seem quite implausible in the light of the widely held view, which Bhaskar seems to share, that the natural sciences are in a historically more advanced state (and not merely more susceptible to rigorous testing) than the human sciences. The

point is rather that in order to explain one (upper) mechanism by another (more basic) one, we first need to discover the upper one. We can't have an explanation until we know what is to be explained. We can't predict the upper mechanism from the lower.

Vertical explanation, like any explanation, requires two terms: that which is to be explained (the explanandum) and that which explains it (the explanans). Hence the explanation of the upper by the lower mechanism does not explain away the upper; the discovery of Mechanism 2 in the above schema does not expose Mechanism 1 as mere appearance. We are left with a permanent ordered multiplicity of sciences, a 'tree' with distinct roots and branches, reflecting the real stratification of natural mechanisms, within and between the objects of the various sciences. The grounds for and implications of this emergence theory will be discussed in chapter 4.

The idea of scientific progress as the deepening knowledge of stratified nature serves as a rough marker of the distinction of transcendental realism from empiricism, on the one hand, and relativistic theories of the object of science, on the other. For each of these has its guiding metaphor too: for empiricism, science *collects* discrete bits of knowledge and accumulates them in its mental bucket; for relativism, scientific changes are like *gestalt switches*, 'coming to see the world differently'. Both these metaphors have their place, but if transcendental realism is right, the metaphor of digging deeper catches far more essential features of the process.

The two dimensions of science

If transcendental realism is depth realism in that it looks beneath the course of events to the mechanisms that generate it, and beneath each layer of mechanisms to the one that founds it, it also recognizes a depth dimension in the *object* of science. As realism, it recognizes that science is *about* something, and about something that exists independently of the science; and as *fallibilist*, it recognizes that the science of any given time can be *wrong* about its object. The 'results' of scientific inquiry at any time are a set of theories about the nature of the world, which

are presumably our best approximation to truth about the world. But the *work* of science at any time takes these theories as its *raw material*, and seeks to transform them into deeper knowledge of the world. These theories are its *transitive object*; that it seeks to transform them shows that its aim is knowledge of its intransitive object, the world that exists independently of it. However much science deepens its knowledge of its *intransitive object*, its product remains a transitive object. This last point enables Bhaskar to allow quite a lot of scope for 'the sociology of knowledge', explanations of scientific results as produced by mechanisms quite extraneous to the project of our deepening our knowledge of nature. But it saves his theory from the ontological relativism that is often inferred from such social studies of science. Rival scientific theories necessarily have different transitive objects, or they would not be different; but they are not about different worlds — otherwise how could they be rivals? They would not be scientific theories at all if they were not aimed at deepening our knowledge of the intransitive object of science. To discuss this more fully, we must move on to a general consideration of Bhaskar's account of the work of science.

The Work of Science

We have seen that to bring about the identity, within a limited time and place, of the domains of the Empirical, the Actual and the Real — to actualize and observe the workings of some underlying mechanism — is a *work of science*. It is an active intervention into nature, made by people with acquired scientific skills, usually using special equipment. It is *work*, not contemplation, not observation, not the taking up of some special kind of scientific 'attitude' (even though it may involve these things, and indeed must involve observation, as *aspects* of the work process). Moreover, it is not the antlike work of collection, nor is it spiderish creation out of the scientist's own mind. But here the metaphor of the bee, which works well enough in contrast to the ant and the spider, begins to break down. For, as Marx said of architects, scientists are

distinguished from bees by the imaginative foresight they bring to their work.

But the production of experimental sequences of events is clearly not an end in itself. The 'product' is not the new arrangement of matter brought about by the experiment, for instance the chemical which has been synthesized, or the reading on a measuring device. It is the deepened knowledge of some mechanism of nature.

Hence scientific work in general (not only its experiments) has the character of *production*, i.e. of the *transformation* of raw materials into (provisionally) finished products, using means of production that are themselves products (as are most of the raw materials), and employing special skills. The transitive object, the existing state of the scientific knowledge, forms the raw material which is to be transformed into a new theory yielding deeper knowledge.

This looks, and is, much like Althusser's theory of 'theoretical practice'. Indeed, when it is recalled that Althusser includes experimental equipment and techniques as part of 'theory', and that Bhaskar himself says that 'science as a process is always entirely intrinsic to "thought" ' (RTS p. 185), it might look as if the two theories are identical. However, there are certain subtle but crucial differences.

First, Althusser says almost nothing about the relation between theories produced by a science, the 'object in thought', or 'object of knowledge', and what they are about, the 'real object'. He says a lot about the sort of answers to this question that *won't do*.[1] But he ends up leaving us looking for the mechanism that brings it about that its product is knowledge, on analogy with the mechanisms of social reproduction that bring it about that what is reproduced is a society; but society is not society by virtue of its relation to some one thing that is not itself society; knowledge is knowledge by virtue of its relation to its real object. We need to know what it is about the process of production which ensures that the product is knowledge *of its object*. The causal interaction with the real object, which takes place in experiment as described by Roy Bhaskar, might provide the basis of an answer. Of course, it is not just any old causal interaction. It is one which, on the basis of existing knowledge, we have grounds for regarding as informative about the real

structure of things. But we cannot regard experiments, as Althusser seems to, as just one device among others whereby science checks its theories. Or rather, insofar as we are forced to do without experiments in some sciences, science is only possible where there are procedures that have at least this in common with experiments: that we have good reason to suppose that if one possible outcome occurs, one postulated mechanism must be real; if another outcome, another mechanism. In other words, we must be able to put questions to nature, and get replies that were not already implicit in the questions themselves.

Furthermore, the concept of a deepening knowledge of stratified nature is absent from Althusser. This is shown by his metaphor of the 'continents' of science — mathematics, physics and history — which allows neither for the irreducibility of one science to another within a continent (for instance, of biology to chemistry within the continent of physics), nor the foundedness of one continent on another (for instance, history on nature). I have criticized this metaphor elsewhere (in *Scientific Realism and Socialist Thought*), so here I shall only add some remarks about the relations between one pair of Althusser's concepts and one pair of Bhaskar's. Insofar as the 'transitive object' refers to the state of scientific knowledge at any time, and the 'intransitive object' to the object which exists independently of the science, which the science is about, they appear to be closely parallel to Althusser's 'object in thought' and 'real object'. But whereas Althusser thinks of the object in thought primarily as the *result* of scientific work, Bhaskar thinks of the transitive object primarily as its *raw material*. This might be considered a matter of emphasis, since for both philosophers the product of science at any one time becomes the raw material of the scientific work that follows it. But I think this difference of usage is symptomatic of the fact that Althusser thinks mainly in terms of *one* 'epistemological break' which founds a science, and thereafter of development without sharp breaks; Bhaskar thinks of *successive* deepenings of our knowledge, as new layers are uncovered. Hence for Bhaskar the task of devising and testing new explanations is always before us; for Althusser, it appears that all we need to do once we have got a science is to spell out and apply its original insights. So while the conception of

scientific activity as a production-process is shared, Bhaskarian theoretical practice can never forget about its intransitive object; Althusserian theoretical practice often has done.

Bhaskar's phrase for his description of the scientific work-process is 'the social production of knowledge by means of knowledge'. This echoes the title of Piero Sraffa's book *Production of Commodities by Means of Commodities*. In Sraffa's model of production the input as well as the output consists of commodities: raw materials, means of production, and goods for the sustenance of the work-force. This suggests a model of theoretical production in which the input as well as the output is knowledge, which can also be divided up as raw materials, means of production and, corresponding to goods for the workers' sustenance, *scientific training*.

> The necessity for a scientific training shows that knowledge is a social product and cannot be conceived as a purely individual acquisition. For it always stands to the individual as something that must be acquired to be used (for scientific work). (RTS, p. 187)

For just as nature does not for the most part produce manifestations of its mechanisms in closed systems, it does not produce people spontaneously capable of perceiving and interpreting such manifestations. We not only have to work to make the mechanism appear, we have to work to make ourselves capable of understanding the appearance. Scientific training is to the 'subjective aspect' of scientific work what experiment is to the 'objective aspect'. It *produces* suitable 'knowing subjects'. It does so by induction into the theory and practice of existing science. Hence, to become a scientifically 'knowing subject' is to acquire a *historically specific* set of ideas, techniques and skills; little can be said about 'knowing subjects' at any abstract, historically unspecific level, after the manner of traditional epistemology. Non-scientific forms of knowledge are equally historically specific; our minds are formed by historically specific societies, and that is the only way they can be formed at all. And of course, different societies will inculcate different ideas, practices, etc. Since at any given time we are full of all sorts of ideas which will later turn out to be false, this also means that no one can become a 'knowing subject' without

being suckled on all sorts of *falsehoods* 'at the breast of the universal ethos' (Hegel). But that is no objection to such suckling, without which our minds would not just be 'blank sheets', but destined to stay that way forever. One sometimes encounters a sort of romanticized empiricism which supposes that if only we were not subject to 'indoctrination' or 'conditioning' we would be able to *see* the truth. But those words should only arouse our anger when they are used to mean the intentional misleading of the young. For it is absurd to imply that any society or culture or generation can do better than pass on its own sincerely held beliefs to its successors — who may *then* be able to criticize, correct and improve on those beliefs. A mind unsullied by second-hand prejudices would be a mind incapable of experience of a recognizably human kind, and hence also of 'finding out for itself'. In William Blake's striking metaphor:

> Establishment of Truth depends on destruction of Falsehood continually,
> On Circumcision, not on Virginity, O Reasoners of Albion!
>
> ('Jerusalem', *Complete Writings*, p. 687)

So far as specifically *scientific* knowledge is concerned, everyone will recognize the need for training. However, there is perhaps a residual shadow of the anti-indoctrination fallacy in the over-sharp distinction drawn by Kuhn and his followers between 'normal' and 'revolutionary' science. For Kuhn, scientific revolutions are discontinuous with other scientific development, with training in 'normal science' being inherently conservative, and something of an epistemic obstacle to scientific revolutions. But something may be an obstacle to an activity, yet also a necessary condition of it, like gravity to highjumping. A training in normal science is a necessary condition not only of normal science but of the revolutionary science that overturns it. (It is true that scientific revolutions are sometimes brought to a science from 'outside', but always from another, usually closely related, scientific discipline, and never by thinkers unfamiliar with the science transformed by them.)

To recognize the character of science as 'the social production of knowledge by means of knowledge' is to place it within history, and so to allow that it shares the impurity and questionability of all human history. One could not take seriously a historian of, say, the English, French or Russian Revolution who *either* denied that it was an episode in the story of human emancipation, *or* denied that there were countless accidents among its causes, vices among the motives of its agents, and iniquities among its effects. And the history of the sciences is no more 'the march of God upon Earth' than is the history of states. On the one hand, the sciences do deepen our knowledge of nature, and only because they do so can they be recognized as sciences, as distinct from other activities. But the mechanism of knowledge production which makes a science what it is does not exist in a closed system. Psychological, economic, political and ideological mechanisms codetermine the history of the sciences with it. What does this mean for 'social studies of science' and the criticisms of science that are sometimes based on them?

It might be said: *insofar as* a science has succeeded in deepening its knowledge of its intransitive object, its process of production is irrelevant. This may be admitted, but it has nothing like the force that is sometimes thought by those who dub any criticism of a product in virtue of its means of production 'the genetic fallacy'. For our reasonable confidence that a science *does* give us genuine knowledge is based precisely on the nature of the mechanisms by which that knowledge was produced. To give a non-cognitive analogy: if a beer-taster pronounced the product very fine, and then changed his opinion when he learnt that some non-real-ale techniques were used in its production, one might suspect he was in the grip of the genetic fallacy; but there are some people who would accuse you of the genetic fallacy if you refused to drink a substance which you knew to have been produced by the method usually used for making sulphuric acid.

In order to be relevant to the assessment of the product, a causal study of theoretical production must distinguish the different kinds of mechanism which may have contributed to the process. On the one hand, there are those that belong with the *intrinsic aspect* of science — experiments and other forms of

reality-testing which make the science what it is: an attempt to deepen our knowledge of its intransitive object. On the other hand, there are the ideological bias of, and political or economic pressures on, the scientific community. A study of the latter may reasonably arouse our suspicions about a science, and lead us to double-check its reality-testing procedures and their interpretation. Would racist theories of 'intelligence' ever have been discredited scientifically if it were not for scrutiny motivated by suspicion of their political bias? But on the other hand, any attempt to bypass the intrinsic aspect altogether may lead to rejecting well-founded views, simply because their proponents had motives for wanting them to be true (the error that C.S. Lewis has dubbed 'Bulverism': *First and Second Things*, pp. 13–18).

Bhaskar describes his position as 'epistemic relativism', i.e. relativism about the transitive object, without relativism about the intransitive object. I shall discuss this matter further with respect to the human sciences. For the truth is that genuine ideologically motivated differences about the *content* of scientific theory (as opposed to the ethics and politics of its application) are hard to find in the natural sciences. In the human sciences, though, they are frequent and intractable. Meanwhile, it can be said that Bhaskar's view (1) gives an honourable place to social studies of science, but (2) gives no place to any arbitrary choice or subjective preference in assessing scientific theories, nor of any selection of them purely on moral, political or aesthetic grounds (even though such grounds may not be subjective or arbitrary).

Throughout this discussion of the work of science, it has been assumed that the essential feature of the product of that work is *explanation*. The mechanisms discovered by science explain what happens, deeper mechanisms explain surface ones, and so on. It might be alleged that this neglects *prediction*, which is widely held to be, (a) theoretically speaking, *symmetrical* with explanation, in that if A explains B, B could have been predicted from A; and, (b) practically speaking, the aim of science, without which explanation would lack any utility. Let us examine these claims.

The symmetry of explanation and prediction, like much that empiricism holds to apply in nature generally, in fact applies

only in closed systems. Whatever a mechanism explains could be predicted from that mechanism plus a given input or stimulus, *provided* that other mechanisms were not interfering. But in the world outside the laboratory (and aside from certain astronomical events), that sort of prediction is not available to science. And certain sciences, whose mechanisms operate only in open systems, can achieve high explanatory power without being able to make a single prediction (evolutionary biology, generative grammar). This inability is not a *failure*; it is a theoretically demonstrable feature of the real object of these sciences that explanations of it will not generate predictions. In the case of evolutionary biology, natural selection only comes into play upon random mutations which are not themselves predictable; and no syntactic theory could predict what will be said, which depends on the conversational situation, etc. (There is a sense in which generative grammar is said to predict, but it does not predict spontaneously occurring events, e.g. speech acts. It 'predicts' [synchronically] which sentences of a language will be grammatical. This in turn predicts the results of characteristic linguistic 'experiments', i.e. questions put to native speakers as to whether this or that sentence is acceptable. The use of 'predicts' here is slightly odd, and I suspect is parasitic on the belief that predictions and explanations are symmetrical.)

The asymmetry of explanation and prediction means that extravagant claims for falsifiability as a necessary feature of science are mistaken. Scientific theories are falsifiable in the sense that they can be shown to be false, but not in the sense that any given 'counter-example' will overturn them. Obviously enough we do make some sort of forecasts about what will happen in open systems on the basis of scientific theories, but nothing that happens in an open system will of itself falsify a theory. We take this for granted outside the classroom. If a doctor tells her patient 'you are out of danger', and the patient walks out of the surgery and under a bus, no one thinks the doctor unscientific. In view of this, the Popperian case against Marxism is forced to shift its ground. For if we ask 'is Marxism falsifiable by events in open systems?', the answer is 'no, no theory is'; and if we ask 'is Marxism falsifiable in closed systems', the answer is 'no, because there are no closed systems

in social science — nor yet in biology, meteorology, etc.' But it does not follow that Marxism is without explanatory power.

It is obvious, though, that this does not make the problem go away, for we need accounts of (a) how theories *can* be tested when closure is not attainable (for Popper is surely right that an *untestable* theory would be unscientific), and (b) how theories can be usefully *applied* if they are not predictive. We will encounter these questions again with regard to social science. But first we must reach a clearer understanding of the nature of explanation in open systems. To this end, I shall now discuss the transcendental realist theory of natural necessity.

The Works of Nature

It will by now be clear that transcendental realism involves a notion of natural necessity that is not reducible to regular succession. It agrees with commonsense that to say that A makes B happen is to say more than that A-type events are generally followed by B-type events; so a real difference of interpretation explained our laughter when, during my schooldays, a teacher snapped at us 'every time I open my mouth, some idiot speaks'. Here I want to spell out what the 'extra' is, and in doing so meet the objection that any such extra must be unwarranted by the evidence. For it was this objection which made Hume think that his *repeated conjunction* account of cause was unavoidable, even though he fully recognized its paradoxical character. For it seemed that any other definition of cause had to include a non-empirical element, and therefore an element which could not be empirically justified. Non-empirical, in a sense, the element certainly is, but the 'therefore' does not follow.

Bhaskar's first marker for this element is the word 'power', which is itself empirical enough — it merely indicates what a given kind of thing can do, given the right conditions: dogs can bark, aeroplanes can fly, cricket balls can smash greenhouses, and so on. So far, the objector is going to say, nothing has really been added: how does 'Rectory Ale sends you to sleep because of its dormitive power' differ from 'Rectory Ale sends you to sleep'?

First, while the latter may count as an explanation of my sleepiness, it does not step forward for the role of explanandum. Such explanations induce an intellectual slumber to match the physical one induced by the Rectory Ale. But once dormitive powers are mentioned, we are alerted to the question 'wherein lie these dormitive powers? What is it about the micro-structure of the beer that generates this power?' One person may claim it is the high alcohol content, another may say no, cider is just as strong and much less soporific, it must be the hops. So the chemists get to work, testing the alcohol content and that of the soporific agent from the hops, and pharmaceutical research can tell us about the effects of these substances, separately and conjointly, on the human nervous system.

Thus the stratified nature of explanation, and the dynamic nature of scientific inquiry, making each result the next matter for investigation, give empirically justifiable content to the non-empirical part of causal claims. Effects are ascribed to causal powers, causal powers to the inner structure (and place in larger structures) of the causal agent. The 'extra' in the causal power is just this structure, which is unearthed only by a second stage of investigation, after the identification of causal power. And this structure is, of course, no more an 'unmoved mover', an unexplainable explanation, than the power it explained. We may dig deeper to discover just what it is about C_2H_5OH (alcohol) which reacts on the human organism in the way it does. With this in mind, let us look at what Bhaskar says about recent non-Humean theories.

At the beginning of SRHE, Bhaskar discusses other recent critics of the Humean account of natural necessity, the 'anti-deductivists'[2] who

> have sought to show how scientific practice yields cognitive items — whether dressed as models, paradigms, heuristics, conceptual schemata or regulative ideals — which are irreducible to syntactical operations upon sense-experience and yet indispensable for the intelligibility and empirical extension of theory. In this way such items function, as it were, as social surrogates for natural necessity. (SRHE, p. 3)

That is to say, they show that science in fact sets up explanatory structures which are not reducible to Humean successions; but

these structures are seen as part of the social practice of science rather than as natural structures discovered by that practice; they belong solely to the transitive dimension. The Humean conception of *the world* (as 'the empirical world') is not challenged; and just as for Kant, structures are supplied by us.

Bhaskar argues that such positions are inadequate on three counts:

> First, to the extent that the surrogate can be empirically described, its independent cognitive role disappears . . . conversely, to the extent that its cognitive role is preserved, its epistemic warrant crumbles (since it now ceases to designate real phenomena). (SRHE, p. 3)

This is a criticism which might easily have been made from the opposite direction, i.e. by a Humean: either the model (or whatever) can be cashed in empirical terms, in which case the account collapses back into Humean constant conjunction; or it goes beyond the empirical data — but in that case what empirical grounds can there be for it? But the Humean would think this argument equally good against Bhaskar himself; hence the use of it here implicitly gives a promissory note that an account will be given how theories which go beyond the data can be empirically grounded. I hope the above arguments show that such a note is creditworthy.

The second and third points concern the anomalous asymmetries, on this account, between a scientific theory and the nature it purports to explain: natural necessity must surely exist, if anywhere, in things independent of us, yet it is being presented as supplied entirely by the human mind; and it is strange that we should have to posit structuredness to explain a supposedly unstructured real world. If the real structure of nature, and its consequent necessities, do not make such structured theories essential to their explanation, the structured theories must be more or less gratuitous. We must either retreat to the flatlands of the Humean succession of impressions, or advance to a theory of real structures generating real necessities.

What, then, is the transcendental realist theory of natural necessity? The four concepts that go to make up this theory are *structures, powers, generative mechanisms* and *tendencies*. So far, I

have said least about tendencies; now they must occupy centre stage.

We have seen that things have the powers that they do by virtue of their structures. As the gospel warns us, we can't gather grapes from thorns or figs from thistles, or, in Roy Bhaskar's slightly revised version,

> It is physically impossible for cabinet ministers to bear figs; that is, nothing which bore figs could properly be said to be a cabinet minister at all. (RTS, p. 223)

— despite the evident woodenness of some of that kind.

These structures can be investigated independently of any particular power which they generate; and the structures, and therefore also the powers which they generate, exist whether the powers are being exercised or not. *Generative mechanisms*, Bhaskar says, 'exist as the causal powers of things' (RTS, p. 50). This does not mean that 'generative mechanism' is a redundant equivalent of 'causal power' though. 'Power' is a non-technical term, designating what something can do. 'Generative mechanism' is a technical term, designating 'a "real something" over and above and independent of patterns of events' (RTS, p. 50), which normally endures longer than any pattern of events it generates. A generative mechanism, we might say, is that aspect of the structure of a thing by virtue of which it has a certain power. For example, that aspect of the structure of an oxygen atom by virtue of which it can combine with two hydrogen atoms to form a molecule of water; that aspect of a DNA molecule by virtue of which it can replicate itself; that aspect of a market economy by virtue of which it can go into an overproduction crisis; that aspect of a person's brain-structure by virtue of which he or she can acquire language.

A generative mechanism will operate when suitably triggered. As we have seen, experiment consists in isolating and triggering a generative mechanism so that it will operate unimpeded. But in open systems, generative mechanisms are not isolated; when triggered, they operate, but in conjunction with other generative mechanisms, producing a complexly codetermined outcome. This feature of natural necessity is captured by saying 'causal laws must be analysed as tendencies' (RTS, p. 50). Things tend

to act in certain ways: that is to say, if triggered, a tendency will come into play and have effects, though these effects may not be the ones it would have had in a closed system. While the word 'power' draws attention to the existence of unexercised powers, the word 'tendency' draws attention to the existence of exercised but unrealized tendencies. And the 'course of nature' consists, for the most part, of the interplay of the tendencies of things, exercised but incompletely realized because of their coexistence — oaks tend to grow tall, but not in Beddgelert Forest because of the wet soil; yet their tendency to grow tall is not without effects in Beddgelert Forest — they do get taller than the gorse bushes, and many of them do fall over.

> It is by reference not just to the enduring powers but the unrealized activities or unmanifest (or incompletely manifest) actions of things that the phenomena of the world are explained. It is the idea of continuing activity as distinct from that of enduring power that the concept of tendency is designed to capture. In the concept of tendency, the concept of power is thus literally dynamized or set in motion. (RTS, p. 50).

Now I have so far been following the usual practice of the philosophy of science by saying that powers, once set going, will be realized 'other things being equal'. But this clause is unnecessary with regard to tendency statements. Such a clause

> does not place a condition on explanation, for one can explain an event in terms of tendencies when the latter are never realized. Rather it places a condition on prediction and falsification. (RTS, pp. 96–7)

Explanation in open systems is in terms of tendencies. And in closed systems, presumably, we do not need to say 'other things being equal', as we have taken good care that they are. So

> a fully realist philosophy of science could in principle dispense entirely with the CP [*ceteris paribus*, i.e. other things being equal] clause. . . . For whatever is conveyed by 'This happens CP' can be equally well conveyed by 'This tends to happen'. . . . This is not a shallow, equivocal, sloppy or mean formulation; but the logical form of all the laws of nature known to science. (RTS, p. 97)

A theory of natural necessity as the working of tendencies enables us to avoid two opposite mistakes in talking about cases in which one tendency offsets and neutralizes another. The first is committed by Mill when

> he argues that 'although two or more laws interfere with one another, and apparently frustrate or modify one another's operations, yet in reality all are fulfilled, the collective effect being the exact sum of the causes taken separately'. Mill's mistake here is to suppose that whenever a tendency is set in motion the effect must be in some sense (or in some realm) occurring (as if every time we ran fast we had to be in some way winning). (RTS, p. 99)

The second is committed by Geach in supposing that 'because neither tendency is fulfilled neither tendency can be in play' (RTS, p. 100). Both are wrong.[3] It is not the same thing for something to be stationary because it is at rest, and because it is pulled with equal force in both directions; Buridan's ass, hesitating between two equidistant bales of hay, is not restive for the same reason as a donkey with no desire to go anywhere.[4]

Bhaskar calls tendency-statements *normic*. It should be obvious that he is not using this word to refer to breakable rules, in the sense that logic and ethics are said to be normic (i.e. normative) disciplines. Neither is he using it to refer to probability statements. A tendency may be manifested as a statistical probability, but a normic statement is universal in a way that probability statements are not. A statement such as 'bodies tend to persist in a state of rest or uniform motion in a straight line' has strict universality, even if no body has ever so persisted. A tendency is having effects when it is not manifested as well as when it is; so a normic statement can be true even if the tendency it refers to is never manifested because it is curtailed by offsetting tendencies (perhaps precisely because the tendency is known to, destructive to and frustratable by humans). Finally, a normic statement is not a hypothetical statement, 'if A then B'; it will entail hypothetical statements, but in itself it is a statement about what is actually going on.

Natural necessity, then, may be characterized as: the necessary working (alongside others) of tendencies once they have been triggered. (I postpone discussion of the distinction between tendencies in this sense and two other kinds of
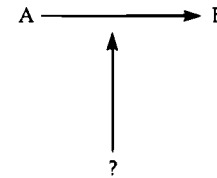
Figure 2.1 Horizontal Explanation, Vertically Unexplained

tendency, i.e. *liabilities* — e.g. paperbacks with 'perfect binding' tend to break in two when opened — and *ontological preference* — e.g. robins tend to eat worms. I believe there are some problems with these distinctions, which I shall discuss in chapter 4.) Just how necessary is natural necessity so conceived, and in what sense?

It is clearly a necessity in the things themselves, not in our judgements about them. It is neither a subjective necessity in us to make such judgements nor a logical necessity in the judgements themselves. It is the necessity that a tendency cannot but work, once the conditions for its working are there: make water hot and it will tend to expand — and if trapped in a sealed container, the tendency will still work, even if the expansion is inhibited.

Some light may be thrown on the relation between the transcendental realist theory of necessity and some classical ones by looking at what Bhaskar calls the Humean, Lockean and Leibnizian levels in the development of a scientific theory (RTS, pp. 171ff). First (the Humean level) some sort of regularity is identified: whenever A then B — though *contra* Hume, this will generally have to be experimentally set up. Also *contra* Hume, 'The scientist never doubts for a moment that something is generating the effect in question. His problem is: what is?' (RTS, p. 172). The state of science at this stage is illustrated in Figure 2.1. The Lockean level is reached when the query is answered. Then we know the structure of some entity that generates this regularity:

> Now it is contingent that x has the nature (e.g. constitution or structure) that it has. But given that it has, it is necessary that it behaves the way it does. (RTS, p. 172).

But in the course of scientific development, the entity whose structure has been discovered may very well come to be *defined* by that structure. It was *discovered* that water = H$_2$O, but we would not now call anything 'water' unless it was H$_2$O. Granted that the molecular structure of water generates various tendencies, it might be asked whether these natural necessities do not turn out, at this Leibnizian level, to be instances of logical necessity.

> At the Leibnizian level statements of law are substitution instances of necessary truths about the individuals to which they refer. (RTS, p. 174).

On this matter I refer the reader to the section of RTS called 'Objections to the Account of Natural Necessity Proposed' (p. 199ff). This section is both longish (seventeen pages) and lucid; it requires no commentary, and to précis it would destroy its clarity. However, I shall make one brief point from it: the logical necessity of a statement and the causal necessity of what it describes are independent questions. Bhaskar says 'Some causal statements expressing necessary connections are logically necessary and some are logically contingent' (RTS, p. 201) and quotes Davidson in support: 'The truth of a causal statement depends on *what* events are described; its status as analytic or synthetic depends on *how* the events are described' ('Actions, Reasons and Causes', p. 90).

Scientific truths are contingent and discovered a posteriori even at the Leibnizian level; they could have been otherwise, and they were not always known. They become 'analytic' because we make them definitions. However, we do not do this arbitrarily *in science*, though we *can* always substitute an analytic for a synthetic statement: 'the American President who betrayed the Iraqi opposition betrayed the Iraqi opposition' is analytic, though it refers to the same fact as 'President Bush betrayed the Iraqi opposition', which is not. But in science, the Leibnizian use of analytic statements is a special one; it is justified only with regard to a particular kind of truth: about inner structures of a thing or kind of things, generating the specific tendencies of that thing or that kind of things. There could be no Leibnizian level to the definition of jade, since the

various stones called 'jade' do not share an inner structure; they share only a cluster of perceptible qualities (colour, smoothness, hardness, etc.) the boundaries of which are drawn on aesthetic/ commercial grounds by dealers in semi-precious stones. But because water, for example, does have a defining inner (molecular) structure, which necessarily generates certain tendencies (e.g. to boil at 100° Celsius), 'water tends to boil at 100° Celsius' can be treated as analytic — i.e. if anything doesn't tend to boil at 100° Celsius, it isn't water.

It is now possible to say a little more about explanation. Clearly explanation does *not* consist in subsuming that which is to be explained under a generalization. If 'all ravens are black' seems to be *some* kind of answer to the question 'why is that bird black?', that is only because it emptily indicates a possible explanation, perhaps by suggesting that genetics may be consulted. A generalization, however strict, which failed to suggest where we might look for *its* explanation would never be considered an explanation ('why is that tea mug dirty?' — 'all the tea mugs in my study are dirty'!). On the other hand, even the loosest generalizations of everyday gossip, which make no claim to exceptionlessness, may indicate the presence of some underlying tendencies, and so be more helpful than the tea mug generalization ('English people eat boiled vegetables', 'referees of academic publications don't understand irony' — presumably there must be some socio-historical explanation for these remarkable facts). However, generalizations only gesture towards the explanatory work, which begins when a mechanism generating a tendency has been located.

It is important to see how this differs from the way some other philosophers of science have accounted for the 'surplus element' over and above generality that makes one generalization explanatory, another accidental. The difference has sometimes been thought to consist in the applicability of a theory in the explanatory case, the absence of such a theory in the accidental case. But what, it must be asked, does the theory contribute that the original generalization did not? Is it just a higher-level generalization? Then the quest for the surplus element is just pushed one step further back: we still want to know how the higher-level generalization differs from an accidental one. Or is it perhaps a *model*? If so, then we can ask if

the model is meant to represent a real mechanism that generates the phenomena about which the generalization has been made. If it is, then we can set about investigating it in other ways; in the course of the dynamic process of scientific development the hypothesized mechanism may be discovered to be real, or not. In this case, 'model' is simply a word for a hypothesized generative mechanism, and we are on transcendental realist ground.

But the model-builder might reply 'my model is only an imaginative construction; it makes things clearer to us, and helps to predict the phenomena; but it makes no sense to go looking for it — it is not that kind of thing'. There may be some peripheral place for such fictions in science, but it should be noted (1) that they don't give us a real surplus element, since nothing is being *postulated* (as opposed to imagined) beyond the original phenomena; in a world without scientists, there would be nothing corresponding to such models; so they can be said not to *explain why* the phenomena are as they are, but only to ease our mental labour. (2) While at a given time a model may have no justification but that it makes given phenomena intelligible to us, in the development of science what were once just such models often come to be discovered as real structures underlying those phenomena, and identified in terms independent of those phenomena. They may even come to be *perceived*, with the aid of new equipment extending our sensory powers. As Bhaskar puts it, 'the hypothetical mechanisms of yesterday may become today's candidates for reality and tomorrow's phenomena' (RTS, p. 159).

It is ultimately Bhaskar's conception of the *development* of science, in which yesterday's explanation becomes what is to be explained, in an ever-deepening stratified account of nature, which warrants — and on thoroughly empirical grounds — going beyond empiricism.

Notes

1. I am referring primarily to Althusser's account in Part I of *Reading Capital*. Althusser deserves great credit for having got the question right here. If he then circles round the question like a cat round a bowl of hot

porridge, I think it is because he has no solution to the dilemma that while experiments are the only mechanism that meet his requirements of internality to the science, and causal power to produce the 'knowledge effect', they do not exist in the science he is most concerned with: historical materialism.

2. Bhaskar lists 'Kneale, Waismann, Hanson, Scriven, Polyani, Toulmin, Hesse and Harré' under this description (SRHE, p. 2).

3. The references are to Mill's *A System of Logic*, Bk. III, chapter 10, section 5; and Geach's 'Aquinas'. For the record, it seems to me that while Geach sometimes commits the mistake identified by Bhaskar (i.e. when he is telling us what Mill should not have said), at others his position (or his reading of Aquinas's) is identical with Bhaskar's. Thus:

A tendency is indeed specifiable, always and exclusively, by describing what happens if the tendency is fulfilled; but not all tendencies do pass to fulfilment, as we readily see if we refuse to muddle ourselves with talk about 'sum of effects', as Mill did. (He was even ready to say that if nothing happens at all, this nothing may be the 'sum' of *actual effects* that are equal and opposite!) We must rather say: Given the natural agents involved, we know their tendencies; given all the tendencies involved, we know what will actually happen. (Thus, given the members of a structure, we know what stresses will be set up; and given all the stresses, we know what deformations will be produced.)' ('Aquinas', p. 103).

A tendency for something to happen is different from its actually happening; but yet a tendency is somehow actual, not a mere potentiality, a 'would happen if'.
Even though the other tendencies involved in a given situation prevent the actual fulfilment of a given tendency, its presence will always make a difference to what actually happens; and the procedure of scientific explanation is to infer natural tendencies from what actually happens, and then predict what will happen from the natural tendencies of the agents believed to be operative.' (ibid, p. 104)

4. The text of RTS has 'Balaam' for 'Buridan'; somehow the talking biblical donkey has been substituted for the indecisive medieval donkey.

3

The Impossibility of Empiricism and Idealism

Machiavelli said: 'Madre di Dio, now I've seen everything', neatly presenting in a single sentence a conceptual impossibility and a denial of the external world.
(Michael Westlake, *One Zero and the Night Controller*, p. 186)

No philosophy exists in a vacuum; there are always particular opposing philosophies which coexist in any historical period, and every philosophy engages, implicitly or explicitly, in controversy with its opponents. Philosophy may seek truth, but it seeks it in an adversarial as well as in an investigative manner. Bhaskar's main contentions are with empiricism and idealism. In the present chapter I shall (1) sketch the historical role of empiricism and the flaws that vitiated it; (2) discuss the *epistemic fallacy* (Bhaskar's phrase), the common fallacy of empiricism and idealism; (3) discuss the nature of idealism as the main modern alternative to empiricism, with particular reference to its twentieth-century forms. These three sections will include longish passages with little explicit reference to texts by Roy Bhaskar, but all presuppose the arguments discussed in the previous chapter. The remaining section, on the other hand, is closely tied to a text by Bhaskar. In it I briefly expound Bhaskar's detailed and intricate critique of one version of empiricism, namely positivism.

The Legacy of Empiricism

Though empiricism has long been the blight of English-speaking philosophy and social science, it is salutary to recall that it was once part of a great liberating movement of thought. This movement had many aspects in different spheres of life and thought, each initially liberating, each flawed in much the same way, and ending as an obstacle to understanding and emancipation. The whole movement may be described as a rejection of the authority of tradition: of established laws and customs, ancient texts, and so on, in favour of turning to 'the great book of the world', and judging for oneself. This attitude, which was later to become one of the features of the Enlightenment, is particularly noticeable in seventeenth-century England. One can see it at work in Baconian science, Leveller politics, Quaker religion. In its extreme forms, it leads to the demand that the slate of received ideas be wiped clean, so that we can start from scratch. In politics, for instance, this gives rise to the idea of a *real* social contract, as in the Levellers and Locke (rather than a hypothetical one as in Hobbes and Rousseau), and hence to the idea that certain individual powers could be 'reserved', kept out of the remit of the body politic.[1]

Now it is easy to see why a 'clean slate' approach appeals to radicals. Instead of having to criticize received authorities piecemeal, they are all swept aside to make way for 'thinking for oneself'. But this approach is wrong-footed for analogous reasons whether in politics, morality or epistemology. For in fact we cannot think for ourselves productively until we have had long practice in thinking other people's thoughts after them; and we think for ourselves precisely in order to resolve problems set by received opinions. The mistake comes with the idea that thinking for oneself would be made easier (rather than impossible) by starting with a clean slate — rather as if a swimmer were to imagine that, since it is the resistance of the water that slows the pace of swimming, one could swim much faster with no water at all.

The clean slate fallacy can take the form, as in Descartes, of a project of wiping the slate clean — of disbelieving whatever can be doubted. But for Descartes and rationalism generally, when the chalk is all rubbed off the slate, we find something written in the slate itself — some innate ideas or common notions or truths of reason that did not need to be derived from experience. For the empiricists, on the other hand, the clean slate (or wax tablet, in the time-honoured metaphor) is a datum; it is taken for a

simple fact that our individual minds start off blank, and that all we need to do is to let nature write on them. This denial that the mind actively contributes anything to knowledge is a rejection of ideas with which the mind comes into the world already endowed. But it also amounts to a systematic repression of the fact that we learn from others how to learn from nature. In consequence, empiricism makes people too radical in one way but too conservative in another. Too radical in that it may undervalue old authorities (as Aristotle was undervalued in the Enlightenment). But too conservative in a much more fundamental way: that it makes one uncritical towards one's own experience, and towards one's own concept of experience. It fails to see how 'received authorities' have moulded our experience, and its own conception of experience. It treats experience as self-authenticating and the concept of experience as self-explanatory. It does not recognize that what we experience is determined not just by what is there, but by what we have already learnt. Hence it can take experience itself to be an authority above criticism, unaware of the way experience can confirm our prejudices, since we may see what we have been taught to see.

Now the foregoing is a general account of empiricism as an instance of the 'clean slate' attitude; the hat does not entirely fit the two great seventeenth-century 'empiricists', Bacon and Locke. Bacon in particular, with his metaphor of the bee, stressed the active work involved in knowledge, and the need for new writing on the 'wax tablet' of the mind before we can erase the old. Strictly speaking, he was not so much an empiricist as an ancestor of classical empiricism, and also, by a different line of descent, of transcendental realism. It took a while for the basic flaw in the empiricist attitude to mature into a whole erroneous philosophy. But in Berkeley and Hume it does so.

Take Berkeley. He shows no particular interest in science, seeking rather to explicate our everyday knowledge of the world. But in doing so, he uncritically assumes a certain scientific model of perception. He assumes, in the first place, that an account of sight can stand in for an account of all the senses; that the paradigm case of sight is the case of 'just looking' (not, for instance, 'looking for ...', or any other practically oriented case of sight); and that what we actually see is coloured shapes. He does not discuss the knowledge, for instance, that we acquire of a tool by using it; he does not consider the effect of our possession of language on our experience. (For instance, the fact that a person with no ornithological vocabulary sees only birds, not pied wagtails or dunnocks. You only see pied wagtails and dunnocks when you have *heard about* pied wagtails and dunnocks.) He seems unaware that it takes a high level of philosophical sophistication to think of our visual field as composed of coloured shapes rather than of continuous particulars, and indeed, even given that sophistication, it is impossible for a sane person with normal eyesight to *see* the world as composed of coloured shapes.

My point here is not 'silly old Berkeley didn't notice these obvious phenomena'; obvious phenomena are often very difficult to notice. My point is that he was precluded from noticing them because he thought that a particular conception of experience was unproblematic, when in fact it was historically specific and learnt.

The empiricist concept of experience can be characterized as follows: (1) it comes through the senses; (2) it is passively imprinted by nature; (3) that on which it is imprinted was previously and would otherwise be a blank page; (4) it is individualistic, in that each person's experience can be understood without reference to other people's; (5) it is atomistic, i.e. each experience is only externally related to other experiences, and so can be understood without reference to them — and consequently it is one-dimensional, i.e. experiences are successive in time and therefore not combined in any structures (though they may fall into linear patterns).

With the possible exception of atomicity, experience has to have these features if it is to be the foundation of an objective knowledge which owes nothing to authority. But (a) this account of experience misreads everyday experience. It has nothing in common with what we mean when we call someone an experienced carpenter or soldier or lover. One important tendency of twentieth-century philosophy has devoted itself to replacing this inadequate account of everyday experience by one that does justice to its practical orientation and structured

character. I refer to the existential phenomenology (whether called that or not) of John Macmurray, Martin Heidegger and Maurice Merleau-Ponty. This tendency of course has quite different preoccupations from those of transcendental realism, but I believe their critiques of empiricism are for the most part complementary rather than contradictory. (b) This conception is also quite inadequate to account for science, for it treats knowledge as simply the accumulation of sense-impressions: but in that case, why are some things significant for science, and not others, why do we need to experiment, not just observe, and why do we need training to become scientists? It is sometimes said to be a strength of empiricism that it gives a single account of knowledge for everyday knowledge and science; I regard that as an error, but it is compounded by the fact that it misrepresents both the one and the other.

So far I have discussed empiricism as an attitude and as a conception of experience. But it can hardly avoid some other doctrines about the world. In particular, Hume's account of causal necessity seems unavoidable once we have posed the question in an empiricist way: 'as we have no idea that is not derived from an impression, we must find some impression that gives rise to this idea of necessity, if we assert that we really have such an idea' (A Treatise of Human Nature, pp. 153–4).

But necessity is not a coloured shape or a sound or a smell. An analysis of any complex practical activity might yield a concept of necessity, but not, of course, if one 'analysed' it into a succession of impressions. So Hume proceeds to derive the notion of necessity from the fact that the conjoining of two events is repeated.

> Did we never see any but particular conjunctions of objects, entirely different from each other, we should never be able to form any such ideas.
>   But, again, suppose we observe several instances in which the same objects are always conjoined together, we immediately conceive a connection between them, and begin to draw an inference from one to another.' (ibid., p. 161)

This is a pretty inaccurate account of our experience. We often infer necessity from a single case — if a light goes on when I press a button, for instance. Of course, the inference may be mistaken, but so may it when based on many cases. Anyway, correctness of judgement is not at issue here; we are talking about how we get a concept. Moreover, repeated conjunction does not always convince us of necessity. If we can think of no possible causal connection, we dismiss it as coincidence. So where *do* we get the concept of necessary connection? Well, if it's a matter of *learning* the concept, from our mums and dads. If it's a matter of *justifying* it, from various forms of practical interaction with the world, of which I have already discussed one (scientific experiment) in the last chapter.

Hume's failure to distinguish these two matters brings us to one of the most enduring and damaging legacies of empiricism: its tendency to ask questions of the form 'how do we know about x?' and think that the answer settles the question 'what is x?' I shall discuss this tendency in the next section. But first let us summarize the points on which, if the transcendental arguments of the last chapter are valid, empiricism as described here is refuted.

(a) The argument from the necessity of experiment shows that the 'spectator' conception of experience as passive observation is inadequate to account for scientific knowledge.

(b) From this also follows the need to distinguish epistemically significant from insignificant experience. Empiricism can't do this, since experience is simply the succession of impressions cast by nature: the more impressions, the more experience, the more knowledge; great knowledge of nature would be a function of old age.

(c) The incapacity of mere successive experiences to ground a theory of causation, since constant conjunctions rarely occur except when produced experimentally by us, shows the non-actuality of causation — i.e. causation cannot exist as no more than a relation between successive events; it must involve the generation of events by enduring structures. Empiricism is irretrievably actualist in its account of causation.

(d) The account of science as an inherently *social* activity, carried out by collaboration in institutions which transmit and transform information from one generation to another, rules out the empiricist assumption that knowledge is essentially an individual product and possession.

(e) The idea that the mind is a blank page *at birth* is an empirical hypothesis, though one on which the weight of evidence seems to be going against empiricism. But the necessity of scientific training shows that scientific knowledge at least can only be acquired by a mind that is already very far from a blank page.

The Epistemic Fallacy

In the last section I said that Hume asks how (or whether) we can know about necessary connection, and thinks that the answer tells us what (if anything) necessary connections are. This shift is so common in the philosophy of the last three centuries that it often goes unnoticed, and it is an important achievement of Bhaskar's philosophy to pick it out, name it, and (I think) refute it. Such a refutation offers an end to what Bertrand Russell has called 'the subjectivistic madness which is characteristic of most modern philosophy' (*A History of Western Philosophy*, p. 773). (Russell often uses such phrases in this connection; he also suggests — correctly, I think — that there is a sort of hubris vis-à-vis nature in this attitude. Nevertheless, the philosophical tendencies influenced by Russell have not escaped this madness and hubris.) In fact this epistemic fallacy pervades not only classical empiricism, where it originates (though Descartes must take much of the blame for setting philosophy off in this direction), but also Kant, the absolute idealists, Schopenhauer, Nietzsche, pragmatism, logical positivism, linguistic philosophy, poststructuralism, and, in a rather different form, phenomenology and existentialism.

Bhaskar defines the epistemic fallacy as 'the view that statements about being can be reduced to or analysed in terms of statements about knowledge' (RTS, p. 36). In casting my mind round for examples of this fallacy, it strikes me that it takes several forms, for example: (1) the question whether something exists gets reduced to the question whether we can know that it exists; (2) the question what *sort* of thing something is gets reduced to the question *how* we know about it; (3) the question whether A has causal/ontological primacy over B gets reduced to

the question whether knowledge of A is presupposed by knowledge of B; (4) the question whether A is identical to B gets reduced to the question whether our way of knowing A is identical to our way of knowing B.

We do not for the most part commit any of these fallacies outside of 'philosophical' contexts (in a non-technical sense of 'philosophical', that is; I suspect they are committed in the 'nocturnal philosophy' of scientists at least as often as by professional philosophers, and they crop up in non-specialist discussions of fundamental issues, too). For instance, it will doubtless never be known whether (as some pre-historians have contended, but most denied) all ancient societies went through a stage of matriarchy, but we have no difficulty in understanding the statement that they did, and recognizing that it may be true or false, and that whether it is true does not depend on us. We do not assume that there is any intrinsic difference between pebbles that have been perceived and those that have not, but whose existence we infer from geological knowledge. We may learn that there has been a gale when we see a fallen tree, but we do not assume that the tree falling caused the gale — or, if the epistemic ordering here seems too contingent: we may know of a magnetic field because pieces of iron are moved, but we do not assume that their movement caused the magnetic field. And if I think a burglar has been because my bike is missing, and you think so because the gate has been forced, we do not conclude that there must have been two burglars — a bike-stealing one and a gate-forcing one. Let us look at philosophical instances of these fallacies in reverse order.

4. Descartes's proof that he is two distinct substances, a mind and a body, involves many questionable metaphysical assumptions that I can't discuss here; but I think that *one* aspect of it is an instance of the 'two burglars' fallacy. Having 'proved' that he is a conscious being (mind) at an early (and supposedly indubitable) stage of the argument, and that he is an extended being (body) only at a later stage and with the aid of theology, he takes the conscious and extended beings to be two beings, instead of concluding that he is a being that is both conscious and extended.

3. The instance I have chosen of the form of epistemic fallacy that passes from the order of knowledge to the order of being is

one which comes from a philosophy which bypasses epistemology in the usual sense. By this I hope to show that the epistemic fallacy does not depend on the epistemological problematic, so that those who reject that problematic still need to be on their guard against it. The 'epistemic fallacy without epistemology' in the form under discussion means treating what is closer to us as thereby prior in itself. I take the example from the undeservedly neglected philosopher John Macmurray, because his text is clear enough to let us see the fallacy easily. But I would argue that Heidegger also fell into the same fallacy, thus constituting one of the many parallels between Macmurray's philosophy and continental existentialism.

Macmurray divides all changes that occur into *actions* (which have reasons) and *events* (which have causes). He then asks:

> What then do we mean by a 'cause'? We mean the source of an occurrence which stands to an event as an agent stands to his act, but which is not an agent. Since in any attempt to understand events the conception of cause must be thought positively, we must say that a cause is a source of occurrences which is a non-agent; an existent which is other than an agent.
>
> The conception of 'cause' is inherently self-contradictory. It is the conception of an agent that is not an agent, the negation of agency. The negative, we know, cannot exist independently, but only as the negative aspect of a positive in the form of the personal. Within action, which is a personal concept, there not merely can, but must be a negation of action; but this negation is in the last analysis a self-negation. If the negative aspect is thought as existing independently of the positive, the result is a contradiction. (*The Self as Agent*, pp. 152–3)

In other words: our understanding of cause is arrived at by subtracting something from our idea of action. Therefore, causes cannot exist in themselves, but only as aspects of actions. But this does not follow. Causes much like the ones we learn about in this way were operating before we existed. (After all, our conception of what it is to be an animal is arrived at by subtracting something from our idea of what it is to be a person, but it would be absurd to think that animals can only exist as aspects of people.)

Macmurray applies this account to experiments:

> Now consider the experimental situation without abstraction. To determine the law which governs the movement of the pendulum, I erect a pendulum, and I set it swinging. Then I begin to take the measurements I need. But during the experiment I do not interfere with its motion. My practical concern is to keep the conditions constant throughout — to prevent interference. When I have made all the measurements I require, I stop the pendulum, and sit down to study the measurements I have noted. The whole experiment is an action of mine: I *do* the experiment. But the pattern of movement I observe and the law that I elicit, refer only to what happens *within* my action. I leave out of account my starting the pendulum when I begin and stopping it when I have finished. The law of the particular instance refers to what happens between these points; to that aspect of my doing the experiment which I do not *do*; that is to say, the negative aspect of my action. If now we call this a causal process, we realize in another way that causality is the negative aspect of agency, and falls within action. (pp. 159–60)

There is much about this account of experiment which is excellent: it grasps the dialectic between active and passive in the experimental situation. But as we have seen in the previous chapter, the experiment would have no point unless it told us what would happen in non-experimental conditions. The uninterfered-with swinging of the pendulum in between the experimenter's actions may be 'abstract' in the sense that *in fact*, in this instance, it is part of a process which, considered as a whole, is an action. But it is not abstract in the sense that it could not occur outside that whole — and if that *were* so, the experiment would tell us nothing about nature. Actions may be prior to events in that *our access to* the knowledge of events is through knowledge of our actions. But that proves nothing about events in themselves being action-dependent.

2. The case in which things known in different ways are treated as different kinds of thing is already familiar from the case of 'theoretical entities'. As we have seen, the tendency to treat entities known only by causal criteria as less real than those known by perceptual criteria is rendered suspect by the frequent use of causal criteria for postulating things of a kind that could

also, under other circumstances, be perceived (the cat that tore open the dustbin liner in the middle of the night); and the repeated discovery, in the history of science, of techniques for perceiving what had previously been 'theoretical entities' ought to undermine anyone's confidence in distinguishing this kind of entity from other kinds.

There is perhaps a faint reflection of this version of the epistemic fallacy in 'commonsense', in the tendency to think that, whereas newscasters can lie, TV cameras can't, and in phrases like 'experience is worth a hundred books' (understand: '*my* experience is worth a hundred other people's').

1. Finally, the case in which questions of existence and non-existence are reduced to questions of knowability or unknowability. The verificationist theory of meaning, espoused by logical positivists, generates many striking instances of this. For example, A.J. Ayer's claim that, since God is not constructible out of sense-data, the word 'God' is meaningless, and the question of His existence does not even arise (see *Language, Truth and Logic*, chapter 6). While Ayer claimed that this view is neither atheism nor theism, since both are defined in terms of a meaningless word, it is effectively a peculiarly dogmatic and a priori form of atheism. The fact that rational dialogue is possible between atheists who believe that the concept of God is intelligible but uninstantiated (whereas perhaps 'Mother of God' is unintelligible), Catholics who believe God's existence to be rationally defensible and Kantians or fideistic Christians who believe that, though God can be proved neither to exist nor not to exist, the issue is of vital moral importance should be enough to refute Ayer's view. 'These language games are played' is a bad argument for anything's existence (another case of the epistemic fallacy, in fact); but it is a good argument against settling questions of what exists by narrowing the sayable to the knowable.

Of course, a logical positivist with a wider definition of experience might arrive at different conclusions from Ayer's, though not more defensible ones. I once had a friend (sadly no longer with us) who had a tendency to logical positivism in philosophy and to mysticism in religion, and who claimed that God is a logical construction out of mystical experiences. But I suspect that this has the unwanted consequence that we ought to worship a subset of our own experiences.

Now we need to ask why the epistemic fallacy is so convincing to so many, and, given that it is so, whether Roy Bhaskar has really refuted it.

It might be argued: when we discuss, say, causation, we are necessarily using our concept of causation. We may change our concept of causation, but the new concept will still be ours. Whatever our concept is, it must be one that we *can* learn, since it is one that we *have* learnt. So it makes no sense to ask what causation is in itself, quite independently of what we can know of it. Hence the question 'what can we know about causation?' really does answer the question 'what is causation?', insofar as any answer can be given. And if we are told: but causation might be quite different in itself from anything we can know about it, we can readily allow this empty Kantian gesture towards 'things in themselves' — but as Wittgenstein has said, what we can't talk about, we should shut up about.[2]

What are we to say to this? Well, first of all, it is obviously true that when we talk about causation we are *using our* concept of causation — what else could we be doing? But we are not *talking about* our concept of causation; we are talking about causation, using our concept, and that is something different. For we *can* talk about our concept of causation; we can say that it differs from Hume's, that it seems to be inadequate to cope with the behaviour of sub-atomic particles, and so on.[3] But that is different from talking about causation, just as a poem about the poet's beloved differs from an (almost certainly bad) poem about the poet's love.

It might still be asked: Does this distinction have any practical import for the theory of knowledge? Are we not as trapped inside *our* concept of causation when we are using it as when we are talking about it? For we can't compare our concept of causation with causation itself, since whatever concept we had of 'causation itself' would just be another concept of *ours*.

The last sentence must be accepted, but this doesn't mean that we can't make an epistemologically useful distinction between talking about cause and talking about our concept of it. For the criteria for deciding what to say, how to test and perhaps change what we have been saying, and so on, are quite different in the

first-order discourse (talking about causation) and in the meta-discourse (talking about our concept of causation). In the former, we focus upon the world through our practical interaction with it; in the latter, we focus on the history of ideas through our reading of its texts. The former can lead us to *revise* our concept of causation, to replace it with another one because it is inadequate. The latter can give us no good reasons for revising our concept.

Now someone might reply that with my talk about criteria for revising our concepts, I have slipped into talking about the concept of cause that we *currently* have, and may revise, whereas the claim is about any concept of cause we might have, since any such concept would be ours, i.e. humankind's. But what is the force of this claim? I concede the tautology that we have whatever concept we have. But if this is to throw any light on the concept of cause that we *must* have, it can do so only by reference to some knowledge of the untranscendable limits of human knowledge for all time, set by some contingent facts of our nature. Yet we know a lot more about cause than we do about any such limits.

But we can say *something* general about our having of concepts of causation, and it is nothing to do with our species-specific limits. We can examine what would be a good reason for replacing one concept of causation by another. Such reasons will have to do with the inadequacy of the discarded concept to account for what happens quite independently of our applying any concept.

In Bhaskar's terms, the claims that I have been considering amount to saying that at any given time our knowledge belongs to the transitive dimension (which is true), and that this in some sense 'traps' us. But if we are trapped inside the transitive dimension, this is only for the reason that whatever new knowledge we produce belongs by definition to that dimension. This is no real trap, since we can always *change* the transitive dimension, and that we do so in the ways that we do is (in the best case) explained by the fact that the transitive dimension is not an end in itself, but produced entirely in order to explain what occurs in the intransitive dimension.[4]

Now I want to draw attention to another slippage in the 'defence of the epistemic fallacy' outlined above. Is the claim

about what concepts we could learn, or about what knowledge we could have? It often appears to be the former, and only derivatively the latter (since we can't have knowledge without concepts). But in the first place, we should be very suspicious of arguments from how we could learn a concept. They are often genuinely guilty of the genetic fallacy, as Professor Mundle has pointed out (*A Critique of Linguistic Philosophy*, esp. section 17) — i.e. they treat the question how a concept is learnt as deciding the question what it means. Yet the process of learning concepts is complex and often 'devious', in the sense that we might *have to* mislearn some concepts before we can correct this inadequate knowledge and get it right, as every teacher knows. In some cases, we may have to learn some concepts (e.g. similarity and continuity) before we can learn another (identity) which is logically prior to them.

Further, the obsession with asking how we could learn a concept, which dominated philosophy teaching in the UK during the third quarter of the twentieth century, was unhelpful because we know and will probably always know much less about how we acquire concepts than about what they mean, and indeed because we seem as a species to be almost infinitely fertile in the production of concepts, most of which have no application in reality.

Finally, if it really were proved that we could not *possess the concept* of, for example, real necessary connection, too much would have been proved, since if this concept had not been taught by philosophers and taken for granted by most people, there would have been no point in Hume's arguing against it. So we must come back to the question: how can we know when to correctly apply our concepts? It is a question about what we can know about the world, not about which concepts we can use. And this already takes us out of the more extreme versions of the epistemic fallacy, e.g. verificationism. For we can quite intelligibly (if not intelligently) ask how many angels can dance on a pin, without having a hope in hell of finding out. We may make many coherent statements which may be true or false without our ever being able to find out which.

If this is so, the question 'what can we know?' is far from being answerable in advance of claims about what there is, which it could then arbitrate. The point is rather, by keeping

questions about what there is open, to put our current knowledge constantly in question; to keep us asking: Is this really true? Does it match the real world better than other theories or not? Let us test it, let us put questions to nature and revise our beliefs in the light of the answers.

So far in this section I have been trying to dispel the plausibility of the epistemic fallacy, rather than confronting it with a knock-down refutation. But if the central transcendental argument discussed in the previous chapter is valid, the fallacy is already refuted. For one conclusion of that argument was the reality of structures in nature, independent of us, which endure longer than the experiments by which we test them.

However, Bhaskar also has an argument from the nature of perception, which, without necessarily committing one to the whole transcendental realist ontology, shows that the epistemic fallacy prevents its adherents from accounting even for this mode of experience.

> If changing experience of objects is to be possible, objects must have a distinct being in space and time from the experience of which they are the objects. For Kepler to see the rim of the earth drop away, while Tycho Brahe watches the sun rise, we must suppose that there is something that they both see (in different ways). Similarly when modern sailors refer to what ancient mariners called a sea-serpent as a school of porpoises, we must suppose that there is something which they are describing in different ways. (RTS, p. 31)

Note that here, as in my response to the reasons that, I have suggested, make the epistemic fallacy plausible to some, the crucial premiss of the realist argument is *change*. A static, 'snapshot' view of perception may well appear to be adequately analysed without reference to beings independent of us. But we can neither describe nor justify cognitive change without reference to independently existing objects. If while approaching a landmark on a walk I say, 'I thought that was a barrow, but it has turned out to be only a clump of bracken', the 'that' and 'it' in the two clauses must have a common referent if the sentence is to make sense; and it must have definite, initially misrecognized characteristics to justify my change of judgement.

This dependence of the rationale of realism on the fact of cognitive change will crop up again. We shall see that modern idealism results from trying to take change seriously *while remaining within the epistemic fallacy*. But only the rejection of that fallacy can make possible a coherent account of change. Empiricist versions of idealism, on the other hand, rest on a twofold perceptual stasis: perception is taken out of its normal context of practical activity; and change in the content of perception is ignored. The same dependence on a static vision can be seen in that shadowy idealism which is sometimes encountered in popular 'nocturnal' philosophy: the hippie at the party who would scoff at objectivity, saying 'if I want I can regard this glass of wine as a rose'. For a few seconds he (always 'he') would stare at the glass contemplatively, having rosy or winy sensations — who is to say? But once action resumed, he would be far more likely to drink the glass of wine than to put it in his buttonhole.

The Triumph of Will: Modern Idealism

> Once upon a time, a valiant fellow had the idea that men were drowned in water only because they were possessed with the idea of gravity. If they were to knock this notion out of their heads, say by stating it to be a superstition, a religious concept, they would be sublimely proof against any danger from water. His whole life long he fought against the illusion of gravity, of whose harmful results all statistics brought him new and manifold evidence. This honest fellow was the type of the new revolutionary philosophers in Germany.
> (Marx and Engels, from the Preface to
> *The German Ideology*, p. 37)

The response of the uninitiated to classical idealist philosophy — to Berkeley or Kant — is to think that they are saying that it is up to us what the world is like, in the sense that we could change it, not by hard practical work, but by seeing things differently. In fact, they held that it is *down* to us, but not *up* to us — that is to say, that it is because of us that the world is as it is (or appears as

it appears); but we could not make it be (or appear) different by anything we might (cognitively) do.

Berkeley, for instance, thought that there was nothing in the universe except minds and the ideas that they perceived. But finite minds did not *choose* what to perceive; he was quite aware that what we see when we open our eyes does not depend on us. Indeed, if when I look in a given direction I can see a herd of fallow deer, then you will normally see a herd of fallow deer if you look in the same direction; Berkeley explains this by the infinite goodness of God, who makes you and me see the same things, so that we have got something to chat about, and don't quarrel. Berkeley's account of perception makes it every bit as passive, and knowledge every bit as objective, as does Locke's.

The case of Kant is more complex. He thought knowledge was a *product* of our minds; that they imposed its form as ordered in space and time, conforming to the categories of causality and substance, etc. Its *content* depends not on us but on the things-in-themselves which are themselves unknowable. We have no choice about what the things-in-themselves are, or how they affect our senses. And though our minds 'work' on this content and transform it, this work all takes place, so to speak, behind our backs: all we know is its final product. And that final product is the same for all of us, because we necessarily apply the same spatio-temporal grid and the same categories to raw material received through the senses from the same things-in-themselves. So for Kant, too, there is neither any *choice* of what to experience, nor any *variety* in the forms of knowledge, contributed by the different working of different minds. Just as for Berkeley, conscious experience of the world is passive, knowledge is objective, and choice has no place in any cognitive process.

So whatever may have been the justice in Marx and Engels's jibe against the Young Hegelians in the above quote, it does not apply to classical subjective idealism, any more than it does to Hegel himself. In the twentieth century, though, the kind of idealism that treats the world as dependent on our cognitive choices (whether those choices are regarded as free, or as determined by our historical situation) has really come into its own. Modern idealism is, in this sense, much *more* idealist than that of Berkeley or Kant or Hegel: it sees the world as more

subjective, mind-dependent in a stronger sense, than any previous philosophy. Bhaskar calls it 'superidealism'. For the first time, it has become respectable to write sentences like:

> *Objects* of discourse do not exist. The entities discourse refers to are constituted in it and by it. (Hindess and Hirst, *Mode of Production and Social Formation*, p. 20)

or

> A fish is only a fish if it is socially classified as one, and that classification is only concerned with fish to the extent that scaly things living in the sea help society define itself. After all, the very word 'fish' is a product of the imposition of socially produced categories on nature. (Keith Tester, *Animals and Society*, p. 46)

It is a curious fact that transcendental realism and super-idealism, which both owe something to Kant, and which have moved away from him in diametrically opposed directions, nevertheless take leave of Kant for very similar reasons, and imply very similar critiques of him. Their common critique of Kant (which indeed has its forerunners in Hegel and Marx) would be along these lines: Kant is right to see knowledge as a product of a cognitive *work*, which transforms its raw materials. But he is wrong to think that that work goes on within each individual mind, and in the same way for all. It is work in a much more straightforward sense than that — time- and energy-consuming work that goes on in the public world, and works on historically specific raw materials, with historically specific means of labour, organized in historically specific institutions. Hence the products are different in different times and places, and these differences *do* depend on our activity and on the particular nature and situation of its agents.

What, then, is the point at which these two critiques come apart? One might say that they dispose of the 'thing-in-itself' in different ways. The super-idealists leave it on one side as something we can't talk about, and take the knowledge that we produce (the transitive object, in Bhaskar's terms) to be the only object; Bhaskar argues that without reference to the intransitive object, we cannot make sense of our activity of knowledge-

production. For knowledge-production is unlike many other human activities which also result in discursive products — poems, sermons, jokes, and so on. These do not seek to explain what goes on in the world independently of them; knowledge does. Hence we produce knowledge by a process that essentially involves taking soundings from the intransitive object, which — since our knowledge of it is thereby deepened — is not to be regarded as a Kantian thing-in-itself ' = x' (as Kant says).

What this difference amounts to is that the superidealists seek to incorporate an account of the variety and change in human beliefs into an outlook that remains within the epistemic fallacy. Hence variety and change come to be seen as applying to the *world* that is known, not just to the knowledge. Not, of course, just in the sense that there is variety and change in the world, which obviously there is; but in the sense that there is variety and change between worlds, depending on the variety and change in our beliefs. As against this, Bhaskar argues that one cannot describe variety or account for change without reference to the intransitive object; they require us to abandon the epistemic fallacy. For it is only 'Once we constitute an intransitive dimension' that 'we can see how changing knowledge of unchanging objects is possible' (PN, p. 11).

We already have here the means to refute the sort of superidealism expressed in the quotes from Hindess and Hirst and from Tester — the sort which involves explicit denial of the relation of knowledge to anything outside it. For it fails to give any satisfactory means of differentiating knowledge from other human products. It leaves one asking what the point of any specifically cognitive production is, since it can't be to widen or deepen our knowledge of anything. The answer, I suspect, is that such views are usually motivated by an opportunistic style of politics: ideas function politically, in fact, by referring to social realities; but if a politically expedient discourse is challenged with the claim 'that is not so, the facts are different', a meta-discourse like Hindess and Hirst's which disavows objective intent serves to protect it, to immunize it against criticism. The 'two science' theory ('there is bourgeois science and proletarian science') was used in this way to defend Stalin's patronage of Lysenko's pseudo-science.

While such pragmatism could be hitched to any political wagon, the implicit idea that to change discourse is to change reality has a close affinity with the bureaucrats' instinct for euphemism: don't start treating pauper lunatics decently, start calling them 'rate-aided persons of unsound mind', and everyone will *think* you have started treating them decently. Tester's remarks suggest a wonderfully cheap way of solving two problems of maritime ecology at one stroke: we could reclassify lumps of untreated sewage as 'fish'.

But there is a very much more intellectually serious version of superidealism, which arises from philosophical work done on the history of science, of which Thomas Kuhn's is the most significant. Kuhn's *The Structure of Scientific Revolutions* does not include ontological denials like Hindess and Hirst's; it is not a work of militant, drum-beating anti-realism, but a ground-breaking study in the way science has progressed by way of sharp breaks between old and new paradigms. However, he claims that the different world-views involved in different paradigms are, or can be, *incommensurable*, and hence that their adherents live, in a sense, in 'different worlds'. Kuhn is usually quite cautious about using this kind of language, but the doctrine that most readers seem to come away with is that there are only our incommensurable interpretations of nature, nothing outside them for them to be more or less true of. This view is three steps removed from the old empiricist or positivist view that there is a core of agreed 'observation statements' by which to judge our theories. It is important to recognize that these three steps are separate; it is sometimes thought that if you take the first two you must take the third, which would leave no room for the transcendental realist position. The first step, shared by transcendental realism and indeed most recent philosophies of science, is to say that the empirical dis-confirmation of a theory is not a two-place relation between a theory and the disconfirming data, but a three-place relation between two theories and the data that disconfirm one theory relative to the other. Thus, one does not abandon a theory until one has got a better one. The second step, also widely shared, is to admit that the data are themselves theory-dependent. This does not by itself, as is sometimes alleged, disqualify them for their critical role in relation to the theories, for the two theories

may agree about the data, which may still be shown to support one of them against the other. The third step is to say that between *incommensurable* theories there can be no relation, and no agreement about data. Hence the transition from one theory to another can only be a leap, pushed by the internal anomalies of the old theory, and landing among new 'data' concerning a different 'reality'.

Now let us consider how Bhaskar deals with this problem of incommensurability, which he does in the chapter of SRHE called 'Incommensurability and the Refutation of Super-idealism' (pp. 70–93). He starts by pointing out that the issue is not necessarily an all or nothing one between the strict incommensurabilty of two languages with no shared meanings, and meaning-uniformity, in which there are no problems of mutual comprehension.

Rather, these are only limiting cases of *meaning-variance* (some degree of divergence of meanings) and *meaning-filiation* (some degree of mutuality of meaning). These normally coexist. Bhaskar mentions in a footnote that he has argued, in PN, chapter 4, that the extreme cases — incommensurability and meaning-uniformity — are nonsensical.

> Communication is impossible unless some descriptive and practical pre-suppositions are shared in common; unnecessary unless there is the possibility of discrepancy (non-identity in objective content) between them. (PN, p. 153)

So if in what follows he assumes a state of incommensurability for the sake of argument, we may take it that that argument is intended as a reductio ad absurdum.

Bhaskar accepts what he calls 'epistemic relativism', i.e. the recognition that our beliefs are socially produced, transient and fallible.[5] But he claims that this does not commit us to *judgemental relativism*, i.e. the idea that 'all beliefs are equally valid in the sense that there are no rational grounds for preferring one to another' (SRHE, p. 72). (See the discussion in chapter 6 below, where I suggest that 'rational' should read 'cognitive', though it is arguable — but not obvious — that the two are equivalent in this context.)

In order to show how judgemental relativism can be avoided, Bhaskar asks whether, if there were two incommensurable theories, there could be grounds for rational choice between them. To answer this, we need to distinguish the *sense* of the two theories — the meanings, definitions, etc. of the terms used — from their *reference*, i.e. their referring to some object, their 'referent'. This is a familiar distinction which goes back to Frege: two expressions with a different sense may have the same referent, e.g. 'the Queen of England in 1990' and 'the eldest daughter of George VI'. The referent need not be a concrete object; it could be a causal mechanism (as in Maxwell's criterion of scientific reality[6]), or even a mathematical number ('the even prime' does not mean the same as 'the square root of 4', but both refer to the number 2).

Bhaskar argues that in cases where incommensurability of sense is claimed, there will nevertheless be a common referent; the sense belongs to the transitive dimension of the theory, the referent to the intransitive dimension. This makes possible a rational choice between the theories on the grounds: which theory explains more of the same phenomena under its own description. For instance, what Priestley called 'de-phlogisticated air', Lavoisier called 'oxygen'; Lavoisier could explain more by his theory about oxygen than Priestley could by his theory about de-phlogisticated air. Hence it was rational to choose Lavoisier's theory.

Bhaskar goes on (SRHE, p. 73) to anticipate the obvious objection that he is assuming two things which are just what are at issue, namely (1) that there *is* an intransitive object for the two theories to refer to, and (2) that a 'subject' could be in a position to choose between them. For the superidealist use of the concept of incommensurability typically involves claiming that there is no *object* outside the incommensurable theories, and that any 'subject' must be too 'inside' one theory to operate with the other.

But Bhaskar argues that, while the superidealist may make such claims, they necessarily proceed as if those claims were false; for if there is no common referent, the theories don't clash at all, and the 'incommensurability' ceases to be interesting. Nobody bothers to say that astrology is incommensurable with monetarism or generative grammar with acupuncture. They are

just about totally different things. Incommensurability, in any interesting sense, is supposed to be a species of *conflict* between theories. And if two theories have no shared meanings, they cannot clash by virtue of one including negations of propositions asserted by the other;[7] they can only clash *about* something, as Priestley's and Lavoisier's theories clash about the gas which the one calls de-phlogisticated air, the other oxygen.

Where two theories lack a common referent and hence do *not* clash, there is no problem about some 'subject' operating with both theories. There is nothing paradoxical about being both a Kleinian psychoanalyst and a Sraffian economist. But what of the case of clashing incommensurables, i.e. theories incommensurable in sense but having a common referent? There might *be* good reasons for preferring one to the other, but, it might be argued, no one could ever *have* those good reasons, since understanding one of the incommensurable theories would preclude understanding the other.

In the first place, it may be said that, as a psychological claim, this is rather implausible. Just as one can switch back and forth between seeing a duck and seeing a rabbit in the famous duck/rabbit gestalt switch, so one may switch, if uneasily, between different world-views, and have no difficulty in *understanding* the one that one is currently *not* in. It has been suggested that King Charles II slipped between Catholicism and scepticism in this way, or Pierre Bayle between Protestantism and scepticism. So far from being impossible, this is quite common. But the superidealist might claim that incommensurability precludes understanding not merely psychologically, but logically or semiologically or 'conceptually'; and therefore that the cases I have mentioned are either misdescribed, or the two world-views are not really cases of incommensurability. But if incommensurability is understood in such an extreme sense, it becomes impossible for the situation which gives rise to the problem of incommensurability ever to occur. For it is precisely in the situation in which a scientist or scientific community is *changing* from one theory to another, or *comparing* their theory with an alternative, that these phenomena are said to occur (or alternatively when two contemporaneous cultures *encounter* each other in such a way that mutual understanding about something becomes an issue for them). Maybe such situations of

theory-comparison and cross-cultural encounter never occur, but if they occur, they occur *for someone* and *about something*. The realist is not committed to denying that there could *be* complete incommensurability, only that anyone could ever be confronted by it. Hence Bhaskar does not preclude

> the possibility of two epistemic communities travelling on, so to speak, semantic world-lines which never meet and know nothing of each other. (SRHE, p. 74)

They could even be described by a third party. But the incommensurabilist is interested in the incommensurability

> between different mundane, i.e. historically realized, human cultures at [some] or over time, especially within science. Moreover the choice situation which excites him or her is not the 'external' one open to a prospective third party, but that actually confronted from within by the communities concerned. (SRHE, p. 74)

And the case in which there could be no subject capable of rationally choosing is the case in which there could be no subject confronted by the choice at all. The argument against objectless (referentless) incommensurability is a conclusive one: if incommensurability is a kind of clash, there must be something to clash about; two theories that are not about the same thing can't clash. But the argument against subjectless incommensurability is an argument about its importance in the philosophy of science; two communities *might* conceivably have no overlap of sense in their theories, and hence not know that they were referring to the same thing. Of such a case, one could still give some sort of account, drawing on Kuhn's idea that the main way that the scientific community changes its opinion is not by the individuals who compose it changing their opinion, but by their dying off, and being replaced by a new generation committed to a new paradigm. However, (1) there are surely numerous instances when we do understand the reference of a sentence whose sense is alien to our own world-view. There is a tale of a citizen of a police state walking down the street muttering 'the hypocritical, lying, murdering, double-crossing swine', and being overheard by a secret policeman who promptly arrests

him for insulting the authorities. And in any argument about, for instance, economics, one may find oneself using phrases like 'what you would call "disguised unemployment"', 'what you would call "growth"', or 'what you would call "an economic rent"', using concepts absent from one's own conceptual scheme (and arguedly so) to designate phenomena identifiable by both parties.

(2) Further if science did change in the way described in the previous paragraph, the change would be a matter of fashion rather than progress. Probably much intellectual change is like this: when did British analytical philosophers ever get the better of British idealist philosophers in argument, or French poststructuralists of French structuralists? One might as well talk about arguments between short and long hair. But (a) if Kuhn's description of the *motives* of scientific change is right, this account can't be true *of science*, for according to Kuhn, a paradigm is not abandoned until it has gone into crisis through the accumulation of anomalies; and those anomalies accumulate in the course of applying the paradigm to explain the phenomena. If what we can reasonably say were not constrained by nature, we could keep a paradigm afloat indefinitely. The 'fashion' theory of scientific change belongs not with Kuhn's account of science, but with Feyerabend's (of which more presently). (b) While fashion may be an important determinant of intellectual change outside science, this is surely to be regretted. For all it produces is an 'eternal recurrence' of opinions, while criticism of a theory that has first been understood and assimilated sometimes produces a better theory. Intellectual progress through criticism is sometimes possible, and the law of fashion — that the last generation's fashions are self-evidently ridiculous, and the ones before that unknown — is the chief obstacle to such progress.

Before passing on to the more overtly voluntaristic superidealism of Feyerabend, it should be said that the realist case against incommensurability does not involve any playing down of the real difficulties confronted by translators, migrants, anthropologists, historians of ideas, or participants in dialogue between widely differing world outlooks. But these are practical problems which can at least partly be solved by skill, ingenuity and hard work. When it is said that a translation of a given text

is 'impossible' (as has been said, to give a philosophical example, of Heidegger's *Sein und Zeit*), this is true if by 'translation' we mean a smooth, linear text conveying all the meanings of the original. But such things as extensive footnotes explaining the connotations of the original can get much of the meaning across (as translators Macquarrie and Robinson demonstrate in the instanced case). As Bhaskar concludes his case against the argument from 'subjectlessness',

> no agents could find themselves in the situation described by the super-idealist, so that the philosophical fancy of inc[ommensur-ability] must cede to appreciation of the real difficulties of meaning-variance-in-filiation in and around science. (SRHE, p. 75)

Now I come to Feyerabend's frankly voluntaristic position — a *celebration* of the 'subjectivistic madness' deplored by Russell. Feyerabend's data from the history of science often look like Kuhn's, but the lifelines left by Kuhn whereby much of his theory can be retrieved from a realist standpoint have been cut by Feyerabend. This he can do with equanimity, and indeed pleasure, because his aim is not really to provide for science a philosophy that does justice to its history, but rather to undermine its epistemic authority. As Bhaskar puts it in an early article (contemporary with RTS):

> Like an undercover agent who works on both sides of the fence, Feyerabend plays the game of reason in order to undermine the authority of reason. His position is not self-refuting because it is clear that Feyerabend is in fact committed, in *Against Method*, to higher-order values. These may be summed up as: *for freedom* and *against science*. ('Feyerabend and Bachelard: Two Philosophies of Science', p. 41)

It is instructive to consider the use Feyerabend makes of John Stuart Mill's 'On Liberty' in this connection. Mill defends freedom of thought and expression on the grounds that the free clash of different ideas is the best guarantee of the progress of human knowledge. Feyerabend supports this plausible idea with evidence from the history of the sciences. But Mill has a fairly clear idea what he means by 'progress'. He expects our

ideas to get *closer to the truth* by this means. This conception is not available to Feyerabend; when he talks as if he means this, his arguments must be treated as ad hominem (in that they argue from premises of Mill's that are not Feyerabend's own). It is the freedom of thought itself that he values, not any epistemic progress that might come from it. Now in a certain sense of 'freedom of thought', Mill and indeed most of us value it as an end in itself; freedom of thought, that is, as a civil right: freedom to hold and express our opinions without interference from the police. But Feyerabend means something else as well: that we can choose what to believe, independently of any grounds for thinking it true. We will have *motives* for choosing our beliefs of course, but these may be boredom with our old beliefs, nationalistic pride in the ideas of our compatriots, or indeed anything whatever.

Now this seems to me to misrepresent what belief is; believing something entails thinking that there are good grounds for believing it. Otherwise we are only *pretending* to believe, even if we manage to fool ourselves too. Feyerabend's conception of belief devalues and degrades human thought; it also undermines one of the strongest arguments for freedom of thought: that since we cannot choose what we believe, we cannot be commanded to believe anything. Finally, it undermines our other, non-cognitive freedoms too:

> For Hume, 'reason is and ought only to be the slave of the passions'. But in Feyerabend the passions lack their necessary complement: an efficient slave. Knowledge may not be the most important social activity, but it is the one upon which the achievement of any human objective depends. Freedom, in the sense Feyerabend attaches to it, depends upon knowledge (praxis presupposes theory); we can only be as free as our knowledge is reliable and complete. We are not free to choose what we believe if we are to attain the kinds of objectives Feyerabend mentions. Only if belief-in-itself was the sole end of human action would Feyerabend be warranted in such an assumption. (ibid., pp. 42–3)

Of course, one may reject the purely instrumental relation of knowledge to freedom characteristic of the utilitarianism referred to in this passage. Bhaskar's own account of freedom is

a much richer one (see chapter 6 below). But any plausible alternative to the utilitarian view will make the relation between knowledge and freedom *more* intimate than this (and Bhaskar's does so).

It will be useful at this point to consider Bhaskar's prolonged critique of one version of superidealism, illustrating the way in which Marx's jibe quoted at the beginning of this section fits it. This critique constitutes the larger part of Bhaskar's book *Philosophy and the Idea of Freedom* (PIF), and the version criticized is that of Richard Rorty, often regarded as the main representative of 'postmodernist' philosophy in the English-speaking world.[8] The title *Philosophy and the Idea of Freedom* — backed up by a cover picture of Jacobinical jollifications — may seem strange for a book which is three-quarters a critique of Rorty, one-quarter an account of critical realism and its relation to Marxism. But its central theme is a confrontation between two conceptions of freedom — an 'in-gear' and a 'freewheeling' conception, held by Bhaskar and Rorty respectively.

From what I have said so far, the 'ideal case' of superidealism might be described as a voluntaristic and discursive remake of Berkeley's philosophy. *Discursive* in that the 'mind' which 'it is all in' is linguistic rather than perceptual in character. And *voluntaristic* in that what there is depends not just on us but on our *choices*. It is possible, though, to hold a version of superidealism which, while having the same voluntaristic and discursive character, resembles Kant in leaving 'the starry heavens above' in the realm of scientific determinism, and claiming for freedom only 'the moral law within', realized within the limits of pure discourse.[9] These positions are both criticized by Bhaskar as held by Rorty at different stages of his work (if not sometimes at the same time).

First as to Rorty's version of Kant's dualism of 'the starry heavens above and the moral law within'. Like many philosophers who are concerned to keep positivist science at bay with regard to the human world, Rorty concedes 'the starry heavens above' to it too easily (the same could be said of Winch and Habermas). Science is seen as being much as the positivists describe it — actualist, predictive, likely to arrive at a one-level determinist account of everything in terms of Humean laws governing physical processes. A good deal of the desire to keep

science at bay stems from assuming that this misdescription of it is true, and recognizing how devastating the consequences would be of reducing our account of the human world to such terms. Once this model of natural science has been shown to be false, there is much less cause for concern at the thought that it might have something to say about the human world. And if Bhaskar's central arguments are valid, that model of science *has* been refuted.

Bhaskar calls one chapter of his book 'A Tale of Two Rortys', and this positivistic Rorty is one of them. The other is the one who seems to matter most to Rorty himself, the 'existentialist' Rorty (spirits of Heidegger and Sartre hold your peace!), the advocate of what I shall call out-of-gear freedom. This metaphor, I hope, is clear enough: in-gear freedom is a matter of interacting causally with the world in order to realize our intentions; it is threatened by any view which denies the efficacy of our intentions in bringing about changes in the real world; out-of-gear freedom is precisely a matter of disengaging our choices from causal interaction with the world, to ward off the threat that the nature of that world might limit or determine them. One instance of an out-of-gear conception of freedom is expressed by Rorty: 'Man is always free to choose new descriptions (for, among other things, himself)' (*Philosophy and the Mirror of Nature*, p. 362n).

Rorty's position as summarized by Bhaskar is as follows:

1. All things may be redescribed, even if they do not change, possibly in terms of an incommensurable vocabulary.
2. All things may exhibit novelty, and so require a new, potentially incommensurable discourse.
3. Only human beings can discourse (normally or abnormally, literally or metaphorically). And:
4. Only human beings can overcome themselves, their past and their fellow human beings — and they do so in and by (creating a new) discourse in terms of a new incommensurable vocabulary. (PIF, p. 62)

This is the only freedom that Rorty allows us: to redescribe the world, each other, ourselves. But his fundamentally super-idealist view means that this freedom is sufficient, since it

changes the 'world' in the sense that a Kuhnian paradigm shift does.

> The final victory of poetry in its ancient quarrel with philosophy — the final victory of metaphors of self-creation over metaphors of discovery — would consist in our becoming reconciled to the thought that this is the only sort of power over the world that we can hope to have. (*Contingency, Irony and Solidarity*, p. 40, quoted PIF, p. 64)

One is reminded of Freud's distinction between health, neurosis and psychosis (in 'The Loss of Reality in Neurosis and Psychosis'). The healthy person changes the world by action, the neurotic withdraws from the world and changes himself, the psychotic 'changes the world' in a more lordly manner, by changing his own perception of it. Here Rorty is presenting psychosis as our true freedom.

However, some sort of Rortian practice is at least recognizable in the inter-personal sphere. While we may redescribe anything, our *self*-redescription becomes self-creation if we can get others to accept it and thus give it the only sort of 'truth' admissible by a superidealist. So that if, I suppose, the word 'gay' for 'homosexual' comes to be generally accepted, society will be transformed in the only way in our power, and as a result bigots and thugs will learn to indulge in gay-bashing instead of queer-bashing. Some of us might have thought that liberation meant more than that. But if Rorty's out-of-gear freedom is the only one we've got, such redescription is the best we can hope for. By the same token, the agonies of the oppressed must appear as 'not drowning but waving'.

Granted, Rorty recognizes that freedom even in his sense is threatened by scarcity of food, and secret police, but this is seen primarily in terms of their depriving us of the leisure and peace of mind necessary if we are to engage in such 'higher' (or self-indulgent) activities as redescribing ourselves.

Bhaskar's alternative (in-gear) conception of freedom is defended in his chapter 'How is Freedom Possible?', but I shall not expound it here because I do so fully in chapter 6. My intention here is only to show the inadequacy of Rortian

freedom. The only human emotion it could be used to gratify is vanity.

If Rorty's account of science is flawed and his notion of freedom unsatisfactory, these faults are compounded by the fact that the two are not even compatible. For as Rorty fully recognizes, even such relatively freewheeling acts as talking and writing and indeed thinking involve some motion of matter which is subject to the laws of physics. If natural causality is the mechanistic and one-level thing that he takes it to be, even a purely discursive freedom is impossible. As Bhaskar puts it:

> The problem for Rorty, as for Kant, is how, if the lower-order level is completely determined, what is described in higher-order terms can have any effect on it. And of course, the fact is that it cannot. If the intentional level, at which we cite reasons for actions and offer justifications and criticisms of beliefs, is merely a redescription of movements which are already sufficiently determined by antecedent physicalistic causes, then the *causal irrelevance of reasons* for the states of the phenomenal world of bodily movements and physical happenings (including the production of sounds and marks) immediately follows. (PIF, p. 53)

Finally, alongside this 'Kantian' dualism, there is a different dualism, or series of dualisms, running through Rorty's work, of which both poles are within discourse: between metaphysics and irony, normal and abnormal discourse (generalized from Kuhn's normal and revolutionary science), or between scientific and literary cultures. In each case, as Bhaskar puts it, Rorty's philosophy is a 'continual posing of dichotomies between, on the one hand, a hard fundamentalist demand usually steeped in actualist folklore and, on the other, a soft deflationary option, usually with voluntaristic overtones or leanings. (PIF, p. 133). However, Rorty *underdescribes* science, playing down both its non-discursive features (e.g. experiment) and its specific differences from poetry or biblical exegesis. In principle, it is treated as no more objective or truth-seeking than other discourses, i.e. it is treated in the voluntaristic way characteristic of superidealism. But it is in the other, 'softer' zones that Rorty really wants to give this voluntarism free play. He is a good enough pragmatist to let science get on with its job of providing prosperous North American academics with their cars and computers and intercontinental air trips. The sort of discourse which finds no room in Rorty's inn is that which seeks to inform the hard work of transforming recalcitrant social realities, which neither technological innovation nor redescription will budge.

Positivism as Ideology

The last of the three long chapters that constitute SRHE is called 'The Positivist Illusion'. It is an explanatory critique (see chapter 6 below) of positivism, extremely intricate in structure, and requiring a detailed commentary rather than the brief introduction which is all I have space to give it here. But I should say something about what Roy Bhaskar is doing and where he is going in this, the first-fruits of that series of sustained critiques of philosophical ideologies which he has long promised us.

We need to be quite clear at the outset about the aim of this critique. First, it presupposes rather than proves that positivism is false. It aims to show how that theory functions, why it has seemed plausible to many, how it affects the practice of the sciences, what place it occupies in the kind of society with which it is associated. That it is false follows not from the arguments in this chapter but from Bhaskar's arguments which I have outlined above in chapter 2, if they are valid.

Second, the positivism that is attacked is not just the explicit positivism of the schools of thought that have used that name. Bhaskar is claiming (as others have before him) that positivism is the dominant ideology (at least of those ideologies relating to science) in capitalist cultures. Positivism as an explicit movement has never remotely approached to that. Both the Comtean positivism of the nineteenth century and the logical positivism of the Vienna Circle in the twentieth have been fringe phenomena — and indeed politically distinctly to the left of centre. Comte had been associated with the utopian socialist Saint-Simon, and his own followers are always turning up in the history of the nineteenth-century left. The Vienna Circle were for the most part progressive democrats in a national culture where reaction was strong; one of them, Otto Neurath, was a revolutionary socialist. But it may still be the case, and I believe

that it is, that certain key doctrines associated with positivism have been *implicit*, in a variety of different ways, in much of the culture of capitalism insofar as it relates to the sciences. Positivist assumptions are often taken for granted by scientists themselves, even when their own scientific *practice* requires a realist analysis, not a positivist one. More importantly, in the *human* sciences, positivism has not only been a common 'nocturnal philosophy' but has influenced the diurnal practice of the workers in these fields. For they assume that natural science is as described by the positivists, and try to imitate this, hoping thereby to match its impressive achievements. It is a recurring theme of Bhaskar's work that the human sciences have been misled by a positivist misunderstanding of the natural sciences; and this has been the basis of several valuable interventions of transcendental realism in the work of the human sciences, as we shall see in chapter 7.

Finally, the anti-positivist strand in the culture of capitalist societies, from the romantics through neo-Kantianism to Rorty, is often parasitic on positivism, both in the way that 'anti-' reactions often are, as mere inversions, and in that it usually accepts the positivist account of natural science, treating it either as an enemy or as 'alright in its place' but irrelevant to the 'things that matter'; they rarely have any interest, as Bhaskar has, in rescuing natural science from its positivist strait-jacket. In all these ways (positive or negative) the influence of positivism is much wider than the ranks of its adherents.

As to the kind of entity he is tracking down, Bhaskar tells us:

> I shall be treating positivism as an abstract but transfactually efficacious (and so real) cognitive structure, mechanism or apparatus — a real tendency of thought which, when I consider it abstractly, I will simply designate as 'P'. (SRHE, p. 229)

Despite my general dislike of acronyms and abbreviations, I think that this one is useful since it helps us to avoid misidentifying or narrowing the object of Bhaskar's critique.

In the first instance, and according to its own self-image altogether, P is an epistemology. It teaches that

> Particular knowledge is of events sensed in perception; general knowledge is of the patterns such events trace in space and over

> time which, if it is to be possible, must be constant (the Humean theory of causal laws). Sense-perception exhausts the possible objects of knowledge. Conversely any object of sense-perception constitutes a possible object of knowledge. Thus the cognitive claims of theory, metaphysics, morality, ethics, politics, religion and hermeneutics alike are rejected; and man is located squarely as an object within the system of objects in which he acts. P is a limit form of empiricism. (SRHE, p. 230)

Now this epistemology is not only false in content; its claim to be a freestanding epistemology, unsupported by any presupposed ontology or sociology or conception of how philosophy may arrive at such conclusions as it has, is incredible given what we know about the nature of knowledge as at once a social product, a discovery of independent realities, and so on. This raises the question of how a false theory operates and has effects in the neighbourhood of the realities which it misrepresents; how does P function as an ideology for scientific practices which in reality have transitive and intransitive dimensions? In fact, it has tacit and unrecognized accounts of these dimensions. Its sociology of science (transitive dimension) is an individualistic, mechanistic and behaviouristic one, and its ontology is an empirical realist one (a one-level ontology of material objects, despite the tension between such an ontology and its belief in sense-data as the basic building blocks of knowledge). As for 'meta-critical' questions about what philosophical practice yields P as its result, they are radically repressed; Bhaskar includes a section on P's 'ideology in the meta-critical dimension' — a section with a heading but no text (SRHE, p. 292).

P therefore leads a double life: consciously, it is a (false) epistemology; unconsciously it is a (contradictory) system covering much wider issues. It hides its repressed contradictions by a peculiar *inversion* in its tacit accounts of transitive and intransitive dimensions, knowledge and being.

> P cannot sustain either the idea of an independent reality or that of a socially produced science. Instead, these ideas, which constitute preconditions for an adequate account of science, . . . become crossed, resulting in a *de-realized reality* and a *de-socialized science*. (SRHE, p. 252)

Reality is de-realized by a radical form of the epistemic fallacy, phenomenalism, for which we know only our sense-impressions, which constitute 'the empirical world'. Science is de-socialized by the reification of facts, according to which nature, so to speak, presents itself to science on a plate, ready cooked and sliced into 'facts'. Thus the epistemic fallacy is complemented by an 'ontic fallacy', reducing knowledge to its object, effacing its process of production. This role reversal between science and nature is the central feature of P as described by Bhaskar.

This system of errors is presented as *ideology*, in the sense that they are not just mistakes, but ones which function in the interest of a particular social system. Results of the social process of science are presented as uncriticizable natural givens, rather as orthodox economics presents historically specific effects of capitalism as universal features of the human condition. 'Normal science' in Kuhn's sense is reinforced, scientific revolutions inhibited. The discontinuity of science from 'commonsense' knowledge is suppressed, as the specialist character of scientific practice (training, experiment) is passed over, while the special character of its results (closed systems, quantitative explanations) is generalized.

> It [P] can generate an ideology of technocratic expertise and managerial authority as well as, and perhaps to go with, its quasi-egalitarian mystique of commonsense and everyman. (SRHE, p. 272)

Notes

1. See the 1647 'Agreement of the People', in which powers are vested in the elected representatives to do 'whatsoever is not expresly, or implyedly reserved by the represented to themselves' (Morton, ed., *Freedom in Arms*, p. 140). Reserved items include religious liberty and freedom from conscription. My criticism of the Levellers is not that they wished to incorporate constraints on the powers of government in the constitution — a wise move which we would do well to imitate in the UK today; rather that the notion of 'reserving' presupposes pre-political powers possessed by individuals.

2. 'Wovon man nicht sprechen kann, darüber muss man schweigen.' *Tractatus Logico-Philosophicus*, p. 150.

3. However, I have noticed a use of 'concept' in managerial jargon which seems to be analogous with the methodological idiom in using a transitive dimension term to refer to the intransitive dimension, as in 'we must develop the concept of public lavatories', meaning we must build more public lavatories, or more likely, in the current economic climate, 'we must get used to a radical new concept around public lavatories', meaning we must close them down. I don't know whether this usage is widespread enough to cause misreadings of philosophical texts.

4. On the idea that the dynamics of scientific progress refutes the notion that we are trapped within appearances, compare Engels's brief remarks in 'Ludwig Feuerbach and the End of Classical German Philosophy', in Marx and Engels, *Selected Works in One Volume*, p. 605).

5. Since 'epistemic relativism' has been held by virtually every philosopher of the last hundred years without calling it 'relativism', the phrase strikes me as an unnecessary concession to the relativists. But it may be that there is more than a verbal concession since Bhaskar says that epistemic relativism involves rejecting the correspondence theory of truth. Since it is mainly in the social sciences that relativism is a live issue, I shall keep my criticisms of this rejection for the final chapter.

6. Maxwell's criterion of scientific reality:

> If a concept can be embedded in the network of laws such that together they yield alternative definitions of the concept, couched in logically independent terms, not built into its original definition, then we feel that the concept tells us something about reality. (cited by Michael Ruse in his paper 'Definitions of Species in Biology')

7. I am not saying that conflicts between theories *with* shared meanings need *not* be about something. The point is rather that those who deny the intransitive object can still give an account of clashes between theories with shared meanings in terms of logical contradictions between their propositions. Not an *adequate* account though, since the notion of contradiction presupposes the notion of truth.

8. I quarantine the term 'postmodernist' in quotes, not just because of its oxymoronic character, but because the term is applied to phenomena as different as the proverbial 'dog-star and the animal that barks' (Spinoza). In architecture and related disciplines, 'postmodernism' is anti-modernism, and so, in an entirely non-pejorative sense (for it is to be welcomed), reactionary; in literature, it is at most reformist in relation to modernism, with which it has many continuities. In philosophy, it is sometimes one and sometimes the

other. Alasdair MacIntyre gets called a postmodernist on account of that staunchly pre-modernist book *After Virtue*. Rorty's postmodernism is of the reformist kind: his work is an outcome — not to say a reductio ad absurdum — of the 'subjectivistic madness' of post-Cartesian philosophy and of the linguistic idealism of its twentieth-century versions.

9. It may be said that the terms 'voluntarist' and 'discursive' apply also to Kant's idealism itself. But they do so in quite a different — and more metaphysical — way than in the recent remake.

=====4=====

Stratification and Emergence

The Irreducibility of Emergent Strata

We have seen in chapter 2 that nature is *stratified*. *Science* is stratified in that it is divided into distinct *sciences* — physics, chemistry, biology, economics, etc. — which are mutually irreducible, but which are *ordered*. Physics is in this sense more basic than chemistry, which is more basic than biology, which is more basic than the human sciences.

This differentiation and stratification of the sciences is not due to any historical accidents such as which emerged first or how university departments are organized. Considered as a social institution, science may well be divided up partly on the basis of such accidental criteria, but there are also intrinsic divisions based on real stratification of the aspects of nature of which these sciences speak.

I have yet to expound Roy Bhaskar's arguments for this view. But before doing so I shall say more about what is being claimed. The aspect of the stratification of nature most easily grasped is that while everything can be studied (under some description) by physics, and every material substance by chemistry, only some of these things are studied by biology (the vegetable and animal kingdoms), and only some of these again by psychology, and so on. This can be depicted as in Figure 4.1. The relations between the more basic and less basic domains are one-way relations of inclusion: all animals are composed of chemical substances but not all chemical substances are parts of animals, and so on.

This means that animals are governed both by biological and by chemical laws. An animal can do all sorts of things which the chemicals of which it is composed could not do were they obeying not the biological laws governing the organism but only
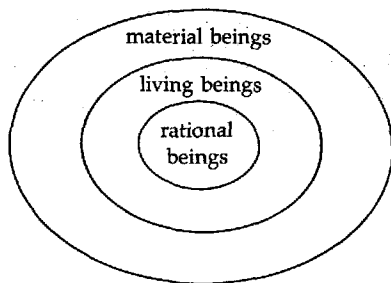
Figure 4.1 Stratification of Kinds of Being

'their own' chemical laws. But of course the animal is not able to *break* the laws of chemistry or physics. As an animal, it has got active powers to do what, as a conglomeration of chemicals, it has only got a passive power to 'do'. If the law of gravity really prevented Macavity (considered as a heavy object) from springing on to the shed roof, there is no way he could do it. It does not, but, considered as a heavy object, he only has the passive power to do it. Considered as a live cat, he has this active power.

Anything belonging to a higher stratum of nature will be governed by more than one kind of law, which is as much as to say more than one kind of mechanism is operating in it. Let us recall here what Bhaskar has said about the objects of the various sciences:

> the predicates 'natural', 'social', 'human', 'physical', 'chemical', 'aerodynamical', 'biological', 'economic', etc. ought not to be regarded as differentiating distinct kinds of *events*, but as differentiating distinct kinds of *mechanisms*. For in the generation of an open-systemic event several of these predicates may be simultaneously applicable. (RTS, p. 119)

As I have pointed out, if it is wrong to apply these predicates to events, it is equally wrong to apply them to entities. For 'biological beings' are also physical beings, and so on. The strictest way to conceive of the stratification of nature is as a stratification of *mechanisms*. There can be no biological mechanisms unless there are chemical ones, while the reverse

does not hold. So at the level of mechanisms, too, there is a one-way hierarchy.

However, at the level of the Actual, relations between strata are not all one-way: they all muck in together. Chemistry can't explain 'more of what happens' than biology — or not necessarily so. Being a more basic stratum does not necessarily mean being a stratum whose effects are more widespread. For though animals are *governed* by zoological laws while inanimate things are not, anything and everything may be *effected* by zoological laws, since animals have effects on the inanimate world. It is because of the cat's powers and tendencies that the (inanimate) contents of my dustbin liner are strewn across the pavement.

The schema and passage I quoted on pp. 48–9 indicate the relations between mechanisms at different strata: the theory of electrons and atomic structure *explains* the theory of atomic number and valency. This is not just a relation between two theories; the theories denote real mechanisms and the relation between those theories maps a relation between those mechanisms. These mechanisms are not spatially locatable objects or events; they are tendencies of certain natural kinds, and operate wherever those natural kinds are instantiated. Each mechanism exists at a different stratum of nature, and the one stratum explains the other. Here we have a problem of ambiguity in the word 'explains', which has to serve to refer both to the relation between one theory and another, and to that between the real strata to which the theories refer. If we lived in the Middle Ages it would be natural to refer to the real relation between the mechanisms at different strata as one of cause; the more basic mechanism causes the less. But in modern usage that would be misleading. The two mechanisms are simultaneous, not successive; neither is an event or action; there is no question of one's being the other's *efficient cause* — the only kind of cause with which most modern philosophy is happy. It is important to distinguish the way in which one mechanism explains another (which we may call *vertical explanation*) from the way in which a mechanism plus a stimulus explain an event (*horizontal explanation*).

Having made this distinction, it is also important to avoid the temptation of thinking that a mechanism which explains

another explains it away, so that the higher-level mechanism drops out of the scientific account. In the quoted passage that I am referring to, Bhaskar points out that we are never in fact able to predict a higher-level mechanism from our knowledge of a more basic one. We always have to discover the higher-level mechanism first; it then becomes the phenomenon to be explained in the next stage of ever-deepening scientific knowledge. There is a tendency in empiricist philosophy of science (unavoidable given its actualist assumptions) to deny the status of explanation to any but the most basic explanatory stratum. Explanations in terms of higher-level mechanisms are seen as mere 'explanation-sketches', standing in for explanations not yet achieved. But this misrepresents the development of science. This is so, in the first place, because when an explanatory mechanism is discovered at a given level, a stage in the work of science has been completed — it does not await the discovery of the next layer down before it can claim to have explained its initial explanandum; in the second place, because, if it did have to await that next stage, explanation would never be achieved, since the newly discovered mechanism always lies open to a further explanation in terms of a deeper stratum in the next stage of scientific discovery. We never reach rock-bottom — so the prejudice that only rock-bottom explanations are real ones would leave us for ever without real explanations. Finally, the fact that a particular explanation (in terms of a mechanism at a given stratum) is itself a good candidate for an explanandum is a point in its favour; an explanation for which the prospects of finding a deeper explanation look bleak is unlikely to be a true one. Therefore, far from rendering an explanation redundant, a deeper explanation underwrites it and reinforces its position in the structure of science.

Bhaskar refers to the relation between a higher-level mechanism and the underlying one in terms of *rootedness* and *emergence*. The higher-level one is rooted in, and emergent from, the more basic one. The term 'emergence' has a philosophical history that indicates that Bhaskar does not regard rootedness as reducibility. Emergence theories are those that, while recognizing that the more complex aspects of reality (e.g. life, mind) presuppose the less complex (e.g. matter), also insist that they have features which are irreducible, i.e. cannot be thought in concepts appropriate to the less complex levels — and that not because of any subjective constraints on our thought, but because of the inherent nature of the emergent strata.

Now some emergence theories postulate a teleology or goal-directedness implicit in the lower strata, by which they tend to give rise to higher strata (e.g. the 'anthropic principle'). Bhaskar makes no such postulates. Higher strata may have arisen, for all we know, by accident; their historical emergence may be accounted for entirely as a by-product of processes describable in fully lower-level terms. But that does not mean that the emergent strata are 'nothing but' those from which they emerged.

> Let us suppose that we could explain the emergence of organic life in terms of the physical and chemical elements out of which organic things were formed and perhaps even reproduce this process in the laboratory. Now would biologists lose their object of inquiry? Would living things cease to be real? Our apprehension of them unmasked as an illusion? No, for in as much as living things were capable of acting back on the materials out of which they were formed, biology would not be otiose. For a knowledge of biological structures and principles would still be necessary to account for any determinate state of the physical world. Whatever is capable of producing a physical effect is real and a proper object of scientific study. (RTS, p. 113)

In other words, the description of the world even in purely physical terms will be different from what it would have been had no living creatures existed and hence no biological mechanisms operated. This difference cannot be explained in purely physical terms: reference to biological laws is an essential part of its explanation. Thus if we want to explain the proliferation of brightly coloured objects (flowers, colourful birds, etc.) at a certain stage of natural history we have to appeal to the laws of natural selection (the value of colours for pollination, mating, etc.); the laws of physics will tell us nothing.

Emergence theories such as Bhaskar's are fighting on two fronts: against dualist or pluralist theories which assert the complete independence of higher strata on lower, and against reductionists who assert the ultimate unreality of the higher strata. He distinguishes three possible senses of 'reduction', two

of which are acceptable to him though not to a dualist, and one of which is unacceptable to him but acceptable to a reductionist.

> There is first the idea of some lower-order or microscopic domain providing a *basis* for the existence of some higher-order property or power; as for example the neuro-physiological organization of human beings may be said to provide a basis for their power of speech. There is secondly the idea that one might be able to *explain* the principles of the higher-order science in terms of those of the lower one. This depends upon being able to undertake at least a partial translation of the terms of the two domains. ... There is finally the sense in which it is suggested that from a knowledge of the states and principles of the lower-order science we might be able to predict behaviour in the higher-order domain. It is important to see that it is to this claim that the strong actualist is committed, if he is to eliminate complex behaviour in favour of its atomistic surrogates. (RTS, p. 115)

Bhaskar has a convincing empirical case against the 'strong actualist' (i.e. one who believes all reality can be reduced to a single basic stratum at which actualism holds). For predictions of this last type have not been forthcoming, and translations of the type which have have never led to the higher-order language becoming redundant (for reasons to be discussed shortly). There is also an argument from the fact that the scientific activities which alone give us access to the more basic strata themselves belong to the higher strata:

> The only way of reconciling experimental activity with the empiricist notion of law is to regard it as an illusion; that is, to regard actions performed in it as subsumable in principle under a complete atomistic state-description. ... Now this has the absurd consequence that the apparent discovery of natural laws depends upon the prior reduction of social to natural science. Or to put it another way, in an actualist world there would be no way of discovering laws which did not already presuppose a knowledge of them. (RTS, p. 116)

I am not sure that this argument, in its present form, is as watertight as it seems. One might have a pre-scientific, provisional explanation-sketch of scientific activity — or indeed one might have scientific activity without theorizing that activity at all; having arrived by that activity at a rock-bottom

explanatory science, one might afterwards redescribe the scientific activity in terms of the newly discovered laws, 'throwing away the ladder one had climbed'. But this will not really save the reductionist's case, for one could not *justify* the scientific activity under its physicalistic redescription. The unreduced description of the scientific activity *as* a scientific activity is necessary for that.

Now let us consider Bhaskar's arguments against the reducibility of emergent strata to those in which they are rooted. He sums them up (RTS, p. 181) as (a) 'the need for a well-defined reductans', and by claiming (b) 'that a reduction left the reality of the higher-order entities intact, at least in as much as they were causal agents capable of acting back on the materials out of which they are formed'. The former means that the emergent stratum could not be predicted or constructed from the one in which it is rooted; on the contrary, only when the emergent stratum has itself been well described can it be explained in terms of a more basic one. We could never predict consciousness from however highly developed a neuro-physiology. We may *now* be able to ascribe it to an organism on the basis of a knowledge of that organism's neurophysiology, but only because we have first learnt what consciousness is and then correlated it with and explained it by certain neuro-physiological formations. If we did not already know about consciousness we would be quite in the dark about what neurophysiology was explaining and why certain neurophysiological facts were significant. I take it that it is at least a fact of the history of science that discovery does proceed in this direction.

The reductionist's next move might be to say that while we need to describe the emergent stratum *first*, once we have explained it we can substitute the more basic description for the emergent one. Once we have learnt that pain is explained by C fibres firing, we can say 'my C fibres are firing' instead of 'I am in pain', just as we have earlier learnt to say 'I am in pain' instead of 'yaaaah!'

Bhaskar's response to this in the passage in question is, I think, elliptical. He says that insofar as the emergent stratum acts upon the stratum out of which it was formed and has effects in it, it cannot be dispensed with: because I am in pain, I take

aspirins, which affect my neurophysiology. This certainly gives some kind of criterion of the reality of a stratum — i.e. if a postulated stratum has effects at a real stratum, it too is real. We have already seen that there can be causal as well as perceptual criteria for reality (and indeed perception itself is only evidence for the reality of its objects because it is caused by them).

However, the reductionist may admit that my pain causes me to take the aspirin, but say that that is just a loose way of saying that my C fibres firing causes me to take the aspirin. Our reply must be along the lines: it only has that effect because it is pain, not because it is C fibres firing; just as, if I read a joke and laugh, I laugh because it is a joke, not because of the configuration of ink on paper. This latter example has the advantage of watertightness — no one will deny the effectivity of the joke, whether written or spoken, and the inefficacy of the inkmarks seen by an illiterate. But it has the disadvantage that the 'vertical explanation' of humour will certainly not be in terms of ink and paper. The vertical explanation of pain, on the other hand, will be in terms of C fibres firing. But it is plausible that the effects of pain are rather like the effects of a joke in that it is features mentioned in the mentalistic description of pain, and not in its neurophysiological one, that make me take an aspirin. Simply knowing that my C fibres were firing would no more make me take an aspirin than seeing a joke written in Chinese (which I cannot read) would make me laugh.

Roy Bhaskar uses the following example in PIF (p. 48), arguing against Rorty's statement that 'Physicalism is probably right in saying that we shall someday be able, "in principle", to predict every movement of a person's body (including those of his larynx and his writing hand) by reference to microstructures within his body' (*Philosophy and the Mirror of Nature*, p. 354):

> Suppose A goes into a newsagent's and says to the proprietor B, 'The *Guardian*, please', and B hands him a copy of it.

Physicalism as stated by Rorty seems to involve saying that B would have handed A a *Guardian*

> even if A had performed some quite different action, such as asking for the *Independent* or for a packet of chewing gum or B to marry him or dancing a jig, and even if A had not been present at all.

Bhaskar's conclusion from the absurdity of such a supposition is that a person's neurophysiology is not a closed system — it is affected by that person's interaction with others. That is true enough. But the further point can be made that the social and semantic nature of the interaction between A and B is irreducibly what causes the action. No physicalistic description of sound waves proceeding from A's vocal cords to B's eardrums would explain how A got the *Guardian*.

Suppose the reductionist says: but we can translate the socio-semantic description of the interaction into a physicalistic one, and whatever X explains, a translation of X explains.

But here the metaphor of translation breaks down. A translation from one language into another gives, in the ideal case, the same sense. The translation can be understood without any reference to the original. But the case here is more like speaking in a language in which one is not fluent, such that one composes sentences in one's mother-tongue and translates them mentally before speaking. For we are not fluent in neurophysiologese, and can only explain social interaction in its terms by translating into it word for word from social-interactionese. And we could not acquire fluency in the former, for it is not just a matter of our greater familiarity with the latter. Here Bhaskar's 'causal criterion' is crucial. It is features of the A–B interchange which are only picked up under a social-interactionese description that explains what happens, just as a joke translated may lose the power to cause laughter. There is nothing funny, in English, about the Italian lady's response to Napoleon's statement that all Italians danced badly: 'Not all, but a good part' (*buona parte*).

On Living in a Stratified World

Stratification and Composition

I have just noted a distinction between two ways in which one stratum may presuppose another: the written joke presupposes the chemical reality of inkmarks on paper, but it is not *rooted in* or *emergent from* them, in that they do not explain the joke. If we

want a vertical explanation of the joke (i.e. an account in terms of some more basic level of reality of why it is a joke) we would have to go, perhaps, to Freud's theory of primary processes, etc. in *Jokes and Their Relation to the Unconscious*. In general, a mechanism at one level will presuppose many levels — perhaps all levels that are more foundational than, or 'below', it. But it will usually be rooted in one or at most two levels; society, for instance, is rooted in biology and not in physics or chemistry, though it presupposes the reality of the world that physics and chemistry explain to us. Rootedness is a relation a stratum has to the one directly below it, or occasionally perhaps to two adjacent strata immediately below it (as it is arguable that psychology is rooted in both the biological and the social levels).

Having made these distinctions, it is worth pointing out that many (though not all) cases of rootedness — emergence relations are also relations of *composition*. Biological organisms, for instance, are composed of chemical substances. It is because they are so composed that they are rooted in chemistry. But they are also emergent from it: they obey laws other than chemical laws, and can do things that could never have been predicted from chemical laws alone.

From what has been said about relations between adjacent strata, it will be clear that the biological organism and the molecules of which it is composed will each be governed by its own set of laws, biological or chemical, which are mutually irreducible. And the same will apply in other cases of part—whole relations, for instance (if it *is* a part—whole relation, which is doubtful), people and society. People are not mere aspects of society, as absolute idealist philosophy has sometimes taught, nor is society a mere collection of people, as the methodological individualists hold. Laws of human behaviour and of social processes will be distinct, and it will not be possible to reduce one to or predict one from the other. Each level is autonomous in the sense of having its own irreducible set of mechanisms, and distinct sciences using different concepts and discovering different laws will be required to study them.

This immediately rules out two methodologies which have often seemed the only alternatives, and have consequently divided the field between them: atomism, which claims that a

reality is only understood when it is resolved into its smallest components; and holism, which makes the opposite claim, that the part is only, and is entirely, explicable in terms of the whole of which it is a part. Both these programmes are *reductive*, denying the autonomy of one or other level. Indeed, if either is taken as a methodology with general validity, it will generate a regress, reducing wholes to greater wholes of which they are parts, or parts to smaller parts of which they are composed, until we reach either One Big Whole (Schelling's Absolute, a night in which, as Hegel said, all cows are black), or a mass of literal atoms, differing only numerically and related only mechanically.

Hostility to reductionism (an entirely legitimate hostility) has most often been directed against atomistic reductionism, so that the threat from the holistic side is often played down. There are exceptions, such as Kierkegaard's protest against Hegel's (alleged) holistic reduction of the individual. But when Marxism, or certain Marxists, are accused of reductionism the accusation is almost always coupled with that of 'mechanism' (generally an atomistic philosophy), though in fact I think it is precisely the least mechanistic — most holistic — Marxists who are guilty not only of that 'suppression of particularity' (Sartre) which says that Flaubert is a petty bourgeois without noticing that not every petty bourgeois is Flaubert, but also of the reduction of geography and biology to the social (with consequent ecological insensitivity). The 'mechanists' Engels and Timpanaro give both nature and human individuality a much better run for their money.

As against atomism and holism, Bhaskar's emergence theory allows us to conceive of real, irreducible wholes which are both composed of parts that are themselves real irreducible wholes, and are in turn parts of larger wholes, with each level of this hierarchy of composition having its own peculiar mechanisms and emergent powers. This in turn allows us to understand dysfunctions in those wholes that are functional in character, since the parts are not pure functions of the whole, but go their own way as well. One word for such a theory of potentially dysfunctional wholes is 'dialectic'.

'Holism', originally a technical term coined by General Smuts, has come into popular parlance in connection with 'holistic'

medicine, etc. It is worth noting that the rejection of holistic methodologies does not discredit such practical holism a priori, nor does rejection of atomism vindicate it. Rather, since this 'dialectical' conception both views a human being as a real whole, and asserts the autonomous reality of our parts, it allows for the coexistence of holistic and analytical approaches. Of course, it does not pronounce on any particular medical or other discipline; they must prove themselves in practice.

Stratified Freedom

> Human freedom, on this view, *if* it exists, would not be something that somehow cheats science (as it is normally conceived) or, on the other hand, something that belongs in a realm apart from science; but something whose basis would have to be scientifically understood.
>
> (RTS, p. 112)

> For science to be possible men must be free in the specific sense of being able to act according to a plan, e.g. in the experimental testing of a scientific hypothesis. (RTS, p. 117)

The theory of stratification makes it possible to situate a freedom which is compatible with these two statements. We have emergent powers, not reducible to physics or chemistry or physiology. These involve, among other things, the power to act on the ground of reasons. If we did not have such a power, we could not make experiments, and science itself would be undermined. If our reasoned actions in planning and carrying out experiments were reducible to physiological reflexes, they could not be assessed as rational or irrational procedures, good or bad or irrelevant experiments, for it is precisely the presence in them of rational as opposed to merely physiological mechanisms which makes them subject to such judgements. It is not enough that they *can* be described in rational terms; it must be features captured only in such a description which are effective in making them what they are. In other words, our rational powers must be genuinely emergent if scientific work is

to be possible. These powers must be *causes* of some of our activities, and irreducible to any subrational causes.

What kind of freedom does the possibility of science demand, and the emergence theory allow? 'Liberty of spontaneity', certainly — i.e. the power to act in accordance with our own natures (in this case, as rational beings), rather than being constrained by the nature of something alien. 'Liberty of spontaneity' has sometimes been contrasted (e.g. by Leibniz) with 'liberty of indifference', i.e. freedom to choose between alternatives with nothing, inside or outside of us, making us choose one way rather than the other. This 'liberty of indifference', which seems to me (as it seemed to Spinoza and Hume) more like randomness than like any liberty we would care to have, is not required as a condition of the possibility of science; and neither, I think, is it shown to be possible by the theory of stratification and emergence, since what emerges are new causal powers, not causal gaps. So to return to liberty of spontaneity: it will be useful to say something about what is unique to us about the sort of liberty we have got, and what is shared with other creatures.

Emergent powers exist at *every* stratum. There is a sense in which a tree is 'free' from mechanical determination. It doesn't *break* mechanical laws (neither do we) but it grows according to its own nature in ways impossible for something subject *only* to mechanical laws, and it has effects at all strata (on lower strata through its transformation of matter, but also by making possible birds' nests, human aesthetic wonder, and so on). It has therefore got a degree of 'liberty of spontaneity'. And this is very much more obvious in the case of an animal, since it is conscious, mobile, and has the power to initiate action relatively independently of external stimuli, and to relate to its own kind.

Possession of liberty of spontaneity is not (as liberty of indifference is) an all-or-nothing thing. It comes with the possibility of restraint, unfreedom, oppression. Hence it makes sense to talk of the freedom and oppression and liberation of animals,[1] as it would not if they were Cartesian machines. To treat animals *as if* they were machines, denying their emergent powers as is done in factory farming, is therefore an infraction of natural freedom and so prima facie objectionable. To a degree, it even makes sense to talk of restraint of the freedom of plants,

rivers, etc. But liberty of spontaneity is quite a different thing at each stratum, and generates a correspondingly different notion of what is harmful in treatment of creatures governed by the laws of that stratum.

To return to ourselves (humans): our liberty of spontaneity does not exist at one level only. We are living beings, primates, and part of our proper freedom is that proper to primates. We are also social beings, members of organized societies who have many powers attendant on being such members. Society is an emergent stratum — societies are governed by social laws not natural ones, and may to a degree control by their organization the mechanisms presupposed and the forces generated by social existence. A society that is highly organized and takes authoritative decisions may be freer than one which lets things take their course, just as the fish that can swim upstream is freer than one that can only float down it. And as members of society, we may participate in and benefit from the freedom inherent in such organization and authority. So that at one level, the power and freedom of the individual varies in proportion to the power and freedom of society, not inversely with it as liberalism supposes. However, as individual agents we have powers which are not reducible to, but emergent from, biological and social ones, e.g. the power to listen and talk, to reason and to act upon our reasons. We have the power to act on our reasons whether they are good ones or bad ones, and also the self-critical power to put right our bad reasons in some measure.

Each stratum in which we have our being has its own proper powers, irreducible to those at lower levels, and generating a liberty of spontaneity which can be subject to greater or less limitation or transgression. So our freedom is a complex thing — stratified freedom, in fact. And there is no in-principle-untranscendable 'ceiling' to the levels we inhabit. If I have stopped this upwards-moving survey at the level of the reasoning powers of the individual, that does not mean that I deny the possibility of higher strata with their corresponding freedoms. Just as we cannot know that physics (our most basic science at present) is in any absolute sense a 'rock-bottom', so there is no a priori limit to the emergence of higher strata. But for reasons which will become clear in what follows, it may be difficult to say anything precise about any such strata.

Problems of Science in Higher Strata

A stratified world, as we have seen, is an open world — a world which does not naturally produce closed systems. But it is also a world in which we can produce closed systems *at some strata*. The 'lower' the strata in the hierarchy of rootedness and emergence, the closer we get to a closed system. For it is possible to isolate, for instance, a chemical process from the interruptions of organic processes, but it is not possible to isolate an organic process from the effects of chemical processes, since it is rooted in them. A science such as evolutionary biology deals with systems which are inherently and in principle open, since, on the one hand, 'random' mutations (i.e. those determined by a purely physico-chemical process, not a biological one) are presupposed; and, on the other hand, the environment which determines what constitutes 'fitness' is governed by geological, meteorological and social processes as well as biological ones. It is often possible to isolate a system from processes generated by 'higher' strata, but never possible to isolate one from those generated by 'lower' strata. Hence the further up the hierarchy we go, the more distant our approximations to closure become.

Naturally this presents problems for the scientific study of the upper strata. Since these problems for the most part affect the human sciences, much of the discussion of them must be held over to Part II of this book, on critical naturalism. However, these problems are not generated by the distinction between the natural and the human sciences as such. Some natural sciences are higher up the hierarchy than others, and hence further from closure. Likewise some human sciences are higher than others, and further from closure. Granted that all human sciences are further up than all natural, there is nevertheless not one Great Divide, but rather many gradations. Evolutionary biology has no experiments, as physics and chemistry have, and while it has considerable explanatory power, it has little predictive power. Among the human 'sciences', economics and (since Chomsky) linguistics can at least claim scientific status without making people laugh. It is doubtful if as much can be said for psychology or semiology. (I have elsewhere coined the term 'epistemoids' for approximations to sciences in these areas, but since Roy

Bhaskar has not adopted this usage, I shall say no more about it here.[2])

Explanation in Open Systems

In the remainder of this section I shall discuss what Bhaskar says about explanation in open systems, and how the idea of the stratification of nature affects it.

He notes (RTS, pp. 121–2) that the typical causal explanation in ordinary life is of a 'transitive verb' kind: 'Tania pushed the door open.' This differs from 'Tania observed the door open' in that it is a causal statement, and from 'Tania pushed the door hard' in that it implies that the door *is* open at the end of the process. The explanation is deductive, and involves action-by-contact, as in classical mechanics. However, this model needs to be modified since in open systems there will be a multiplicity of causes. Bhaskar gives an example (p. 123) of a historical narrative in which a multiplicity of transitive verbs maps a complex causal sequence. The pattern of explanation of such complex sequences is a four-stage one, which Bhaskar calls RRRE: resolution, redescription, retrodiction, elimination.

> *Resolution*: the process is analysed into its various causal components;
> *redescription*: granted that we have a background of theory about the various mechanisms operative in this open system, we can redescribe the causal components in terms of this theory. We will then be in a position to
> *retrodict* the causes of these components. However, since we are in an open system, there may be any number of possible causes that could have codetermined these events. We need to
> *eliminate* such of these as we can, by means of independent evidence about the antecedent events. Skills in RRRE are precisely the skills of a detective (preferably assisted by a forensic laboratory).

Although Bhaskar's example is from a historical narrative, the RRRE model is meant to apply in all open-systemic disciplines.

Hence the applied scientist, in the natural sciences too, has to have quite different skills from those of the pure scientist, who is trained in theory and experiment rather than in RRRE.

> The applied scientist must be adept at analysing a situation as a whole, of thinking at several different levels at once, recognizing clues, piecing together diverse bits of information and assessing the likely outcomes of various courses of action. (RTS, p. 120)

While the pure scientist abstracts from levels of reality he or she is not concerned with, the applied scientist always keeps a weather eye on them.

It may be noted in passing that many of the *mis*applications of science that are so hazardous to health and the environment are cases in which the open-system-oriented skills of the applied scientist have *not* been duly exercised: where the pure scientists' discovery has been transferred straight from the laboratory to the factory, without proper consideration of the 'side effects' of the discovered mechanism on other strata.

Finally, on the stage of *redescription*: this seems to presuppose some pure-science account of the causal mechanisms involved. These may of course be at a variety of strata. And mechanisms at lower strata than the ones immediately concerned may be relevant, in that the active powers available to an entity under one stratum cannot exceed the passive powers it possesses at lower strata (e.g. an animal can't do what its chemical components can't have done to them, etc.). Nevertheless, if we are explaining a process in human history and the only pure sciences that we can draw on are the natural sciences, we are going to be very under-informed at the redescription stage. This problem will be discussed in chapter 5.

Now I shall pass to the discussion of what seem to me several unresolved problems about stratified determination.

Problems about Tendencies, Conditions and Determinism

In an appendix to chapter 3 of RTS, Roy Bhaskar distinguishes two kinds of tendency: we are already familiar with the first, a

'power normically qualified' (p. 229) such as (perhaps) dogs tend to bark when they hear burglars breaking in. The second kind (tendency$_2$), Bhaskar metaphorically (as he puts it) characterizes as 'ontological preference'.

> To attribute a tendency (in the second sense) is not just to normically qualify the exercise of a power; but to say that some of the intrinsic enabling conditions of a relatively enduring kind for the power's exercise are satisfied; that the thing is predisposed or oriented towards doing it, that it is in something of a state or condition to do it. (p. 230)

Fido tends$_2$ to bark — i.e. he is a vociferous cur. Bhaskar notes that a tendency$_2$ is typically ascribed to *some* of those possessing the tendency$_1$ to do the same, as in this case, in which some dogs tend$_2$ to do what all dogs tend$_1$ to do, i.e. bark.

This distinction is intuitively plausible, but I doubt if it can be consistently maintained as a distinction of kind outside of human affairs. It is the distinction between, for example, someone who will do something if appropriate circumstances arise, and someone who will do it given half a chance. Bill will go to the pub if a good friend calls and invites him, he hasn't got any more pressing engagements, and he is not short of money; Joe will go there anyway unless he is skint. Where human actions are concerned, the distinction can I think be maintained in terms of desires: Bill's desire to go to the pub is aroused by events, Joe's is never absent. But when applied to non-humans, I think it breaks down:

> the way in which Fido expresses his generic tendency$_2$ to bark may depend upon the particular factor that excites it. He may bark viciously at an intruder but conventionally at the postman, fearfully at the moon but affectionately at another dog, arrogantly at a cat and playfully at an old shoe. (p. 233)

And if Gelert barks only at intruders, we are probably inclined to attribute only a tendency$_1$. Yet I can see no difference in principle. *Maybe* Fido is just waiting for the excuse to bark, but then maybe Gelert is too, only he requires a better excuse.

It is still clearer that the distinction breaks down for inanimate objects. The match has the tendency to light if struck, but a piece of phosphorus will ignite if removed from the water. It is easy for us to think anthropomorphically of the phosphorus being *prevented* from lighting while the match needs a special stimulus to make it do so. But one can imagine a world in which everything was normally under water, and we would say about phosphorus what we now say about the match.

In both cases, things will exercise a tendency in appropriate circumstances, and not otherwise. There is a difference of degree only, and external to the nature of the thing, between the two cases. In short it is *only* the anthropomorphic metaphor that sustains the plausibility of the distinction.

I think this should alert us to a more serious problem in Roy Bhaskar's account of causality. He wants to make an ontological, not just an epistemic or methodological, distinction between causes and conditions. For while he does say in one place 'conditions is an epistemic, not an ontological category' (RTS, p. 78), he immediately says something which implies the opposite: 'conditions change, but they do not have the power to change'. Whether a kind of entity has a given kind of power or not is not an epistemic question. And elsewhere, he always treats the causes/conditions distinction as ontologically grounded.

First let us note that the usage in which we refer to *the* cause of an event is prompted by practical concerns, it is not constrained by scientific ones. 'The cause of the fire was a lighted cigarette end dropped on the carpet' (the combustible carpet was only a condition). But if it were an *ashtray* which was so readily combustible that it caught light when a cigarette was put in it, it would be the combustibility of the material that was 'the cause' of the fire. In itself, aside from our forensic or other practical concerns, there is nothing that is 'the cause', only *causes*. And these include 'conditions'.

The model of causation that separates causes and conditions may be called the *interventionist* model. It is a model of idling conditions suddenly stimulated into operation by an agent: there is a living room, wood laid out in the grate, Susan sitting in her armchair; then Susan gets up, strikes a match and lights the fire, bringing the wood's tendency to burn into play.

But in reality Susan is no 'unmoved mover'; the room has been getting chillier, Susan has been getting less and less comfortable, till she is moved to action. If tendencies were not always in motion, nothing would ever get started at all. But many of these tendencies would naturally be classed as conditions (the chilly room). And cases of human interventions into nature, like this fire-lighting one, are the cases in which the causes/conditions distinction seems most plausible. In nature itself, most causal processes are almost imperceptible operations of almost indiscernible, and complexly interacting, tendencies. And the same applies at the macro-social level. Outstanding human actions may appear like 'causes' in the midst of 'conditions': Caesar crossing the Rubicon, Luther nailing up his Ninety-five Theses, the storming of the Winter Palace. But the 'conditions' were themselves composed of thousands of (largely routine) human actions in which (easily overlooked) tendencies operated.

If it is at all possible to produce an account which saves the causes/conditions distinction, that account must make a place for three points arising from this discussion.

1. Causes have causes; there are no 'first causes' on earth.
2. A cause of the obvious 'agent intervening' kind may have as *its* cause a long-operating tendency which, to the casual view, fades into the furniture of 'conditions'.
3. That which, in relation to a given cause, may appear as a mere condition may in itself be the site of miscellaneous tendencies at work.

This suggests that the notion of conditions is a relative one: conditions are such in relation to some agent's intervention to which they form the background; in themselves, they always involve tendencies already at work, which will codetermine the outcome with the 'cause'. The limit-case of *pure* conditions — in which no tendency is actualized — cannot be instantiated: the law of inertia is itself a tendency. I am not sure whether even this relative distinction can be given a definition unrelated to our expectations or interests.

Before passing on to consider Bhaskar's critique of determinism, though, something should be said in mitigation of

the above criticisms. The project of providing an adequate philosophy for science, for the human sciences, and for the work of human emancipation that can be based on them, requires, according to Roy Bhaskar, more and subtler distinctions to be made than those offered by classical idealist or empiricist theories. In distinguishing between different kinds of causal factor, tendency, power, etc., he is avoiding over-simplifications and mapping what is perhaps a gradation of realities. There may be distinctions of degree here, even if not distinctions of kind. And it is possible (as I have suggested in the case of tendency$_1$ and tendency$_2$) that what is only a distinction of degree in the natural world may become a distinction of kind in the human world.

Now let us consider the various determinisms that Bhaskar discusses. In the first place, there is one kind that he accepts: 'ubiquity determinism'. This is a version of the old principle that nothing happens without a cause, the 'principle of sufficient reason'. It should be said that this is not to be regarded as an a priori ontological principle in Bhaskar's philosophy, for it contains no such principles. The case is more like a Kantian 'regulative idea', that we must treat 'as if' true for the purposes of cognitive inquiry. It is possible that there are aspects of nature in which indeterminacy holds — in which 'God plays dice', in spite of Einstein. But we could not *know* that there are, though we might suspect it; we continue to regard unanswered causal questions as just that — not as proven cases of indeterminacy. (Of course indeterminacy, as a theory in physics, is beyond the scope of philosophical criticism, though Bhaskar is right to reject appeals to it in making the case against determinism. If the stake in the debate about determinism is the real effectiveness of our rational powers, indeterminacy is quite irrelevant to that debate.)

So much for ubiquity determinism. But there are other versions of determinism which, while presupposing ubiquity determinism, make more far-reaching claims. One we have already seen refuted by Bhaskar's arguments, namely actualism. According to this, the actual, observable course of events occurs in constant conjunctions, 'whenever A happens, B happens'. This we have seen to be false, since were it true, scientific experiment would be incomprehensible. But there is another

kind which, while it seeks closure, does so, as it were, with a nostalgia for actualism. This is *regularity determinism* — a doctrine which Bhaskar often slips into calling simply 'determinism'.
Regularity determinism is defined as the view that

For every event y there is an event x or set of events $x_1 \ldots x_n$ such that x or $x_1 \ldots x_n$ and y are regularly conjoined under some set of descriptions. (RTS, p. 69)

In other words, it sees all causation as being describable in terms that in fact apply only in closed systems. That is to say, that for all events (since all are caused) some closed-system-like description can be found. So the regularity determinist holds that, although there is no one kind of event that always precedes, or is always followed by, a squirrel running up a tree, one could in principle spell out the events involved in any given case of a squirrel running up a tree, and its antecedents, in such a way that the one could be explained by the other as an instance of an exceptionless 'whenever A then B' generalization. Such a project would involve including all relevant factors in the description. It involves, so to speak, a purely descriptive accomplishment of closure — closure without experimental intervention. Is such closure possible? It might indeed happen that that a particular set of things constituted a 'closed system' in the sense of being cut off from non-constant external influences (see RTS, p. 69). And unless and until the description reaches such a cut-off point, it is necessary to go on extending it to include all relevant factors. But such a 'closure' would not be enough to set up a closure in Bhaskar's sense; it would not generate constant conjunctions. For such a system might be 'internally open' by virtue of including some complex entity whose behaviour could not be predicted from its environmental stimuli alone (e.g. a squirrel). Any system including (and that means also *affected by*) a person, animal or any other being which is complex enough to initiate action not fully determined by its immediate environment will be an open system. Hence the regularity determinist programme will have to involve eliminating complex entities by reducing them to their atomic components whenever they cannot be physically excluded, just as it will have to 'eliminate' external factors by internalizing

Table 4.1 Limit Conditions for a Closure, i.e. for the Stability of Empirical Relationships

Conditions for a Closure	(1) Epistemically Dominant Case	(2) Epistemically Recessive Case
(A) System	Isolation	Constancy of Extrinsic Conditions
(B) Individuals	Atomicity	Constancy of Intrinsic Conditions
(C) Principle of Organization	Additive	Constancy of Non-Additive Principle

Source: RTS, p. 76.

them, whenever it cannot physically isolate the system. And reducing complex beings to their atomic components will be no use unless a purely additive account can be given of the organization of those atoms. If complex beings have really emergent powers, they cannot be taken into account. The regularity determinist's norm is set out in Table 4.1.

If these conditions are not generally realized in the world, the regularity determinist account of causes collapses. But isolation, atomicity and the additive principle are *not* generally realized, i.e. things generally interact, there are really complex beings which can initiate action without external stimulus, complex beings really do have emergent powers, and so on. So the regularity determinist is thrown back on the second choice, the 'epistemically recessive case'. But

it is easy to see that once an actual isolation and an atomistic description are set up as norms two regresses are initiated, viz. to systems so vast that they exclude nothing and to individuals so minute that they include nothing. (RTS, p. 77)

The case against regularity determinism is a strong one. But it sometimes looks as if Bhaskar is claiming to have shown that the world is even more 'open' than he actually has.

Laws leave the field of the ordinary phenomena of life at least partially open. (RTS, p. 111)

If this is read as saying that *particular* laws leave things open, it is true. But it certainly has not been shown that the *totality* of laws

leaves things open. What is not caused by one mechanism is caused by another. In the end, Bhaskar recognizes this:

> Thus the behaviour of e.g. animate things is not determined by physical laws alone. But that does not mean that their behaviour is not completely determined: only that an area of autonomy is marked out which is the site of a putatively independent science. And because the forms of determination need not fall under the classical paradigm this in turn situates the possibility of various kinds of self-determination (including the possibility that the behaviour of men may be governed by rational principles of action). (RTS, p. 114)

That is indeed what the theory of stratified determination shows. But I think that a very natural reading of the section on 'autonomy and reduction' up to this quoted passage is as depicting a world in which, while things do not *break* the laws of nature, they do behave in ways unaccounted for by any laws. One of my aims in this section has been to show that Bhaskar's 'anti-determinist' arguments should be read with care not to jump to conclusions which are ultimately incompatible with ubiquity determinism. His arguments establish 'liberty of spontaneity' at various levels — animal and human, individual and social. But they do not establish 'liberty of indifference'. For the most part, the impression that they do so comes from the use of words to the effect that laws constrain but do not determine events — expressions which can and should be read as meaning that *particular* laws do so. But Bhaskar also says that it is an error 'to think that because something happened and because it was caused to happen, it had to happen before it was caused' (RTS, p. 107). If my remarks about causes, tendencies and conditions are correct, the phrase 'before it was caused' is a very odd one.

The Battle of the Trees

The idea of the stratification of nature implies that there is a definite answer to the questions how the strata are ordered, which presuppose which, which explain which, etc. Bhaskar

says relatively little about this matter, though the following passage lays down one principle for approaching it:

> Now in general a reduction [i.e. a vertical explanation — A.C.] is possible because the entities in terms of which the behaviour of the thing is explained occupy a different volume of space, either larger or (more usually) smaller. Thus the possibility of a reduction implies in general that the individuals of the different kinds cannot be said to occupy the same place at the same time and one not be part of the other. This gives us a general criterion which imposes limits on regresses of strata, i.e. upon the possibility of a sequence of (explanatory) reductions. For one could define a branch of science as a series of theories within which this criterion is satisfied. On it, quantum mechanics and chemistry would belong to the same branch. But electromagnetism and mechanics, neurophysiology and psychology and (it will be argued) psychology and sociology would belong to different branches. (RTS, pp. 181–2)

This idea of ordering in various branches seems to refer to cases where entities at one stratum are composed of those at another — yet the relation of psychology to sociology, which might be thought to be such a case (insofar as sociology is about societies, psychology is about people, and societies are widely believed to be composed of people), is explicitly excluded. And relations *between* different branches are not discussed, though such relations there must be — psychology and sociology at least ontologically presuppose one another, of which more shortly.

It is helpful to distinguish three relations which may hold between strata. The most general is *ontological presupposition.* One stratum ontologically presupposes another if it could not exist unless the other existed. I have suggested, I hope uncontroversially, that biological strata[3] ontologically presuppose physical and chemical ones, and social strata ontologically presuppose biological ones. In these cases, the presupposed strata existed first, but this is not necessarily the case: two strata, one of which ontologically pre-supposes the other, may have come into being simultaneously. This is important as it allows for the possibility that two strata may ontologically presuppose each other.

Second, there is the relation of *vertical explanation*: mechanisms at one stratum explain those at another. Obviously,

?
psychological and semiological sciences
social sciences
biological sciences
molecular sciences
?

Figure 4.2 Tree of Sciences and Their Objects
(*Source: Scientific Realism and Socialist Thought*, p. 45)

the explained stratum in such a case ontologically presupposes the explaining one, and all those that the latter ontologically presupposes (ontological presupposition is a transitive relation). But vertical explanation is not a transitive relation: chemistry explains biology and biology explains sociology, but chemistry does not explain sociology. Third, we have seen that there is the relation of *composition*: living organisms are composed of chemical compounds, etc. Clearly any stratum ontologically presupposes any from which it draws its components. In some cases (as in the last mentioned) the component stratum also vertically explains the composite one, but that is not always the case, as we shall see.

In my book *Scientific Realism and Socialist Thought*, I suggest the tree of the sciences (and their objects) illustrated in Figure 4.2. At a conference on 'Realism in the Human Sciences' (Southampton, 1990), Caroline New pointed out that an alternative tree has been suggested (with some plausibility, since it seems closer to the 'tree of composition') in which the order of the top two layers is reversed. She herself proposed situating those two on a level, as separate branches. How should we go about resolving these questions?

My reason for proposing the ordering in Figure 4.2 (which I did introduce with the clause: 'Oversimplifying a lot, it would presumably contain some such ordering as this') was that I think that some psychological and semiological mechanisms are vertically explained by social ones, but not vice versa. Social mechanisms are not what they are because of any psychological or semiological ones. This does not mean that the nature of human individuals places no constraints on what social mechanisms can exist, but it is our *biological* nature that does so.

Facts such as our long childhood and our need to produce in ways not given by instinct explain some of the more general mechanisms of human society. Of course there is no question of every social mechanism having a direct biological explanation: most of them are specific to particular *kinds* of society. Attempts to vertically explain social mechanisms in terms of psychological ones, on the other hand (e.g. 'applied psychoanalysis'), are all wildly implausible, while many psychological mechanisms are specific to certain societies, and hence presumably socially determined. Of course all these statements are empirical generalizations, and this is not the place to prove them; I only mean to make it clear that my ordering principle is vertical explanation.

But the relations between the levels are more complex than this suggests. The vertical explanation of the psychological and semiological levels must be in terms of biological as well as social mechanisms. And where ontological presupposition is concerned, the social, psychological and semiological levels all ontologically presuppose each other. (I certainly did not mean to suggest that society *existed before* language or mind; the three could only emerge together.)

Finally, as regards composition: it is not implausible to say that society is (at least partly) composed of people (though there is much more to be said about this in the next chapter). So Figure 4.2 suggests that vertical explanation is not always from parts to whole, it can be from whole to parts. In the case of language, the levels to which its parts belong seem to have no particular explanatory role, for it can be composed of virtually any elements, just as long as they can be distinguished from each other: sounds, gestures, letters, ideographs, and all with infinite possible variations. Language is vertically explained not by its elements but by biology and society; it ontologically presupposes *some* elements, of course, but since any elements would do, it does not presuppose any particular ones. These components get their value entirely from the structure of the language. In that sense, language comes as near as can be to a holistic system.[4]

I conclude that any realistic 'tree' is going to be an untidy one, and perhaps not representable in two dimensions, since the order of composition does not coincide with the order of vertical

explanation; some strata are vertically explained by more than one other stratum; and relations of ontological presupposition are not all one-way. I hope this discussion will have prepared the ground for an account of Roy Bhaskar's social ontology, to which the next chapter is devoted.

Notes

1. In defending the philosophical coherence of the notion of animal liberation, I am not committing myself to the movement that calls itself by that name. In general, the measures proposed under the name of 'animal rights' are much more defensible, though that concept itself is less so (philosophically speaking).

2. See chapter 4 of my *Scientific Realism and Socialist Thought*, which Roy Bhaskar refers to in the Postscript to PN (p. 176, second edition).

3. Throughout this section I use the 'methodological idiom' (see p. 43) freely. But I don't think it creates any damaging ambiguities, since the hierarchy of sciences maps the hierarchy of strata, and whatever can be said about one, something parallel can be said about the other.

4. For this reason, linguistic structures should not be used as a model for the understanding of social structures: the error of structuralism.

==== *PART II* ====

Critical Naturalism

Social Being and Social Knowledge

From a certain point of view, it might seem strange to place sections on social being before sections on social knowledge. Since Descartes, it has been customary first to ask how we can know, and only afterwards what it is that we can know. But this Cartesian ordering has been a contributory factor to the prevalence of the epistemic fallacy: it is easy to let the question how we can know determine our conception of what there is. And if in a certain respect the epistemic question does seem prior, in another it is secondary to the ontological one: knowledge exists as an aspect of our being in the world, and before we can know how we know, we need to have some idea how we interact with that world in such a way as to acquire knowledge of it.

Bhaskar states at the outset of the chapter of PN on 'societies',

> I shall concentrate first on the ontological question of the properties that societies possess, before shifting to the epistemological question of how these properties make them possible objects of knowledge for us. This is not an arbitrary order of development. It reflects the condition that, for transcendental realism, it is the nature of objects that determines their cognitive possibilities for us; that, in nature, it is humanity that is contingent and knowledge, so to speak, accidental. (PN, p. 25)

Of course, in order to ask 'what properties do societies and people possess that might make them possible objects of knowledge for us?' (PN, p. 13), we need to have some idea what societies and people are. For we must be able to distinguish these questions from the questions what language or God or leek soup must be in order for us to have knowledge of them. But we do after

137

all have some implicit knowledge of what societies and people are, since we *are* people and people *are* social beings. It is the task of some transcendental arguments to make this 'connatural knowledge'[1] explicit.

> For the possibility is bound to arise of posing transcendental questions of the form 'what must be the case for ϕ to be possible?' for social practices other than science. (PN, p. 7)

Roy Bhaskar's conception of social being is, I take it, the result of such transcendental arguments. Their premises are observations which would command widespread assent; but they involve transcendental refutations of widely held alternative views of social being.

The Relational Conception of Society

The foregoing chapter will already have led readers to expect that Bhaskar's account of society will not be an atomistic one — that knowledge of society does not reduce to knowledge of people. For the conception of the stratification of nature involves the recognition of real complex wholes, with emergent powers not predictable from the powers of their parts. Insofar as 'methodological individualism' stems from a general assumption that complexity is ultimately unreal, that complex wholes must be resolved into simple parts before they can really be understood, Bhaskar's argument for stratification and emergence has already undermined it. But methodological individualism in social theory has other motives too. Indeed, as Bhaskar points out, if methodological individualism is really just one instance of a general theory of explanatory atomism, the reductive programme would not stop at human individuals, who are themselves complex wholes about whom we may ask the question whether their powers are reducible to those of their simpler components. 'Seldom does it occur to subscribers to this [methodological individualist] view that an identical train of thought logically entails their own reducibility, via the laws and principles of neurophysiology, to the status of inanimate things!' (PN, p. 26). Methodological individualism has mainly

been popular because of the philosophical underpinning which it seems to give to certain substantive (and false, as Bhaskar would argue) social theories: utilitarianism, liberal political theory, and neo-classical economics. But it also gains plausibility from its starting point in the important truth that 'society is made up or consists of — and only of — people' (PN, p. 30), or, as Bhaskar amends this formula '*the material presence of society = persons and the (material) results of their actions*' (PN, p. 30).

Now if one reads methodological individualist texts, one finds that they tend to assume that once this point is conceded, methodological individualism follows (see, e.g., A.G.N. Flew's *Thinking About Social Thinking*, pp. 42–3). This assumption only looks plausible on the further assumption that there is one and only one alternative to methodological individualism, namely some form of methodological collectivism which treats *groups* as the fundamental, irreducible social reality and the 'bedrock' of social explanation. There have sometimes been such theories. Some of the English and Italian Hegelians tended to treat nations as somehow more fundamental than individuals. Among Marxists, Milton Fisk, while resisting the tendency to see individuals as mere aspects of groups, and pointing out its dangerous (fascistic) political consequences, seems to treat groups (i.e. classes) as the fundamental *social* entity (see his *Ethics and Society*).

Bhaskar's case against methodological individualism starts by pointing out that there is a third possible social ontology:

> Sociology is not concerned, as such, with large-scale, mass or group behaviour (conceived as the behaviour of large numbers, masses or groups of individuals). Rather it is concerned, at least paradigmatically, with the persistent *relations* between individuals (and groups), and with the relations between these relations (and between such relations and nature and the products of such relations). (PN, pp. 28–9).

Compare Marx:

> Society does not consist of individuals, but expresses the sum of interrelations, the relations within which these individuals stand. (*Grundrisse*, p. 265)

In the light of this relational model of society, the examples considered by methodological individualists as likely candidates for irreducibly social entities look quite beside the point — riots and orgies, for instance. The use of such examples shows that methodological individualists think that irreducibly social entities could not exist unless individuals had somehow sacrificed their individuality. The relational model suggests quite different examples: one is a worker only because of one's relation to an employer, a husband only because of one's relation to a wife, a buyer only because of one's relation to a seller, a property-owner only because of one's relations to non-owners. Our social being is constituted by relations and our social acts presuppose them. Yet relations and the related individuals may be ontologically independent, in that, for example, the relations between the head of the philosophy department at Southampton and the members of the department existed before I entered into them in a new way by becoming head of department — and of course I also existed before. Such relations can form the subject-matter of specifically *social* sciences (as distinct from the wider category of human sciences), without 'melting down' individual reality. The critic of methodological individualism can deny that there can be de-socialized people without asserting that there could be a de-populated society.

Once it is seen what the relational alternative to atomistic models of society does and does not imply, its superiority is easily shown by its capacity to cope with the counter-examples that have often been presented against methodological individualism: 'A tribesman implies a tribe, the cashing of a cheque a banking system' (PN, p. 28). In short, the problem is 'how one could ever give a non-social (i.e. strictly individual-istic) explanation of individual, at least characteristically human, behaviour!' (PN, p. 28).

If an atomism of individuals is untenable, an 'atomism of relations' would be equally so. Relations presuppose other relations, relations are related to other relations. The lattice-work of relations constitutes the structure of 'society'. It is possible to focus study on the relations (which may endure through changes of the related individuals), or on individuals (who may circulate around the network of relations that is

society). The social sciences and the psychological sciences therefore have distinct subject-matters. But if these two groups of sciences are really to get off the ground, we need to know something about the distinctive kinds of generative mechanism they each discover. Above all we need to know how these two kinds of being (social and personal being), of which neither can exist without the other, can be governed by two kinds of law, and how these different laws are themselves related. Bhaskar gives us such an account, which in turn will provide new grounds for accepting the relational model of social being. He calls this account the *transformational model of social activity* (TMSA). This model forms the core of Bhaskar's social ontology, with which the rest of it stands or falls.

The Transformational Model of Social Activity

As in the case of the relational model of social being, Bhaskar introduces the TMSA polemically, as an alternative to three other models. However, it is not just a plausible alternative: it accounts for the facts by which each of the other models defends itself. It is consistent with *all* these phenomena in a way that none of the other models is. It is the conclusion of a transcendental argument from the premises of its rivals, and hence a transcendental refutation of those rivals. It should be clear by now that this does not mean that it claims to be a priori or indubitable. But it does, I think, show it to be the best available model to date.

There are two opposite conceptions of social explanation related to, though not identical with, methodological individual-ism and collectivism. They may loosely be described as *humanism*, which sees human agency as everything, and *structuralism*, which sees social structure as everything. Each can appeal plausibly to a set of widely agreed facts. Let us consider the 'widely agreed facts', divided into two lists: those that seem by themselves to support humanism and those that seem by themselves to support structuralism, respectively.

(A) Whenever we look at social reality, we see nothing but human actions and their effects: acts of producing, exchanging

and consuming, of voting and issuing decrees and storming the Winter Palace, of investing and going on strike and repossessing houses, of preaching the gospel and committing simony and burning heretics. In every case these actions are carried out by people, separately or in groups. They may sometimes be acting in an official capacity on behalf of an 'artificial' or 'corporate' person, such as a limited company or a government department, but these artificial or corporate persons do nothing without 'natural persons' acting for them. There cannot be a war without soldiers fighting, a slump without bosses sacking workers, inflation without sellers raising their prices, a crime wave without villains assaulting and robbing people, a revolution without people taking to the streets. 'Structures don't take to the streets.'

(B) Whatever happens in society happens as it does because social structures are as they are: one cannot be a producer without a mode of production, or a seller without a market; one cannot vote without an electoral system or pass laws without a constitution. And this is not just a matter of enabling conditions; the social position one occupies largely *determines* what one does: a worker must work, a seller must sell. And the social structure largely determines the developmental tendencies of the society: a capitalist economy *must* progress technologically, undergo inflation and periodic recessions and so on. One can sometimes predict a crisis from the tendencies inherent in the structure, even though everybody is trying to avert one. One can explain the bourgeois outcomes of the English and French Revolutions in terms of the structures of the societies that gave rise to them, irrespective of their 'agents'' ideas about the rule of the people or the saints or the enlightened. One can explore the causes of crime and the conditions of its increase without investigating 'the criminal mind', and, anyway, no knowledge of 'the criminal mind' could tell you why there were more of them about in 1990 than in 1970. In general, facts about human agency don't tell you why people do different things from one epoch to another, while facts about the social structures which differentiate those epochs do.

In describing list (B), I am already slipping into saying things which seem to entail 'structuralism'. For these facts suggest that we could dispense with the concept of human agency for the purposes of social explanation: Robespierre is but the 'bloody instrument', not indeed of Rousseau's philosophy, as Heine thought, but of the balance of class forces in France and Europe. From this point, it is an easy step either to analyse agency, reasons and motives as mere effects of the structure, even its epiphenomena; or to present the conscious belief that we are purposive agents as an illusion that lubricates the wheels of the structurally determined process of which it is an aspect. This last has been the view of some Althusserians, and it is a natural reading of some passages in the work of Althusser himself.

But, on the other hand, if we take group (A) by itself, it would be easy to conclude that society is 'nothing but people', and that the concepts that we need for social explanation are those of purposive action, motivation, rational deliberation, and so on. From this point of view, the phenomena picked up by list (B) are to be seen either as mere shorthand summaries designating the cumulative effects of many human acts; or, if they are admitted to be genuinely explanatory phenomena, this is seen as the result of a kind of 'alienation' or 'reification' by which people *submit* to inherently unreal necessities; unreal in that since there is nothing to (for instance) the state of the market but a mass of individual acts of exchange, that state is wholly created and sustained by our agency, which it consequently cannot constrain. 'Alienation', in this sense, is seen as being like the child who paints a picture of a ghost and then becomes frightened of it.

Humanism and structuralism, as I have presented (and perhaps caricatured) them, can't both be true. But further: there is some difficulty even in making them engage in debate. Like Carthage and Rome, it is the war of the whale and the elephant. Yet it is actually not difficult for the same person to alternate between these two views, as in a gestalt switch: indeed, it is difficult not to, and this, I think, is significant.

Clearly, we need a theory which will accommodate both groups of facts, a 'both and' theory, not an 'either/or' one. Suppose we were to say: both kinds of causality are real — purposive agency has effects and so does structural causality, people make societies and societies make people. Roy Bhaskar presents this third model, which he attributes to Peter Berger and his associates (Berger aand Pullberg, 'Reification and the
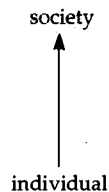
Figure 5.1 Model I: The Weberian Stereotype, 'Voluntarism'
(Source: PN, p. 32)

Figure 5.2 Model II: The Durkheimian Stereotype, 'Reification'
(Source: PN, p. 32)
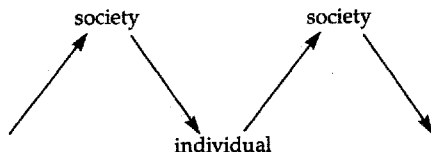
Figure 5.3 Model III: The 'Dialectical' Conception, 'Illicit Identification'
(Source: PN, p. 32)

Sociological Critique of Consciousness), along with the other two in Figures 5.1 to 5.3.

Now in one sense of the word 'make' it is quite true that societies make people and people make societies. However, they do not make each other out of nothing, or with nothing; and society does not make people in the same way as people make society. Society produces us as the people that we are, 'out of' a biologically given raw material, and it continues to transform us throughout our lives. We in turn make new societies out of old societies by our actions, whether
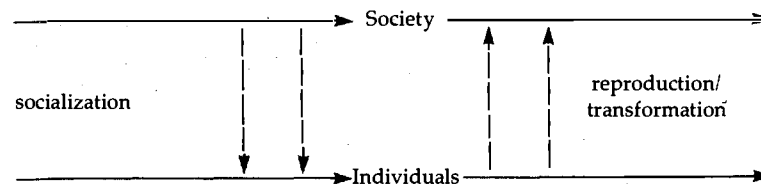
Figure 5.4 The Transformational Model of the Society/Person Connection (Source: PN, p. 36)

intentionally or not, and to whatever extent the new society either replicates the old one or is radically different. Societies (composed as they are of relations between people, and ramifications of those relations) can only exist as the outcome of human agency. If we were not reproducing/transforming social relations all the time, they would not exist: that is the truth of 'humanism'. But all human action presupposes the pre-existence of society and makes no sense without it. Its social context determines what actions are possible and what their outcomes will be. That is the truth of structuralism. But this account differs from Model III (Figure 5.3) in that the total social process is not a single linear sequence of causes (some social, some individual), but rather the interaction of two distinct kinds of entity, societies and people. So the study of social structures may be an autonomous 'structuralist' one, and the study of personal agency another autonomous 'humanist' one. Bhaskar represents the TMSA in Figure 5.4. He distinguishes this from the other three models on the grounds that 'on Model I there are actions, but no conditions; on Model II conditions, but no actions; on Model III no distinction between the two' (PN, p. 37)

But since societies and people are mutually ontologically dependent, separable only by analysis, there is more to be said. For an interactionist account of the relation between inseparables presents problems (with which we Spinozists have long been bothering our heads: see my 'The Materiality of Morals'). Bhaskar's solution is to say that there is a real (analytical) distinction, not only between human practice and social structure, but also between two aspects of each.

Society is both the ever-present *condition* (material cause) and the continually reproduced *outcome* of human agency. And praxis is both

work, that is, conscious *production*, and (normally unconscious) *reproduction* of the conditions of production, that is society. One could refer to the former as the *duality of structure*, and the latter as the *duality of praxis*. (PN, pp. 34–5)

Language is a good example. Unless we have learnt a pre-existent language with rules that exist independently of us we could not talk at all (structure as condition). We talk not as a rule to reproduce or transform the language but for personal ends of which we are conscious (practice as production). But our language only continues to exist because we talk, for it has no existence apart from people talking (structure as outcome). So our acts of talking do reproduce and transform the language, without our for the most part intending it to do so. For example, our children pick up the language to a large extent without its being taught (reproduction); and the language they learn is different from the one we learnt, since our usage differs (mostly without our noticing it) from the usage we learnt — e.g. the disappearance in current English of the distinctions between 'haven't got' and 'don't have', and between perfect tense and past historic, with the latter in each case supplanting the former (transformation).

There is, I think, a certain asymmetry between the two dualities as they relate to the two groups of sciences (social and psychological). Society as the condition of action and society as its outcome both belong to the subject-matter of social science, which is concerned with the mechanisms whereby the former develops into the latter. So far as people are concerned, the distinctively personal concepts — consciousness (and Unconscious in the Freudian sense), agency, reason, motive, desire, belief — are all connected with the aspect of action as production. Action as reproduction/transformation is generally action as taken over by social mechanisms. The only exception is when production and reproduction coincide, i.e. when people intentionally reproduce or transform their social structure. The duality of practice, then, is a duality between social and personal aspects of practice:

Now the autonomy of the social and the psychological is at one with our intuitions. Thus we do not suppose that the reason why the

garbage is collected is necessarily the garbage collector's reason for collecting it (though it depends on the latter). (PN, pp. 35–6)

This distinction between psychological and social knowledge is not arbitrary. We are concerned here with two distinct *strata*, though mutually ontologically dependent ones. This is a real ontological difference: people are not relations, societies are not conscious agents. Different strata, as we have seen, are characterized by different kinds of *mechanism*. Social science is centrally concerned with structural explanation because this corresponds to the kind of being that society has, as a network of relations. Purposive explanation, if it appears in it at all, does so only in exceptional circumstances, when a whole social organism is organized around some conscious purpose, like British society in World War II. On the other hand, intentional action is absolutely central to the study of people. Structural explanation here appears only at a second level, when one is explaining intentions. But I shall discuss intentions further in the following sections: here I am drawing attention to what Bhaskar calls 'an ontological hiatus between society and people' (PN, p. 37).

The importance of distinguishing categorically between people and societies, and correspondingly between human actions and changes in the social structure, should now be clear. For the properties possessed by social forms may be very different from those possessed by the individuals upon whose activity they depend. Thus one can allow, without paradox or strain, that purposiveness, intentionality and sometimes self-consciousness characterize human actions but not transformations in the social structure. The conception I am proposing is that people, in their conscious activity, for the most part unconsciously reproduce (and occasionally transform) the structures governing their substantive activities of production. Thus people do not marry to reproduce the nuclear family or work to sustain the capitalist economy. Yet it is nevertheless the unintended consequence (and inexorable result) of, as it is also a necessary condition for, their activity. Moreover, when social forms change, the explanation will not normally lie in the desires of agents to change them that way, though as a very important theoretical and political limit, it *may* do so. (PN, p. 35)

This has the consequence that differences that are crucial at the personal level may be of no account at the social one, and vice versa. Marrying contributes equally to reproducing the nuclear family whether it is done for love or for money or, like Luther, 'to spite the Pope and the Devil'. A miser in a pre-capitalist world and a financial magnate in a capitalist one may be driven by the same psychoanalytically 'anal character', but with quite different consequences.

At this point I would like to answer a possible objection to the transformational model of social activity. The idea that human activity is production, i.e. the transformation of raw material with tools, is familiar from Althusser's Marxism, and indeed implicit in the work of Marx himself. This model has been criticized as indicating an objectionable technologism. Thus Heidegger, criticizing Marx in his 'Letter on Humanism' (*Basic Writings*, p. 220), says:

> The essence of materialism does not consist in the assertion that everything is simply matter but rather in a metaphysical determination according to which every being appears as the material of labour. . . . The essence of materialism is concealed in the essence of technology, about which much has been written but little has been thought.

Elsewhere (in 'The Question Concerning Technology', *Basic Writings*, pp. 283–317), Heidegger expands on this notion of a technological world-view which ends by reducing everything, nature and human nature, to raw material for the economic production-process. And this world-view precedes industrialism by a couple of centuries. The world-view that Heidegger describes is, I think, real enough and dangerous enough. But the following should be noted.

(a) It is one thing to say that human action typically takes the form of production with and out of pre-existing entities, quite another to say that these entities are *nothing but* a stock of raw materials for our use. The former follows from the kind of being we are — we are not the sort of animal that can live without producing, nor are we gods that can create out of nothing. Heidegger's own excellent account of our work-world in Division One of *Being and Time* presupposes as much.

(b) Is it not significant that the world-view in question predates the technology that most clearly embodies it? Is not the determination to treat all being as nothing but a stockpile of raw materials a feature of production for the market, of capitalist relations of production rather than industrial forces of production?

(c) If there is a single philosophical idea which reflects more closely than any other this commercial (rather than technological) spirit, it is the epistemic fallacy, which reduces nature to our cognitive appropriation of it, just as this spirit reduces it to our economic appropriation of it. This epistemic fallacy has dominated philosophy for just the same period. In offering us the chance to break decisively with this fallacy and the consequent anthropocentric world-view (Russell's 'three centuries of subjectivistic madness'), Bhaskar's realism makes possible (though it does not actually entail) a much greater respect for the integrity of things independent of us.

If there is an ontological hiatus between society and people, Bhaskar nevertheless holds that:

> we need a system of mediating concepts, encompassing both aspects of the duality of praxis, designating the 'slots', as it were, in the social structure into which active subjects must slip in order to reproduce it; that is, a system of concepts designating the 'point of contact' between human agency and social structures. Such a point, linking action to structure, must *both* endure and be immediately occupied by individuals. It is clear that the mediating system we need is that of the *positions* (places, functions, rules, tasks, duties, rights, etc.) occupied (filled, assumed, enacted, etc.) by individuals, and of the *practices* (activities, etc.) in which, in virtue of their occupancy of these positions (and vice versa), they engage. I shall call this mediating system the position-practice system. Now such positions and practices, if they are to be individuated at all, can only be done so *relationally*. (PN, pp. 40–1)

This last sentence, while I believe it is true, is perhaps not obvious. The non-obviousness is due, I think, to a pervasive tendency to reify relations. While no one forgets that parenthood is constituted by a relation to a child, it is possible to forget that holding a job is a relation to an employer (and others whom the work serves) since much of the practice of the job

may not involve direct interaction with them. And it is very common (both in everyday consciousness and in political theory) to forget that ownership of property is not a binary relation of owner to property, but a relation to all those whom one's ownership excludes from access to the property. However, as Bhaskar goes on to point out, while all relations *include* interaction between the related, they do not all *consist in* such interaction. The relation between a citizen and the state, for example, comprises a whole range of rights and duties which are by no means always being exercised.

So this conception of 'my station and its duties'[2] gives support to the relational model of social being, of which Bhaskar goes on to sketch in a few details. From the standpoint of social science, the relations to be studied are those between 'positioned-practices' rather than those between individuals (though the latter there must be, and they may be studied by the psychologist and the historian). For it is the relations between positioned-practices that endure through changes in individual bearers. Bhaskar also gives a brief discussion of internal relations, i.e. relations such that the related being is what it essentially is by virtue of the relation. Many philosophers have denied the existence of internal relations, while others have claimed that all relations are internal. Bhaskar very reasonably takes the view that some relations are internal and others not, and points out that a relation may be internal on one side and external on the other.

Two more points must be noted before summing up and passing on to the account of human agency. (i) Bhaskar hints in PN at various ways in which this relational and transformational view of society will be important for politics and other potentially emancipatory practices. This part of his theory is developed fully in SRHE, and I shall devote my next chapter to it. (ii) Bhaskar claims that his relational and transformational view entails certain differences between social and natural structures. I shall criticize this view in the final chapter, but it behoves me to note the differences here, which I shall do by the following quotation:

1. Social structures, unlike natural structures, do not exist independently of the activities they govern.

2. Social structures, unlike natural structures, do not exist independently of the agents' conceptions of what they are doing in their activity.
3. Social structures, unlike natural structures, may be only relatively enduring (so that the tendencies they ground may not be universal in the sense of space-time invariant). (PN, p. 38)

In general, the TMSA (with its associated relational model) as presented by Roy Bhaskar looks a very well-founded theory. In the first place it is unembarrassed by the data which support either atomism or collectivism, humanism or structuralism, while each of those theories is embarrassed by the data which support its opposite. Indeed none of those theories would command credibility were it not for the incredibility of its opposite. Confronted with a credible alternative, they lose their attraction. And this alternative is a genuine dialectical synthesis in which the truth in each of the surpassed theories is preserved.

But the TMSA is no mere compromise. It is an original theory which can ground the autonomy, coexistence and conjoint application of the psychological and social sciences. Above all, it gives an account of how we interact with society, being both its effects and its causes, yet not mere links between social causes and social effects, but the beings by which a unique kind of causal power comes into the world: the causal power of reasons. It is to this that I now turn.

Agency: Reasons as Causes

Bhaskar's chapter in PN on agency (pp. 80–119) is concerned with the explanation of human action in terms of reasons, and the place of such explanation in a stratified account of explanation in the human world. He aims to show both that the human power of acting on reasons is irreducible, and that it is 'naturalistic' in the sense that this is a kind of causal explanation which takes its place among others, as an emergent power.

The concept of reasons for actions belongs in a 'mentalistic' language which includes concepts like those of beliefs, desires, intentions. If the psychological sciences are to be autonomous in

the sense of having their own subject-matter — a specific set of emergent powers — these mentalistic concepts will be central to those sciences.

Bhaskar introduces this mentalistic language, somewhat unfortunately in my view, by saying:

> The powers most naturally invoked here are those that involve consciousness, that is, those states of persons in virtue of which mentalistic predicates are applicable. (PN, p. 80)[3]

In fact, the mentalistic concepts which Bhaskar uses are not tied to the concept of consciousness, as he himself argues in this chapter and elsewhere. One may act on reasons, desires, etc. of which one is *unconscious*. So while the zone that Bhaskar stakes out for psychology is not reducible by redescription in physicalistic or behaviourist terms, it is not a Cartesian 'pychology of consciousness' either. In this respect as in many others, Bhaskar's philosophy is particularly welcoming to psychoanalytical approaches.

The central tenet of the non-reductive naturalism defended here is an affirmative answer to the question *'Can reasons be causes?'* (PN, p. 80). This requires defence on two fronts. For reductionist theories dispense with the notion of reasons for actions, and explain actions by other types of cause. And several twentieth-century philosophical positions (both in the Anglophone world and in continental Europe) treat reasons and causes as different sorts of explanation, which cannot overlap or cohabit in the same explanatory account.

What does it mean to say that reasons can be causes? Bhaskar suggests that

> When something is cited as a cause it is, I think, most typically being viewed as that factor which, in the circumstances that actually prevailed, 'so tipped the balance of events as to produce the known outcome'.[4] Clearly such a concept is non-Humean and generative. But any full transcendental realist defence of the naturalistic status of reason explanations will need to show not only that reason explanations function in our discourse in a causal kind of way, but that reasons are analogous to the causal structures of nature and that empirical knowledge of them is possible. (PN, p. 83)

Intentional actions involve beliefs and desires. In explaining such actions 'if a cognitive item, such as a belief, is mentioned, a conative one, such as a desire, is presupposed, and vice versa' (PN, p. 83). It is hardly open to dispute that, given a desire for something, coming to have a belief about the way to get it may 'tip the balance', and so be naturally described as 'the cause'. If the belief was already held, coming to have the desire may tip the balance. In any adequate explanation, both factors would have to be mentioned as causes. Why then should anyone want to deny that reasons for actions, which are generally specified in terms of beliefs and/or desires, can be causes? A reductionist might want to drop talk of reasons altogether and explain actions in terms of some physicalistic system, but we have already seen that this won't work (*re* buying the *Guardian*). But many philosophers seem to think that by denying that reasons can be causes, they somehow preserve the autonomy of rational action from causal reduction. Some of these philosophers might want to accept that it may be a reason that 'tips the balance', so that what is done would not have been done without that reason. In this case, the denial that the reason is a cause can only express a peculiar definition of cause — generally a Humean one. And we have already seen that a Humean notion of cause won't do. But if they are prepared to deny that, in this ordinary sense of cause, reasons can be causes, then they are 'saving' reasons by removing them from the world of real events. Hence on such views

> the very distinction on which the language-stratum theorist pitches his brief, between things that we do (a), like catching buses, and things that happen to us (b), like catching colds, becomes impossible to sustain. For it is only if we are the cause of some but not other of our bodily movements that such a contrast can be maintained; and that we can properly be said to act at all.
>
> For the transcendental realist there is no problem in sustaining such a contrast, and such a concept. For in the (a), but not the (b) case, the agent's reasons are a necessary condition for the bodily movements that occurred, in the straightforward sense that had the agent not possessed them (and unless the bodily movements were overdetermined) they would not have occurred. (PN, p. 89)

And as for the claim (very common among British analytical philosophers a few years ago) that reasons could not cause

actions since cause and effect must be logically distinct, 'Logic connects statements, not events, actions and the like, which are connected, when they are, by relations of natural necessity' (PN, p. 85).

That reasons can be causes is also a necessary condition of the phenomenon known as *rationalization* (in the Freudian sense). This occurs when the reasons sincerely given for an action by the agent are *not* the real reasons. Thus we may suppose that Henry VIII sincerely believed that he had his marriage to Catherine of Aragon annulled because it was contrary to canon law, whereas the *real* reason was that she had not provided him with a male heir, or perhaps that he fancied Ann Boleyn. What is the force of '*real* reason' here? Surely, causally efficacious reason.

Now someone might want to claim that rationalization is a fringe phenomenon — after all, the word only came into the language with Ernest Jones's paper 'Rationalization in Everyday Life' (1908). But Bhaskar wants to claim that the distinction between a real reason for a belief or action (one which is causally efficacious) and a possible reason (that is, I take it, something that has the logical standing of a reason for it, whether or not it is *anybody's* reason), is fundamental to our whole way of thinking about thought and action. For any self-critical thinking depends on recognizing the possibility that one is in error; doubt, conjecture and hypothesis about one's own and others' mental states must be possible.

> In this way the logical possibility of error about, misdescription and misrecognition of one's own state of awareness, and hence *inter alia* of one's reasons, is a condition of any reflexive intelligence. (PN, pp. 91–2)

In questioning one's mental states in this way, one is, among other things asking whether one's putative reasons are one's real reasons, i.e. the reasons that are effective. For instance, I may come to question whether my believing a scandalous story about an odious political leader is really caused by the evidence for the story or my desire to vilify that leader, or whether my depression is caused by the state of the world or the state of my digestion.

The case of practical reason is similar:

> unless a reason could function as a cause, there would be no sense in a person evaluating (or appraising) different beliefs in order to decide how to act. For either a reason will make a difference to his/her behaviour or it will not. In the former case it counts as a cause. In the latter case it is logically redundant, and deliberation, ratiocination (and indeed thought generally) become practically otiose. (PN, p. 92)

Now a reason may provide the 'balance-tipping' cause of an action when it is the new element — as when someone convinces you of the truth of some belief, which then precipitates an action. (Openness to effect of such reasons on one's behaviour is surely what we mean by freedom, when we contrast freedom with the compulsive action of a psychopath, or of a normal person in an abnormal state, such as sleepwalking or post-hypnotic suggestion.) But a reason may also have effects as a long-standing disposition.

> Thus the *possession* of a reason, conceived as a more or less long-standing disposition or orientation to act in a certain way, may itself be a cause — as being a social democrat gives an agent a reason for voting Labour. (PN, p. 93)

Such reasons 'have to be analysed normically, that is, as tendencies' (PN, p. 93) — tendencies which can exist unexercised or be exercised unrealized, like any tendencies. And like other tendencies too, they can themselves be explained in terms of deeper structures, and so on. Thus one's tendency to vote Labour may itself be normically explained: one is a trade unionist, and trade unionists tend to vote Labour. This tendency of trade unionists may in turn be explained in terms of theories about class and politics, and so on.

So reasons belong to the causal order, cohabit and interact with other causes in the open system of the world. They are explicable in terms of, but irreducible to, deeper strata of the social (and also ultimately the natural) world.

This conception of reasons entails a certain kind of philosophy of mind, and a non-Humean one. Reasons are beliefs, but

beliefs are not external to the ongoing life of desiring and acting; for desires, emotions and intentional actions all presuppose beliefs.

> Reasons, then, are beliefs rooted in the practical interests of life. And a person's essence consists just in what she is most fundamentally disposed to do (or become): that set of effective beliefs that determines her psychic (and behavioural) identity, and fixes her in her particularity as a kind. (PN, p. 96)

Synchronic Emergent Powers Materialism

What is the relation of mind to matter in this stratified conception of nature and our place in it? For Bhaskar, it is an instance of non-reductive materialism, which he calls by another unpronounceable tetragrammaton, 'SEPM' (synchronic emergent powers materialism).

SEPM is actually a very open-ended and 'permissive' theory. Its main point is that mental powers are emergent powers, not occurring in the absence of matter, but not reducible to material powers. He explicitly leaves it open whether these powers are (a) not the powers of any substance, but emergent from complex forms of matter; (b) the (non-material) powers of a material substance (perhaps the brain); or (c) the powers of an immaterial substance (PN, p. 98).

In what sense then, it might be asked, is SEPM materialist? Indeed, Bhaskar often uses the word 'materialism' pejoratively in philosophy of mind contexts, as if the term were not included in the title he gives to his own position. When he does so, I take it that 'materialism' is really short for 'central state materialism', i.e. neurophysiological reductionism. I think that SEPM is materialist only in the sense that, while it does not rule out mind as an immaterial substance, it would insist that any such substance ontologically presupposed material substances. But why, one may ask? Probably because the criteria for the existence of any imperceptible entity (which an immaterial entity must presumably be), must be *causal* criteria — the capacity to produce effects on matter. The criterion for the existence of a poltergeist is that it breaks the china. More seriously, the criterion for the existence of a belief in social

democracy is that the believer votes Labour. (In passing: we should beware of the epistemic fallacy here. An immaterial substance which did not have material effects could not be known by us; but we can't assume a priori that everything that exists can be known by us.)

Finally, SEPM is a theory of *synchronic* emergence. Emergence theories first emerged in connection with 'emergent evolution' — clearly a diachronic notion. According to such theories, new higher strata emerged from lower ones at certain times in natural history; and an account is sometimes then given *why* they should have emerged, for instance some sort of immanent teleology in the lower strata as in the Hegelian tradition, or a tendency of life forms towards greater complexification, as in Freud's Eros (see 'Beyond the Pleasure Principle'). By dubbing his emergence theory 'synchronic', Roy Bhaskar brackets off (a) questions about temporal priority. In principle (though the empirical data generally indicate otherwise) an emergent stratum could have existed from all eternity alongside the one in which it is rooted, just as for Thomas Aquinas it was in principle possible (though the data of revealed theology indicated otherwise) that the universe could have existed for all eternity alongside its Creator.

This is not a purely academic question, for while all the strata that we know about *do* seem to have emerged at some particular time, there are instances where it is arguable that two or more strata, one of which is rooted in and emergent from the other, *must* have emerged simultaneously, since they ontologically presuppose each other. I have suggested that society, mind and language are related in this way.

Bhaskar is also bracketing off (b) questions about the causes of emergence. There is nothing in the nature of synchronic emergence which answers the question whether a given stratum emerged by accident, design, or some sort of immanent teleology or teleonomy.

The defence of SEPM against central state materialism can be summarized briefly, since it is an instance of a general case against reductive programmes, which has already been discussed. The 'reduction' of B to A could be taken to mean only (i) that A provides a *basis* for B; this is acceptable: it is uncontentious that the brain provides a basis for (is a condition

of the possibility of) the mind. If 'reduction' is taken to mean (ii) that A *explains* B, we must distinguish between explaining how B came to be ('diachronic explanatory reduction' — e.g. speculations about the origin of life) and a synchronic explanatory reduction which would 'explain it away'. Diachronic reduction is compatible with synchronic emergence: even if we could show the origins of life in the chemical structure of the primal soup, that would not mean that life *is* nothing but primal soup; it is only synchronic reduction which SEPM needs to refute. This it can do along the lines of the need for a well-defined science-to-be-reduced, which I have already discussed. Just as we have to identify an action in social terms (for instance, as an act of voting Tory) before we can explain it psychologically (for instance, as an expression of an authoritarian personality), so we must identify a psychological phenomenon in psychological terms before we can explain it in terms of neurophysiology — and this prerequisite is not a ladder that we can climb and then throw away, since inter-stratum explanation involves constant access up and down the ladder. A third type of explanatory reduction, which tries to *predict* B on the basis of A, is ruled out as requiring closure where none is to be had.

Bhaskar sets out two criteria for successful synchronic explanatory reduction, and claims that neither of them is satisfied. The first is that if the individuals of the two kinds occupy the same place at the same time one must be part of the other. I doubt whether this criterion is very helpful. Some important examples meet this criterion yet reduction does not succeed: a union meeting and the members present are related in this way, yet one cannot be reduced to the other; likewise an organism and its cells. And there is not a lot of point in saying that reduction holds between a pint of beer and the four gills of which it is composed. Furthermore, this criterion does not help in the present instance: Bhaskar claims that it does because 'it makes no sense to locate an economy or a set of beliefs at some point in space' (PN, p. 99). But of course this means that these examples *pass* this test, since it implies that beliefs and brain-states (or people and economy) do *not* occupy the same space at all, and hence do not occupy it without one being a part of the other.

The second criterion is that the terms of two sciences must be partly intertranslatable or their reference states must overlap. But social states, psychological states and neurophysiological states possess properties such that the attribution of higher-order properties to lower-order entities makes literal nonsense. I may be angry about the plans to privatize British Rail, but the anger-rhythms mapped by my EEG cannot be *about* anything. In general, Bhaskar holds, social processes are teleonomic, psychological processes teleological and neurophysiological processes mechanical (PN, p. 100). This argument against physicalistic reduction is elaborated very effectively in the section of PN 'In Defence of Transcategorial Causality' (pp. 101–7). As this title suggests, Bhaskar is defending an interactionist account of the relations between mental and physical strata.

SEPM is significant in that it safeguards the irreducibility and effectivity of reasons in social life. But one should not confuse it with either of two other conceptions of the place of reason in social explanation. In the first place there are theories which treat 'rationality' (e.g. 'economic rationality') as itself an explanation of social activity. The conception of rationality involved is almost always instrumental rationality towards goals that themselves lie outside rational determination, and are generally taken to be the maximization of some sort of utility. Bhaskar's section on 'Rational Explanation' (PN, pp. 107–14) shows such theories to be (psychologically) either trivial or false, and (sociologically) irrelevant since social relations pre-exist and do not express rational agency. Human agents are located in and both empowered and constrained by social structures (plural), which often place inconsistent demands on them.

> It follows from the theoretical dislocation of society and persons and the hypothesis of the stratification of mind that in the field of the human sciences one is dealing, in opposition to sociological individualism and psychological empiricism (or rationalism), with a *double decentring* — of society from man, and of mind from consciousness. (PN, p. 112)

> Now it follows from this second decentring that though psychic unity may be a goal, and is certainly an accomplishment, it is not

(and cannot be) a presupposition of any science of psychology. (p. 113)

Bhaskar's own conception of rationality is a much deeper one, as we shall see in the following chapter.

Second, there are theories which quite simply exaggerate the powers of human reason, either in society as such or, more commonly, in some projected ideal society. Such theories neglect four limits to rationality which are to a greater or lesser extent present in all human action. These Bhaskar lists (SRHE, p. 126) as *unintended consequences* (sometimes of a systematic nature: 'alienation', 'counterfinality'); *unacknowledged conditions* (i.e. aspects of the social world which enable the action but are unknown to the agent); *unconscious motivation* (as in Freud); and *tacit skills* (e.g. the rules of syntax or rhetoric, which we use whenever we speak intelligibly or effectively, yet could not spell out). So that, while we may know under one description what we are doing — and must do so if it is really an action — there may be other true and relevant descriptions under which we (literally) do not know what we are doing. A public speaker may know that he is insulting a visiting potentate before a large audience. He may not know that he is sparking off a diplomatic incident (unintended consequence), using the newly installed amplification equipment (unacknowledged condition), re-enacting a childhood trauma (unconscious motivation), and using metonymy (tacit skill). 'Corresponding to each of these cognitive limits, human scientific knowledge promises a distinct emancipatory benefit' (PN, p. 126).

Social Knowledge

Bhaskar's aim in the passages discussed in this chapter so far has been to fill in the ontological background to social knowledge — to say what sort of thing society must be if knowledge of it is to be possible, and what *sort* of knowledge of it is possible. And we have arrived at some answers to these questions: the life of society is governed by laws which can interact and codetermine events with other laws; these laws operate at a multiplicity of emergent strata, rooted in but

irreducible to natural strata. Since social entities presuppose a natural environment and natural components, and since they exist only in symbiosis with social entities at other strata (societies with people, and so on), we can find only open systems here. So social science must search in the open systems of social life for the various emergent mechanisms that codetermine them. Since we are ourselves the social beings, our own social consciousness may be the starting point — albeit the *corrigible* starting point — of our inquiries (of which more later).

The social sciences, it may be said, are (a) explanatory sciences; (b) sciences without closure; (c) sciences with hermeneutic premises. This sets the scene (and some of the problems) for social epistemology. But it may also be said: this supposedly preliminary, hypothetical and purely philosophical inquiry has actually yielded some *results* for social science. The relational structure of societies and their irreducibility to individuals, the effectivity of reasons and their socially conditioned nature, the in-gear freedom of human agents and their non-transparency — aren't these all substantive, if rather general, social scientific results, about which different schools of social science contend?

I think in fact that they are, and that this is not surprising. For the inquiry was not into how knowledge in general is possible, but how knowledge of society is possible. The object, society, is among the premises of this argument. Some acquaintance with society is presupposed. It was not purely in jest that I referred to the Thomist conception of connatural knowledge (i.e. for example, the knowledge of virtue that a virtuous person has, not through having studied ethics, but through being virtuous). And if we can acquire some (very general) theoretical knowledge by spelling this out, might we not acquire some more specific knowledge about societies and people using similar transcendental (or at least retroductive) arguments from more specific premises in social practice (the labour contract, exchange of goods for money, usury), by asking 'how are these possible?'

First, though, to the bad news: social sciences are sciences without closure. They cannot do anything like shutting off the effects of processes which are not being tested, in order to isolate and test a single mechanism. They can't even secure

constancy of other processes. At most they may *think* constancy: *if* other factors were constant, the rate of profit would fall, or whatever. There may be efforts towards statistical neutralization of variables, such as are beloved of 'empirical psychologists', but statistics brings its own problems (and I would argue that they are much worse than usually imagined).

The impossibility of closure means that there can be nothing in the social sciences like the test between two theories seen in the elegant experiment with needle, cork and glass of water described in chapter 2. It also means that while we may postulate quantitative variations, we can't measure them. Any attempt in the social sciences to imitate the use of maths that is so central to the natural sciences is a blind alley.

How important is all this? Surprisingly, given the central place of the argument from the possibility of experiment in Bhaskar's philosophy, he does not think it need worry us too much. For the social sciences can

(A) inquire into open systems in the same way as the concrete or applied natural sciences do, in the manner dubbed RRRE (resolution, redescription, retrodiction, elimination);
(B) find a partial analogue to experiment; and
(C) find a compensator for its absence.

We have already (in the previous chapter) encountered the RRRE model, which Bhaskar introduces, using an example from historical explanation, in RTS. As he expresses it in PN (p. 129):

> Now explanation in open systems is in general accomplished by a four-phase process:
>
> 1. *Resolution* of a complex event into its components (causal analysis).
> 2. *Redescription* of component causes.
> 3. *Retrodiction* to possible (antecedent) causes of components via independently validated normic statements.
> 4. *Elimination* of alternative possible causes of components.

I mentioned earlier that RRRE normally presupposes a stock of concepts tested under experimentally controlled conditions: the 'pure' science which RRRE *applies* ('*independently validated*

normic statements'). In SRHE, Bhaskar contrasts the procedures of theoretical and practical (applied) sciences, in that the latter is RRRE while the former is 'DREI', that is:

> *description* of law-like behaviour; *retroduction*, exploiting analogies with already known phenomena, to possible explanations of the behaviour; *elaboration* and elimination of alternative explanations; issuing (ideally) in the empirically-controlled *identification* of the causal mechanism(s) at work. (SRHE, p. 68)

Elaboration and elimination (a single stage) appear in both models. DREI starts with description whereas RRRE has *re*description as its second stage, indicating the presence of an already established stock of concepts, well enough defined (presumably by pure, theoretical science) to justify using them for revisionary description.

Retroduction (not to be confused with retrodiction) we have already encountered (in chapter 1) as the genus of which transcendental argument is a species. Bhaskar has characterized it in these terms:

> Typically, then, the construction of an explanation for, that is, the production of the knowledge of the mechanism of production of, some identified phenomenon will involve the building of a model, utilizing such cognitive materials and operating under the control of something like a logic of analogy and metaphor, of a mechanism, which *if* it were to exist and act in the postulated way would account for the phenomenon in question (a movement of thought which may be styled 'retroduction'). (PN, p. 12)[5]

And identification appears only in connection with theoretical science, since it is tied to experimental closure (at least 'ideally').

Now if we apply these models to the social sciences where closure is inconceivable, what do we find? In the first place, the second and third stages of RRRE seem to contain sockets into which leads from pure science need to be plugged. In the natural sciences, we have got such leads: leads from physics and chemistry can be plugged in to engineering and meteorology, leads from biology and chemistry into medicine, and so on. In the social sciences, apart from the ontological premises that we have derived from our transcendental arguments, we seem to

$$
\begin{array}{c}
R \\[4pt]
\left\{ \begin{array}{c} R \\[6pt] R \end{array} \right. \\[4pt]
\end{array}
$$

DRE $-\,-\,\to$

E $-\,-\,\to$ I?
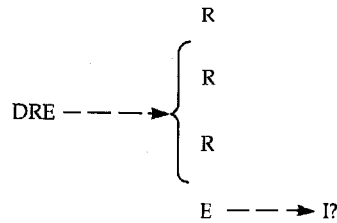
Figure 5.5 Procedure in Social Sciences

lack such leads. Do we have to substitute *speculative* redescriptions and retrodictions for scientific ones, guided and constrained only by these ontological premisses? Such was my suggestion in *Scientific Realism and Socialist Thought*, where I coined the term 'epistemoids' for such practices, but Bhaskar rejects this epistemic pessimism. The alternative is that something may be left of DREI in the social sciences, even without experiment. And in fact the first three stages of it do seem to stand without necessity for experiment. At the fourth stage, leading to I, however, we have got a loose plug. Figure 5.5 illustrates what we get if, in accordance with the maxim that in the social sciences the applications are the only 'experiments', we plug DRE in to RRRE's two empty sockets and the conclusion of RRRE into I. We remain without 'crucial experiments', and without accurate measurements, but not without everything:

> once a hypothesis about a generative structure has been produced in social science it can be tested quite empirically, although not necessarily quantitatively, and albeit exclusively in terms of its explanatory power. (PN, p. 49)

Now let us come to Bhaskar's 'analogue' and' compensator' for the absent experiments (PN, p. 47). He says rather little about the analogue, though I believe the idea could be developed to great profit.

> It might be conjectured that in periods of transition or crisis generative structures, previously opaque, become more visible to

agents. And that this, though it never yields quite the epistemic possibilities of a closure (even when agents are self-consciously seeking to transform the social conditions of their existence), does provide a partial analogue to the role played by experimentation in natural science. (PN, p. 48)

Here Bhaskar refers in a footnote to a passage of mine (in *R. D. Laing*, p. 132) where I introduce a phrase that Bhaskar has also adopted elsewhere, 'the methodological primacy of the pathological'. By seeing how something goes wrong we find out more about the conditions of its working properly than we ever would by observing it working properly. Neurotic human beings, as Freud says, are more instructive psychologically than normal ones. An economy in crisis is more 'transparent' than a smoothly functioning one — it 'reveals codes', shows its works like the pipes in the Pompidou Centre. Mechanisms which are normally disguised by their close interaction with other ones break loose and so are actualized, whereas they normally operate unactualized — just as the law of gravity operates unactualized in your house until one day the roof falls down on your head.

Finally we come to the 'compensator' for the lack of experiments, to which this section has been leading up. In the first place, it is clear that we need some account of theory construction in the social sciences, since otherwise we are confronted by a mere mass of data. But we have in fact always got 'proto-scientific' or ideological theories about society, since such theories are an essential part of social practice. Hermeneutic accounts of social science often take those proto-theories as themselves authoritative for the testing of a social-scientific theory: economists must use the same concepts as businesspeople, and so on. But it is essential to transcendental realism that theory can be counter-phenomenal. The question is therefore how to *transform* proto-theories into scientific theories which can explain and possibly contradict their own theoretical raw material. Bhaskar says (using P for proto-theory, T for a social scientific theory):

> The first step in the transformation P→T will thus be an attempt at a real definition of a form of social life that has already been identified

under a particular description. Note that in the absence of such a definition, and failing a closure, any hypothesis of a causal mechanism is bound to be more or less arbitrary. Thus in social science attempts at real definitions will in general precede rather than follow successful causal hypotheses — though in both cases they can only be justified empirically, viz. by the revealed explanatory power of the hypotheses that can be deduced from them. (PN, pp. 49–50)

Thus the process of theoretical transformation includes empirical testing, and for that reason among others nothing in its premisses is unrevisable. But more theoretical work needs to be done to set up a test by explanatory power in an open system than to set up a test by predictive power in a closed one. The power of abstraction, as Marx said, must replace chemical reagents.

The sort of theoretical work that Bhaskar has in mind here consists of transcendental arguments:

Now the substantive employment of an essentially apodeictic [demonstrative — A.C.] argument should occasion us no surprise. For transcendental arguments are merely a species of which retroductive ones are the genus, distinguished by the features that their *explanandum* consists in the conceptualized activities of agents and, as becomes an arena characterized by a multiplicity of causes, that they isolate necessary not sufficient conditions for it. (PN, p. 50)

These transcendental arguments belong to social science, not to philosophy, for the supplementary considerations which 'will be needed to establish the validity of the analysis' will

include the provision of independent empirical grounds for the existence (and postulated mode of activity) of the structural mechanisms concerned, whereas, in philosophy, in the nature of the case, this is impossible. (PN, pp. 50–1)

I take it that this is Roy Bhaskar's central positive contribution (alongside the important negative one, the exclusion of the quest for closure) to the methodology of the social sciences: the idea that a great part of their theoretical work will consist in

transcendental arguments from premisses familiar from social practice. He cites Marx's *Capital* as an instance of this method:

Capital may most plausibly be viewed as an attempt to establish what must be the case for the experiences grasped by the phenomenal forms of capitalist life to be possible. (PN, p. 51).

Bhaskar's account of social scientific work can be highlighted by contrast with the two main rivals, positivism and hermeneutics, to the discussion of which he devotes the final chapter of PN. Like hermeneutic theorists but unlike positivists, he holds that the study of any social practice must start with the agents' conceptions of it. But unlike the hermeneuticist and like the positivist, he holds that social science can go on to refute these conceptions. He holds social explanation to be both causal (as does the positivist) and interpretive (as does the hermeneuticist), denying their shared premiss that these two notions will not cohabit. And he rejects their shared acceptance of a Humean account of causality.

In the next chapter I will consider the account of the role of social science in human emancipation which Bhaskar bases on this account of social science. I shall then look at the way his theories have been used in particular human sciences. And in the final chapter I shall raise some critical questions about critical naturalism.

Notes

1. See, for instance, Thomas Aquinas, *Summa Theologiae*, Secunda Secundae, Quaestio XLV, Articulus II.
2. 'My Station and Its Duties' is the title of the crucial chapter in F.H. Bradley's *Ethical Studies*.
3. A better definition of mentalistic predicates might be as characterized by *intentionality* in the sense of essential reference to something ontologically independent of them. Bhaskar uses the term intentionality in the ordinary sense of purposiveness rather than this technical sense, though he calls mental states 'referential' which is essentially the same.

4. The phrase quoted by Bhaskar is from Scriven, 'Causes, Connections and Conditions in History', p. 245.

5. Bhaskar refers here to Durkheim, *The Rules of Sociological Method*, p. 2, and to Berger and Pullberg, 'Reification and the Social Critique of Consciousness'.

======6======

Explanation and Emancipation

Schopenhauer . . . would have sickened, become a *pessimist* (which he was not, much as he would have liked to be) had he been deprived of his enemies: of Hegel, of woman, of sensuality, of the human will to survival.
>   (Nietzsche, *The Genealogy of Morals*, p. 241)

Whether or not Schopenhauer was a pessimist personally, Schopenhauer's philosophy, which entails that the only emancipation from misery is extinction, and that the moon is to be preferred to the earth since there is no life on it, is, by *almost* general consent, the epitome of pessimism. But Bryan Magee, early in his book on Schopenhauer, says:

> Even professional philosophers tend to see him in this light, as is evidenced by the title of Frederick Copleston's book *Arthur Schopenhauer: Philosopher of Pessimism*. Yet this is odd, because it is an elementary point in logic that no truth claim can entail a value-judgement. If a valid argument has a value-judgement anywhere in its conclusions this can only mean that the same value-judgement was already to be found somewhere in the premisses: you cannot derive an 'is bad' from an 'is'. No general philosophy — no ontology, epistemology or logic — can entail pessimistic conclusions. Professional philosophers ought always to have known, without having to read Schopenhauer to discover it, that in this sense his pessimism is logically independent of his philosophy; and so it is. (*The Philosophy of Schopenhauer*, p. 13).

I thought it worth quoting this at length because it shows how an intelligent and fairly representative modern philosopher could be so convinced of this dogma that you can't argue from a fact to a value as to be led to say ridiculous things by it.

For this reason, I want to start this chapter by plunging (almost) straight in to an argument of Roy Bhaskar's which seems to me to make a clear and irreparable breach in the Hadrian's wall which modern philosophy has built to keep those nasty Pictish facts from marauding within the boundaries of the empire of value. Once this breach is made, the invasion can be extended in all sorts of directions.

But one preliminary point needs to be made. The arguments from facts to values are more like evidential or scientific than deductive arguments — unsurprisingly, for values exist in open systems, and value-judgements are normic, always (or almost always) holding *other things being equal*. Introducing a previously unrecognized premiss may vitiate a validly derived conclusion from a true premiss. In deductive reasoning, if P implies Q, then P and R implies Q (e.g. if 'the sheep's in the meadow' implies 'the sheep isn't in the fold', then 'the sheep's in the meadow, the cow's in the corn' implies 'the sheep isn't in the fold'). But in evidential reasoning, 'he was seen running from the scene of the murder with a smoking gun' may imply 'he probably did the murder'; but add 'his gun could not have fired the bullet that killed the victim', and the conclusion no longer follows. It has often been pointed out that moral reasoning is more like the latter: 'taking money from that Coca-Cola machine would be theft' may imply 'you shouldn't take money from that Coca-Cola machine' — but combine it with 'it is the only way to get coins to phone the President and stop a nuclear war', and the case is altered (I allude to the film *Dr Strangelove*). Bhaskar takes this into account by including a *ceteris paribus* ('other things being equal') clause in the conclusion of all his fact-to-value inferences.

Explanatory critiques in social science

In this section, then, I want to concentrate on a single argument that establishes the credentials of explanatory critiques as breaching the fact/value divide.[1] In Bhaskar's texts, this argument is embedded in a general account of fact—value relations (in PN), or of the ways in which theory can affect practice (in SRHE). But I focus on the central argument, on the

principle that if we can first storm the castle, these lower terraces can easily be taken. To this end I state the case in my own words, and conclude by quoting the passage where Roy Bhaskar sums it up most lucidly.

Social science, like any science, presents ideas, claimed to be *true* of the object studied, i.e. of society. Unlike the objects of natural sciences, the object it studies, society (or any concrete society), includes ideas. For society can only exist insofar as human agents act, reproducing and transforming the social structure. And human agents act in accordance with ideas. Even though ideas may be causally secondary to economics (at least in the dimension of 'vertical explanation'), and history may be the history of class struggles, as Marx claimed, there can nevertheless be no understanding of the English Civil War and Commonwealth without understanding Puritanism, or of modern Iran without understanding Shi'ite Islam, or of American foreign policy without understanding B movie westerns. So an account of the ideas prevailing in a society will be an essential part of a social-scientific account of that society.

Now many of the most significant ideas in any society will be ideas *about* features of that society. For instance, in Britain in the 1980s, a large number of people believed that unemployment was the result of the fecklessness of the unemployed. Any account of social attitudes, political behaviour, etc. in that period would need to mention that fact. But it would also need to mention the *real* causes of unemployment in the structure of British financial institutions, the world market, government policy, etc. Hence the explanations that were part of the social-scientific study, and the explanations that were part of the society studied, would contradict. If the social science had got it right, then the people it described who had the opposite explanation must have got it wrong. Hence the social science criticizes (part of) its object. There can be no equivalent of this in the natural sciences. Black holes may be unpleasant things to contemplate, but that is no criticism of them. They exist — or don't — and there's an end of it.

Further, the social scientist will not be content with noting the existence of a false belief in the fecklessness of the unemployed; he or she will want to explain it. And whether the explanation is something subtle and socially pervasive, like the atomistic

nature of social relations in a commercial society, or something crude and contingent, like lying press-lords, the criticism of the belief will rub off on to its cause. To say that some institution causes false beliefs is to criticize it. Given that (other things being equal) it is better to believe what is true than what is false, it is also better (other things being equal) that institutions that cause false beliefs should be replaced by, or transformed into, those that cause true ones.

Further still, particular institutions and false beliefs about them may be in a *functional* relation, such that the false beliefs serve to preserve the institutions that they are about. Where institutions oppress a substantial number of people, they will only be stable if protected by such false beliefs. In such cases, to propound the truth is not just to criticize, but to undermine the institution.

Hence, the production of explanations of social institutions is not only, as a general rule, a precondition of criticizing and changing them; sometimes, it *is* criticizing them, and beginning the work of their subversion. One classic example of this kind of explanatory critique to which Bhaskar refers is provided by Marx's account of the wage form. Wage-labour only occurs where the workers do not possess the means of labour (tools, workplace, raw materials), and therefore have to sell their power to work to someone who does. This initial separation of means of labour from worker is not given by nature, but the result of history. It perpetuates itself, since the product of the worker's labour belongs to the owner of the means of labour, and only a portion of it is paid to the worker — in general, too small a portion for the worker to be able to acquire the means of labour.

However, because the worker's pay takes the form of the price of the labour-power he or she has sold, it appears as if 'exchange is no robbery', and, while pay levels may be the subject of negotiation, some wage level or other would be 'fair'. Wage-labour spontaneously generates this ideology of 'wages as payment for labour', which, however, is *false*, in that (a) what is actually paid for is labour-power, (b) labour-power can only be a commodity when *labour* is not possible for the worker without such an exchange, since he or she is deprived of the means of labour, and (c) only a portion of the product of labour goes to the

worker — and the surplus accruing to the owner ensures that the worker's deprivation of the means of labour is perpetuated.

In this case, not only does the institution of wage-labour *cause* false beliefs about itself, it also *protects itself* from the wrath of the workers by this illusion. To expose it is to criticize the wage system (i.e. capitalism), and to spread this word is to stir up dissent from capitalism, which of course is just what Marx intended.

Another sort of case — slightly less clear-cut as an instance of fact-to-value argument, but very important for social science and its political implications — is that in which the causally efficacious institutions or distinctions in society are not the emotively charged ones. Lévi-Strauss reports that Bororo villages are arranged in circles and divided between two moieties, the Cera, who live in the northern half, and the Tugaré, who live in the southern. The men of each moiety must marry into the other one, funerals must be conducted by someone from the other moiety than that of the deceased, and elaborate mythological and ritual distinctions are associated with this division. Cutting across it is the division between the 'upstream' (eastern) and 'downstream' (western) halves of the village. And within each moiety, there are different clans, each with their traditional functions. All these distinctions are highly charged, and regulate the cultural and sacred life of the village. They give the villagers a self-understanding based on symmetry, complementarity, fraternity. Yet cutting across all these three charged distinctions, there is the division into three unequal endogamous groups, upper, middle and lower.

> Three societies which, without realizing it, will remain for ever separated and isolated, each imprisoned in a kind of pride which is concealed even from itself by a smokescreen of institutions, so that each is the unconscious victim of devices, the purpose of which it can no longer discover. (*Tristes Tropiques*, pp. 319–20)

And as the Bororo are, so are we Europeans, with our 'Europe of fatherlands' and our national prejudices, trailing a bloody history, and obscuring even more effectively than the wage form the class lines along which our interests really divide.

The dissonance between causal power and emotive chargedness of institutions does not of course involve a formal contradiction between two beliefs; so it is possible for someone, without formal inconsistency, to recognize, for instance, that Britain is only nominally a monarchy, yet to get excited about royalty. However, this is a phenomenon very close to *displacement* in the psychoanalytic sense, which I shall discuss in two sections' time as susceptible to explanatory critique.

The hardened fact/value dichotomist might respond: the argument jumps from fact to value when it introduces the assumption that it is best to believe what is true. However the questions 'what should I believe about x' and 'what is true about x' are not logically independent questions. In fact they are equivalent, in the sense that the answer to one is necessarily the answer to the other. It simply doesn't make sense to say 'that is true, but I shouldn't believe it' or 'I should believe that, though it is not true'.

This may seem to prove too much. For it looks as if it implies that true belief is always better than false belief, and it was only intended to prove this *other things being equal*. It is better that a would-be murderer should have *false* beliefs about his victim's whereabouts.

But the absolute character of the inference from 'it is true' to 'I should believe it' applies only in the first person case. I cannot separate the question of something's truth from the question whether I should believe it, but someone else, who has reason to believe that I might misuse the knowledge to do evil, or even just be deeply hurt by it, may judge that it would be better if I had false beliefs on a subject. (The relation between the tight argument from 'it is true' to 'I ought to believe it' and the looser argument from 'it is true' to 'he or she ought to believe it other things being equal' looks tricky. Deductions do not change validity according to who makes them. But the point is that since to believe something is to hold it true, 'I ought to believe it' can have no other grounds than 'it is true' has; 'he or she ought to believe it' can. I should note that this form of the argument is mine rather than Bhaskar's.)

As I have given this account of explanatory critiques in my own words, I shall now conclude it with a longish quote from Roy Bhaskar which sums it up lucidly:

If, then, one is in possession of a theory which explains why false consciousness is necessary, then one can pass immediately, without the addition of any extraneous value judgements, to a negative evaluation of the object (generative structure, system of social relations or whatever) that makes that consciousness necessary (and, *ceteris paribus*, to a positive evaluation of action rationally directed at the removal of the sources of false consciousness). Might it not be objected, however, that the fact/value distinction only breaks down in this way because one is committed to the prior valuation that truth is a good, so that one is not deriving a value judgement from entirely factual (natural) premises? But that truth *is* a good (*ceteris paribus*) is not only a condition of moral discourse, it is a condition of any discourse at all. Commitment to truth and consistency apply to factual as much as to value discourse; and so cannot be seized upon as a concealed (value) premise to rescue the autonomy of values from factual discourse, without destroying the distinction between the two, the distinction that it is the point of the objection to uphold. (PN, p. 63)

I have lifted this argument about explanatory critiques out of its context for the sake of clarity. Now it has to be said that the section in which this passage occurs is supposed to defend arguments *both* from facts to values and *from values to facts*. Indeed, he starts by saying that it is 'now often conceded that the facts are in some sense tainted by, or contingent upon, our values' (p. 55). He intends first to support this view, then to show, more contentiously, that some fact-to-value arguments can *also* be valid.

But this raises a doubt as to whether he may not be cutting off the branch he is sitting on. For if facts are already valuey, it is no great matter that they entail values. If we can argue from values to facts and then back to values again, the conclusions of the whole argument will be of the same evaluative nature as the premises, which will not surprise anyone. In this case it will be quite plausible to argue that the intervening, supposedly factual stages are a bit valuey. Either the fact/value gap has not been bridged, since the whole argument is valuey, or it has not just been bridged, but the distinction abolished altogether, which is not what Bhaskar is claiming. Let us consider his argument.

He discusses value-to-fact arguments under the following heads:

(a) from the standpoint of the subject of investigation
   (i) concerning the selection of problems,
   (ii) concerning the conclusions,
   (iii) concerning the standards of inquiry;
(b) from the standpoint of the object of investigation;
(c) from the standpoint of the relation between subject and object.

I shall argue that his argument is inconclusive since (as we shall see) under (a) (i) and (ii) he argues *against* those kinds of value-to-fact argument; under (a) (iii) he subsumes this value-to-fact argument under *relativism*, and then defends one, very restricted form of relativism, while refuting the more general kind; however, the kind of relativism he defends is not the kind that licenses value-to-fact arguments. Under (b) he does defend value-to-fact arguments — but in a way that only works on the assumption that there are *no* valid fact-to-value arguments. There is no separate discussion of (c); instead, he goes on to defend fact-to-value arguments, in the manner summed up in the last quote above.

My reason for criticizing Roy Bhaskar's argument now, rather than sticking to paraphrase and exposition and leaving criticism till afterwards, as elsewhere in the book, is that this criticism defends the radical and far-reaching nature of his fact-to-value argument, against concessions that would tend to weaken it.

(a) (i) It is sometimes argued that, in the social sciences, the complexity of the subject-matter forces us to be selective, and the selection is value-determined; Bhaskar argues that such complexity is equally to be found in natural sciences, and is selected from on *practical* criteria only in the *applied* sciences, whether natural or social. In pure sciences, principles of selection are not imposed but discovered. 'Thus while it is practical interests which determine which out of the infinite number of possible compounds of carbon are studied, it is theoretical interests which motivate the identification of its electronic structure' (PN, p. 56). The evaluative selection argument confuses the natural/social distinction with the pure/applied distinction. So *this* case for value to fact arguments doesn't work.

(a) (ii) A stronger case is argued on the basis of *'interference* between the subject's interest in the object and its knowledge of it'* (PN, p. 56). If human interests were bound up with geometric theories, said Hobbes, we would fight wars about them. In the case of social-scientific theories, they are and we do. But if we are *conscious* of such interference, we can correct it; if we are not, it is no use stating our evaluative premises, since we will be misled about them. Hence explicit evaluative premises for social science are either unnecessary or misleading. 'Interference' remains as a problem to be overcome, but not as a source of acceptable premises.

(a) (iii) This view 'posits a relativity in the methodological norms secreted by different conceptual schemes or paradigms, together with a value-dependence of such conceptual schemes of the sort already discussed under (ii)'. Bhaskar does not say why he treats this view more favourably than (ii). On the surface, it would seem that in this case, too, we could correct the interference of interests if we were conscious of it, while unconscious interference would be a problem to be overcome. But his strategy is to describe it as a special case of relativism, and to criticize anti-relativist arguments for confusing *'epistemic relativity*, which asserts that all beliefs are socially produced', with *'judgemental relativism*, which asserts that all beliefs (statements) are equally valid, in the sense that there can be no (rational) grounds for preferring one to another' (PN, p. 57). (It may be useful for the present argument to substitute 'cognitive' for 'rational' in the last sentence, since the value-to-fact relativist typically claims that there are moral or political, but not cognitive, reasons for preferring one theory.)

It seems to me that Bhaskar has misread the polemical situation here. Those who are called or call themselves relativists generally hold that epistemic relativity does imply judgemental relativism. Once these are distinguished, as Bhaskar does, the characteristic position of relativists is undermined. The epistemic relativity which Bhaskar accepts is widely held by anti-relativists. And this epistemic relativity is of no help at all to those who want to argue from values to facts.

I conclude that none of the arguments from the nature of the *subject* to value-to-fact inference work.

(b) The issue here is whether some features of the object

studied in social science require that it use evaluative language. That values are among the objects studied does not by itself require their description to be couched in valuey language, as Bhaskar rightly notes; a student of canine behaviour does not have to bark. But certain features of the object may require such language, Bhaskar claims. Now of course if the argument which I have already set out (though in PN it comes after) succeeds in showing that we can argue from facts to values, some social-scientific language will indeed be value-laden. But it will be so not in advance of or in addition to but just *by virtue of* being descriptive and explanatory. In this case, there is no question of *bringing* values to the discourse, and hence no real value-to-fact inference.

It seems to me that both the example and the general argument which Bhaskar gives to show the need for evaluative language are really cases of fact-to-value, not value-to-fact argument. Thus he cites Isaiah Berlin's example, that of the following four true statements about what happened in Nazi Germany: 'the country was depopulated', 'millions of people died', 'millions of people were killed', 'millions of people were massacred' — the fourth is both the most evaluative and the most precise and accurate; it gives more of the truth than the others. That is so, but the evaluative force arises entirely out of the factual content. It is not that by bringing values into the discourse one makes it a fuller statement of the truth, but that by making a fuller statement of the truth one implies more values.

At the theoretical level, the argument is that there is an irreducible, but corrigible, hermeneutic moment in social science; that one cannot get started without understanding the *meaning* that actions had for their agents, that institutions have for their participants, etc. But these meanings may be systematic delusions. To understand the Bolsheviks' actions at the time of 'War Communism', one has to understand that they thought they were initiating a rapid transition to a fully communist society; but one must also understand that, in fact, they were irreparably destroying the worker–peasant alliance on which the prospect of socialism in Russia depended, and transforming themselves into a self-perpetuating elite. By incorporating both understandings into one's account, one inevitably *criticizes* their self-understanding and consequent actions — and hence

becomes evaluative. What is this but an explanatory critique? The problem here is that Bhaskar has not at this stage introduced the notion of an explanatory critique as a way of arguing from facts to values; hence he is producing good arguments against people who insist that social sciences must have no evaluative *conclusions*; and treating these arguments as if they showed that social sciences may have evaluative *premisses*.

Bhaskar does not discuss (c) separately, but goes on to discuss fact-to-value arguments instead. If my assessment of his arguments is correct, he has found no real place for value-to-fact arguments — and so has not undermined the far-reaching consequences of his notion of explanatory critiques for ethics and politics. Why was he so keen to find defensible value-to-fact arguments?

He wants to take his distance from two mistaken views, each of which he sometimes calls 'scientism'. The first is the idea that a theory could, so to speak, create values where none had been before. Theories can have practical consequences, but only because we are all already valuing various things, as an inevitable part of living. His argument about the value of truth does not deny that truth is a value for us, but claims that it is a value that is presupposed by all our doings as cognitive beings. Non-cognitive explanatory critiques — to which I shall come shortly — likewise depend on our having values — needs, wants, desires, emotions — which may indeed be radically *transformed* by the work of theory, but can in no way be created by it *ex nihilo*. It is doubtful whether anyone ever thought it could; the Fabian example he quotes[3] is a telling instance of the Webbs' elitist arrogance, but does not fit the description since an evaluative input is assumed (the masses can describe their grievances, though not prescribe their remedies). But at least Bhaskar is forestalling a possible misreading of his own work by criticizing this view.

The second mistaken view from which Bhaskar is taking his distance is that which denies the legitimacy of sociological studies of science, and the political struggles over science that may arise from them. While we cannot understand science without understanding that it is an attempt to deepen our knowledge of its intransitive object, the scientific community is

also a social group subject to similar constraints and pressures to other such groups. This may affect its findings. At worst, there are cases of deliberate falsification, as in some studies of supposed racial determinants of intelligence, or in Soviet biology in the Lysenko period. Even when this is absent, it is possible to find what you want to find, and easier still to miss what you don't want to find. And even assuming all the results of a research project are objectively true, the area chosen for investigation may be determined by contentious ideological assumptions or practical interests. Thus it is likely that drug companies have concentrated on artificially synthesized drugs to the detriment of research into those occurring naturally in plants; and it is certain that military might and commercial profit are the chief determinants of which secrets of nature get uncovered. In a world where science was funded with a view to satisfying human needs and conserving planetary resources, quite different discoveries might be made — neither more nor less objective than the findings of modern science, but useful for different purposes. (I am certainly not belittling intellectual curiosity as a legitimate motive for science — but its economic efficacy is minimal.) Hence social studies of science may be of value in alerting us to likely sources of error; in well-established experimental sciences, this is a marginal role, but in the human sciences it is very significant. And such studies may inform political struggles over allocation of resources, and over the applications of science.

But these points do not mean that we can argue from values to facts. Research is motivated, but it is not the motivating values that determine its factual findings (or if it is, they are placed under suspicion of being 'false facts'); here, Bhaskar's formalization of the issues is less than helpful; he sets out to defend both 'F→V' and 'V→F' arguments, but the arrows do not mean the same in the two cases. As he says himself (using 'F' for facts, 'V' for values, 'T' for theory, 'P' for practice):

> the asymmetry between the F→V and T→P relationships, on the one hand, and the V→F and P→T relationships, on the other, stems from the consideration that whereas factual and theoretical considerations not only predispose and motivate, but, in favourable epistemic circumstances . . . and subject to the operation of various

> ceteris paribus clauses, logically entail value and practical judgements; value and practical commitments, while they may (and in general will) predispose and sometimes motivate, do not (non-trivially) entail factual and theoretical judgements. (SRHE, p. 173)

But this is to accept what he had rejected in PN (pp. 54–5), that 'no factual proposition can be derived from any value judgement', and 'any factual conclusion depends upon premises containing at least . . . one factual proposition'.

My exposition of Bhaskar's account of cognitive explanatory critiques in social science, and my defence of their status as unilateral fact-to-value arguments, is now complete. In the following section I discuss some approximations to and extensions of explanatory critique, with wider implications for social science, politics and ethics.

Other kinds of explanatory critique

Now that the fortress of non-naturalism (the doctrine that facts can't imply values) has been taken, it is possible to extend the notion of explanatory critique, and thus to begin to develop a naturalistic theory of practical reason in general. In this section, I shall discuss three ways in which this can be done.

(A) There are other and worse ills than cognitive error and inconsistency; social sciences can also uncover them. Roy Bhaskar writes of extending the pattern of argument 'to accommodate more interestingly specific forms of false consciousness, and indeed more generally of *defective or unfulfilling being*' (SRHE, p. 178, my italics). Social sciences may generate values and motivate practices by exposing these phenomena too.

(B) Explanatory critiques based on knowledge (not necessarily scientific knowledge) of human emotions have sometimes been presented as the basis for a practice of personal emancipation — without the concept of an explanatory critique being explicitly formulated — notably in the ethics of Spinoza and in Freudian psychoanalysis. An explicit theory of explanatory critiques such as Bhaskar's can throw

light on these projects, and perhaps form the groundwork of a naturalistic moral philosophy. As Bhaskar puts it: 'A transcendental realist ontology requires, it will be seen, as much readjustment in ethics as in epistemology' (SRHE, p. 187).

(C) Having loosed the stranglehold of non-naturalism, it may be possible to construct a general theory of practical reason in all its varieties, showing the differences as well as the similarities of familiar kinds of practical reasoning to that involved in explanatory critiques. Bhaskar lists seven levels of rationality and discusses them (SRHE, pp. 181ff).

In these ways, it can be shown that Tolstoy's remark quoted at the head of the chapter on facts and values (SRHE, p. 169) is mistaken:

> Science is meaningless because it gives no answer to our question, the only question important to us, 'what shall we do and how shall we live?'

Bhaskar shows how science, and more generally knowledge, *can* help us with this question, and thus resumes the great tradition of philosophy exemplified by Socrates and Spinoza, which endeavours to be at once logical and scientific in method, and (if I may be allowed the word) existential in content.

(A) 'But the human sciences are not only concerned to explain "*cognitive* ills"', says Roy Bhaskar, and goes on to list numerous others under the categories practical ills, communicative ills, irrationalities and injustices (SRHE, p. 191). Insofar as these all involve some avoidable frustrations of human needs, one can draw a parallel with the explanatory critiques already discussed: social science does not only bring into view beliefs, their falsehood and their causal relations with the social structure; it also reveals human needs, their frustration, and the relation of those needs and that frustration to the social structure. This aspect of social science is also critical of its object. For while there is no formal contradiction involved in admitting that something is a human need but denying that it should (other things being equal) be satisfied, such a position can be said, in a looser way, not to *make sense*. One could here appeal to an

(inverted) use of G.E. Moore's famous argument against naturalism in ethics. Moore claims that any definition of 'good' must be mistaken, since it always *makes sense* to say (for instance, to a utilitarian) 'I know this action will promote the greatest utility, but is it good?' It seems to me that, once it is conceded that, for example, children have a basic need to play (will have wretched childhoods and become inhibited and miserable adults, lacking in skills and social skills if they are prevented from playing), then it makes no sense to ask 'but *ought* children be allowed to play?' — unless on the basis of some exceptional 'other thing' that is not equal (e.g. 'in the present famine, we will all starve if the children don't spend all their time helping to get food').

Social sciences, then, generate practical emancipatory projects by showing there to be (a) a need, (b) some obstacle preventing its satisfaction, and (c) some means of removing this obstacle. This is not a matter of mere technical imperatives, coming into play only *if* you want the projected good; given that a social science can tell us not only about the means of satisfaction but also about the need itself, it may ground *assertoric imperatives*, i.e. *since* you need this, remove that obstacle thus.

As in the case of cognitive explanatory critiques, there may be a functional as well as a causal relation between the frustrated need and the frustrating institution. The frustration of the need may be not only *generated by* some social institution, but also *necessary for the reproduction of* that institution. So the exploitation of frustrated needs is not always a mere epiphenomenon of the frustration (like commercial pornography for the sexually frustrated); for instance, the frustrated need of workers for possession of the means of their labour is the essential foundation of the system (capitalism) that perpetuates that frustration, since it is what drives them to sell their power to work to a capitalist. Thus there is an exact parallel with the cognitive explanatory critique, with frustrated need replacing false belief.

A few words are required here about the Marxian notion of 'contradictions of capitalism'. It is clear that Marx's intention in using this concept is to provide an explanatory critique of capitalism; his claim to be a 'scientific socialist' largely means that his case for socialism consists entirely in an explanatory

account of capitalism — and this explanation is a critique because it unearths contradictions in capitalism.

Some of these contradictions are cognitive, as we have seen — involving 'contradictions' in the logical as well as the dialectical sense. But there are also what may appear at first to be two other kinds: (1) contradictions between the requirements of capitalism and those of human needs — e.g. 'alienation', exchange-value/use-value contradictions, exploitation; and (2) internal contradictions, causing capitalism to malfunction, in its own terms: for example, the falling rate of profit, overproduction crises. In fact I think that all contradictions have both aspects, since (i) (a) the needs with which capitalism contradicts are not abstract human needs in general, but the historically complexified needs of people in capitalist societies, and (b) capitalism presupposes for its own functioning these needs which it frustrates. In both these ways, the needs are *internally* related to capitalism, though this does not mean that they are wholly constituted by it. Furthermore, (ii) internal malfunctions (a) are arguably only possible given that people are not infinitely malleable, since their needs are rooted in biology, and hence have a 'coefficient of adversity' to full incorporation into the functionality of the system; (b) are only objections to the system given their adverse effects on human needs. I wouldn't worry at all about a stock market crash if it didn't lead to unemployment, etc.

Hence, accounts of the contradictions of capitalism are a subset (probably the most important subset) of need-based explanatory critiques. And the *cognitive* contradictions of capitalism are essential to that system precisely because they obscure the need-based ones. It should be added that Bhaskar regards it as unlikely that a unified notion of contradiction can be arrived at (SRHE, p. 197); rather, a number of kinds of non-cognitive contradiction may be 'clustered around' the notion of logical contradiction from which they derive their name, united perhaps by a sort of 'family resemblance'.

(B) Spinoza's ethics is noteworthy as being, on the one hand, a system of ontology and psychology motivated entirely by moral concerns, and, on the other, a system of morality entirely in the indicative; Spinoza does not say 'we ought to . . .' but 'the free person/one who is led by reason will . . .'. This is made possible by the following triad of doctrines: (a) that an emotion can only be overcome by another emotion; (b) that emotions are not simply data, which cannot be criticized — they involve beliefs, which may be more or less adequate, and the emotion consequently more or less rational; and (c) that we are free to the extent that we have rational emotions, based on adequate ideas. This we achieve not by an 'act of will', taking sides with existing rational emotions against existing irrational ones, but by a work of reason transforming irrational into rational ones, by substituting adequate for inadequate ideas.

This may be restated in these terms: the work of personal liberation is a work of transforming one's emotions by means of explanatory critiques of them. As one comes to understand one's emotions better, one can eliminate contradictions and misconceptions from them. This understanding is never achieved by pure 'introspection', for our emotions are what they are because of our interaction with the world. The increase of self-understanding is equivalent to the increase of our powers both to act on the world, and to be affected by it through the senses.

For Spinoza, the explanatory critique, if genuinely seen to be true, of itself transforms the emotion; for once we see the beliefs involved in an emotion to be ill-founded or inconsistent, those beliefs are necessarily changed, and the emotion thereby transformed. Here at least, explanation *is* emancipation; however, the production of the explanation is a process that also occurs under a non-cognitive description: an increase in interactive powers.

If Spinoza's idea of the work of moral thinking is that of explanatory self-critique, his case against rival moral outlooks is also a sort of explanatory critique. For he regards, for instance, the kind of moral blame which assigns ultimate responsibility to agents as of a piece with vindictive emotions, and to be undermined along with them by an understanding of human motivation.

Roy Bhaskar does not explicitly draw parallels between his theory of explanatory critiques and Spinoza's approach to moral questions, but I believe Spinoza provides the best historical paradigm for that 'readjustment' of ethics that transcendental realist ontology requires, and I think that the possibility of a neo-

Spinozist ethics opened up by critical naturalism is a fruitful and exciting one. And Bhaskar *does* refer, as an example of explanatory critique, to another project of personal self-emancipation the similarity of which to Spinoza's has often been noted: psychoanalysis. Freud himself wrote 'I readily admit my dependence on Spinoza's doctrine', and though he 'did not seek philosophical legitimation', he 'never claimed priority' (letter to Dr Lothar Bickel, 28.6.1931, quoted in Hessing, ed., *Speculum Spinozanum*, p. 63).

Let us consider a long passage in which Freud explains to his patient the Rat-man how psychoanalytic treatment works. If the passage lacks the finished look of some of Freud's accounts, since it is an informal exposition in response to the Rat-man's questions, it has the advantage of being at once a concrete piece of therapeutic work with a concrete symptom, and an explicit application of Freud's general theory of our mental structure. The mismatched emotion which sets off the discussion was the Rat-man's self-reproach at not having been present at the moment of his father's death — a reproach so intense that it made him unable to work.

> When there is a *mésalliance*, I began, between an affect and its ideational content (in this instance, between the intensity of the self-reproach and the occasion for it), a layman will say that the affect is too great for the occasion — that it is exaggerated — and that consequently the inference following from the self-reproach (the inference, that is; that the patient is a criminal) is false. On the contrary, the physician says: 'No. The affect is justified. The sense of guilt cannot in itself be further criticized. But it belongs to another content, which is unknown (*unconscious*), and which requires to be looked for. The known ideational content has only got into its actual position owing to a mistaken association. We are not used to feeling strong affects without their having any ideational content, and therefore, if the content is missing, we seize as a substitute upon another content which is in some way or other suitable, much as our police, when they cannot catch the right murderer, arrest a wrong one instead. Moreover, this fact of there being a mistaken association is the only way of accounting for the powerlessness of logical processes in combating the tormenting idea.' I concluded by admitting that this new way of looking at the matter gave immediate rise to some hard problems; for how could he admit that his self-

> reproach of being a criminal towards his father was justified, when he must know that as a matter of fact he had never committed any crime against him?
>
> At the next sitting the patient showed great interest in what I had said, but ventured, so he told me, to bring forward a few doubts. — How, he asked, could the information that the self-reproach, the sense of guilt, was justified have a therapeutic effect? — I explained that it was not the information that had this effect, but the discovery of the unknown content to which the self-reproach was really attached. — Yes, he said, that was the precise point to which his question had been directed. — I then made some short observations upon *the psychological differences between the conscious and the unconscious*, and upon the fact that everything conscious was subject to a process of wearing-away, while what was unconscious was relatively unchangeable; and I illustrated my remarks by pointing to the antiques standing about in my room. They were, in fact, I said, only objects found in a tomb, and their burial had been their preservation: the destruction of Pompeii was only beginning now that it had been dug up. — Was there any guarantee, he next inquired, of what one's attitude would be towards what was discovered? One man, he thought, would no doubt behave in such a way as to get the better of his self-reproach, but another would not. — No, I said, it followed from the nature of the circumstances that in every case the affect would for the most part be overcome during the progress of the work itself. Every effort was made to preserve Pompeii, whereas people were anxious to be rid of tormenting ideas like his. ('A Case of Obsessional Neurosis', pp. 313–15).

Let us take this point by point. (1) We start with an inappropriately intense emotion. It is recognizedly irrational, since the affect (the feeling of self-reproach — hereafter 'F') is recognized to be stronger than warranted by the idea to which it is attached (of his absence from his father's deathbed — hereafter 'Y'). (2) Freud postulates another idea, X, which is the *real* cause and object of F, since something must explain it, and Y does not. (3) X, since it is unknown yet effective, must be repressed; thereby F, dissociated from it, was *displaced* on to Y. (4) In reply to the question 'how will the discovery of X (to which *ex hypothesi* F was appropriate) help get rid of F?', Freud answers that only the unconsciousness of X enabled it to persist unaltered. Once conscious, it would be subject to 'wearing-away'. Freud goes on to identify the unconscious with the

infantile, preserved by repression. X turns out to be an infantile wish that his father would die. Once the infantile wish is brought into adult consciousness, it loses its terrors; the original emotion Y+F has disappeared, and the infantile residue X+F can be coped with when brought into the context of an adult's sense of reality and proportion.

This constitutes a kind of explanatory critique in which the emotion Y+F is (a) characterized as mismatched, (b) explained as a displacement of X+F, (c) replaced by abreaction of X+F, which is then (d) weathered away by the 'daylight' of reason.

It is worth mentioning that, along with the often discussed assumptions that there are mental causes and unconscious ideas, there is here the interesting assumption that ideas and feelings can be mismatched, and underlying this an ideal of rationality as the alignment of the relation of mental phenomena to their causes with their relation to their objects. When it is revealed that the object of an emotion or belief is not its cause, rectification is in order. In this respect, Freud is a card-carrying Spinozist. But in two ways his account is less 'cognitivist' than Spinoza's. First, in that for Freud the mere *knowledge* of the true origins of the mismatched emotion in repression and displacement will not by itself undo these processes; unless that knowledge has so to speak come up from the unconscious, complete with its attendant feelings, assent to it will merely be a 'second registration' of the knowledge, not an abreaction capable of effecting a cure. Second, in that the emotional tie with the analyst, 'transference', is one effective element in the process of bringing unconscious ideas into the light of day, which can often succeed where pure Spinozist reflection would fail.

(C) In the section 'Reason and the Dialectic of Human Emancipation' (SRHE, pp. 180–211), Bhaskar lists and discusses seven levels of practical rationality. The fourth is explanatory critical rationality of the sort already discussed. The fifth and sixth strike me as being special cases of it rather than distinct levels: under level V, 'depth-explanatory critical rationality', he discusses Marx's account of ideology, with its characteristic elements of theoretical ideology (the rival explanatory account) which reflects and rationalizes the practical consciousness which is itself a mystifying reflection of the social

reality of which it is a necessary element. The passage on level VI, 'depth-rationality', is an account of a possible kind of depth inquiry undertaken by two people with a view to understanding and remedying some frustration to which one of them is subject. This looks built to accommodate psychoanalysis, though the account is generalized to include the case where the frustrating agency is an external circumstance rather than a neurotic symptom. The open-ended nature of this quest, with its possibilities of discovery and disillusion, is brought out. These two levels are grouped together as emancipatory reason, presumably because both set out to explain with the explicit intention of thereby helping the work of emancipation. Level VII, 'historical rationality', is concerned with questions about the unactualized powers and transformative tendencies already present, which may generate the possibility of human emancipation. It is mentioned only to say that these questions can only be answered in the context of some theory (presumably a theory of history as the progressive realization of human potential, after the manner of Kant, Hegel or Marx).

I now turn to the 'lower' levels, for even these familiar forms of practical reason foreshadow the critical and emancipatory reason that has been our concern so far. The first is technical rationality — the only sort of practical rationality known to positivistic 'neutral science': instances of this concern means to some external end. Bhaskar says that they only seem to do more than this if they implicitly suppose human purposes. It may be noted in passing, though, that if human sciences provide an *explicit* account of such purposes, they may transform technical into assertoric imperatives by supplying an extra (factual) premiss, and are then on their way to the level of non-cognitive explanatory critiques. However, Bhaskar makes a different point about the potential of instrumental rationality, which takes us to level II: 'explanatory knowledge increases the range of real (non-utopian) human possibilities, which may of course also mean decreasing the range of imagined ones, by showing certain of these to be purely imaginary' (SRHE, pp. 181–2). Such knowledge is empowering to a movement of the oppressed. Of course, it may also be empowering to the oppressors, but not unambiguously so, for the latter have an interest in obscuring the real range of available possibilities from

the oppressed, hiding possibilities of a better life that depend on transformed structures, and holding out unreal possibilities of a better life within existing structures. This is not necessarily a cynical dodge: the oppressors may equally obscure the unwanted possibilities from themselves. But all this means that even purely instrumental knowledge (including, it might be added, some supplied by the *natural* sciences) is not necessarily neutral. As Althusser put it: 'true ideas always serve the people; false ideas always serve the enemies of the people' (*Lenin and Philosophy*, p. 24).

Finally, there is level III, intra-discursive critical or practical rationality: every theory implies criticism of incompatible theories and the practices based on them. 'X is false' entails 'don't believe X', and, other things being equal, 'don't act on X'. This point, which is the first condition of explanatory critiques proper, also has some practical import even in the absence of an explanation of the disproved or contested beliefs.

All seven of these levels share a common structure, in that they are ways in which an already existing and ongoing practice is transformed by a theory which supplements or contradicts some of the ideas implicit in the practice. None of them can create a practice out of nothing, but all of them can transform practices in ways that could not have occurred without them. The 'primacy of practice' holds, historically and ontologically. But it does not imply that theory is redundant or epiphenomenal or merely explicative or neutral as to ends.

A Non-Cognitive Model of Emancipation; A Cognitive Model of Ethics?

It is clear that Roy Bhaskar is anxious to avoid the misreading of his theory of human emancipation which, on the basis of the prominence given to explanatory knowledge in that theory, would take it as a purely cognitive process. There are of course special cases where it is. When it is *just* a set of false beliefs that enslaves, their replacement by true beliefs *is* liberation. But the vast bulk of human bondage, misery and oppression is not like that. The extension of explanatory critique from cognitive error

to unsatisfied needs makes it clear that false belief is not the only chain that binds us, and it is massively outweighed by others in terms of urgent human problems. Peasants who grow food they cannot afford to eat, unemployed workers, homeless families, bullied wives, tortured prisoners, may all know exactly what would make them free, but lack the power to get it. And Roy Bhaskar has something to say about the nature of emancipation, based on his conception of the way we interact with the structured world outside us.

But first it should be said: (1) that though the oppressed *may* understand their oppression quite well, they may not. In the example from Marx, workers who take wages to be payment for work done may or may not perceive their wages as unjust, and would most likely welcome a rise, but will not recognize their systematic exploitation, rectifiable only by a change of social structure. They will not undertake political action to take over the means of production so long as they see the existing system as only accidentally exploitive. Their cognitive deception is the first line of defence against their social emancipation. Hence cognitive enlightenment is a necessary, though not a sufficient, condition of their emancipation.

(2) It should also be said: that workers who *have* seen through the wage form to the relations of exploitation that lie behind it are so far unfree, that they have an uphill struggle ahead, and may be less 'happy' in a superficial sense than the forelock-touching Tory Working Man; '*dissonance*, not liberation . . . may be the immediate result of enlightenment' (SRHE, pp. 204–5).

Yet to a degree they are already more emancipated. No one with any self-respect would prefer to be a contented dupe than a clearsighted dissenter. But it remains true that the main part of the work of emancipation is not cognitive, but consists in toil and trouble, conflict, changes in power relations, the breaking up of some social structures and the building up of others.

The etymology of the word 'emancipation', almost always favoured by Roy Bhaskar over its near-synonym 'liberation', emphasizes more than the latter the idea that it is always *from* some previous bondage that one is emancipated. Hence it is distinguished from simple *empowering*, which may also, of course, be the result of (applied) new knowledge.

Bhaskar characterizes emancipation in the following way:

It is my contention that that special qualititative kind of becoming free or liberation which is *emancipation*, and which consists in the *transformation*, in self-emancipation by the agents concerned, *from an unwanted and unneeded to a wanted and needed source of determination*, is both causally presaged and logically entailed by explanatory theory, but that it can only be effected in *practice*. Emancipation, as so defined, depends upon the transformation of structures, not the alteration or amelioration of states of affairs. In this special sense an emancipatory politics or practice is necessarily both grounded in scientific theory and revolutionary in objective or intent. (SRHE, p. 171)

There are a number of points to ponder here.

1. The italicized phrase 'from an unwanted and unneeded to a wanted and needed source of determination' encapsulates a theory of what freedom is. It is 'no more the simple recognition [of], than escape from, necessity' (SRHE, pp. 170–1). That is to say (taking the points in reverse order), freedom cannot mean that we escape the causal order of the world, not only because of the intrinsic incredibility of such a notion, but because (a) an uncaused action could no more be *my* action than something that happened to me without my will would be. My actions are those that I — my character, opinions, desires — cause. Certainly, as has often been pointed out, an action has reasons, not just causes — otherwise it would not be an action. But those reasons must *also* be the causes of the action; for if they are not, then either that 'action' is uncaused, i.e. an accident, and therefore not an action, or it is caused by something other than the reasons for it, in which case the 'reasons' are mere rationalizations, and the 'action' once again a mere happening, that we mistakenly think we cause. We are free only if our reasons have effects — and what has effects is a cause. (b) If we are either to know or to act upon the world — and neither is possible without the other — we must both be affected by the world through our senses, and affect the world through our bodily movements. To do either, we must be no disembodied spirits, but made of the same stuff as the world about us, subject to the same causal laws.

So freedom must be 'in-gear' rather than 'out-of-gear' freedom; it is not a matter of disengaging ourselves from the

world so that it gets no grip upon us — for by the same token, we would get no grip on it. We do not escape from necessity in that what we do we do in ways governed by causal laws.

If we *could* disengage ourselves *mentally* from the causal nexus (for it hardly makes sense to think we could disengage ourselves physically), we would actually not be *escaping from* necessity, but rather simply recognizing it — the former of the notions Bhaskar dismisses. Such recognition of necessity would no more be freedom than the prisoner who 'comes quietly' is freer (though he may be less bruised) than the one who resists arrest. However, it is worth mentioning in passing that Engels, to whom Bhaskar attributes this conception of freedom, meant something else by this phrase 'recognition of necessity'. He did not mean accepting being dragged along willy nilly; the image is rather of the yachtsman, whose knowledge and skill enable him to sail near the wind, while the person who does not know how to use the force of the wind will be driven in whatever direction it happens to be blowing. 'Necessity' here, as for Bhaskar, stands for the necessary tendencies of things, not some inevitable fate.

2. The idea of a 'wanted and needed source of determination' is so strikingly discordant with 'out-of-gear' concepts of freedom that it warrants comment. The adherent of 'out-of-gear' freedom may see this idea as just as inadequate as the 'coming quietly' idea of freedom. To extend the metaphor: you get arrested by a decent cop instead of by a real pig. But this rests on the misunderstanding of causation as a kind of compulsion by an outside agency. In special cases, indeed, a causal mechanism may be an alien force, conquerable or not. But among the 'sources of determination' are the laws of our own being, and of the environment which makes it possible for us to be. To take an everyday example, I have not chosen the fact that tea refreshes me, while coffee sets off a slight allergic reaction. But given this fact, I am freer if I can find somewhere that serves tea than I am if I can only get coffee. While this is not an instance of emancipation, the following may be. (a) (At the personal level) if I am cured by psychoanalysis of a disabling obsession or inhibition, I am no less necessitated to act without it afterwards than to act in accordance with it before. Yet I am surely freer. (b) (At the micro-social level) if I am part of a strife-torn household

that makes daily life a nightmare for me, I am less free than as part of a loving one, and may emancipate myself by getting out of the former into the latter; yet each will involve its own kind of constraints (and corresponding enablements). (c) (At the macro-social level) different kinds of society are governed by different kinds of laws. I don't only mean legislative enactments (though of course that is also true), but social mechanisms generating different possibilities and tendencies. The future of the area where I live may be determined by market forces, or by plans made by a neighbourhood meeting. In the latter case I can participate in determining my future environment, and live in some confidence that it will not become uninhabitable. Of course, I lose the possibility of speculating on the property market. But in both cases, there is a generative social mechanism determining what happens — and in both cases, that mechanism works only through the actions of human agents. And of course, in both cases, there are material constraints: build a house upon subsiding subsoil, and it will crack. Yet the transition from market forces to neighbourhood meeting would clearly be experienced by most people as an emancipation.

It should be evident that emancipation into such 'in-gear' freedom can't be achieved either by pure cognitive enlighten-ment or any other purely 'inner' or 'mental' change. It 'can only be effected in *practice*', i.e. it requires hard work, transforming recalcitrant structures, with the technical and social means at our disposal, into other, more congenial structures. This brings us to the third point.

3. There is an important distinction between 'amelioration of states of affairs' and 'transformation of structures'. There can of course be freedom-enhancing ameliorations of states of affairs. I would like a holiday in Greece next year, but can't afford it; if I had a rise in salary, I could afford it, and so that amelioration of my state of affairs would to a degree increase my freedom. Furthermore, it might take practical activity to achieve this, whether collective (trade union militancy) or individual (getting promotion). But it would be absurd to call this 'emancipation'. This term implies that there are objectively existing, effective, relatively enduring, but alterable structures constraining one's possibilities: political tyranny, class exploitation, apartheid, patriarchy, bureaucracy, press monopolies, the property

market, and so on. Emancipation involves transforming them; and the whole depth-realist theory indicates that there is a real hiatus between reforms at the level of the actual, retaining existing structures (e.g. pay claims, tax reforms, electoral reform, a bill of rights) and structural changes (e.g. the socialization — or privatization — of the economy, the transfer of political power from one class to another, the break up of the nation-state). There is a hiatus in the sense that one will never change structures by the cumulative effect of reforms in accordance with those structures: tax reforms will not abolish class privileges, and so on.

I should say here that, though I have given examples that I consider plausible, realist philosophy cannot as such tell us which changes are structural, which not; only empirical social-scientific inquiry can do that. And there are disagreements about this issue. For example, I have heard it said that the replacement of patrilineal by matrilineal inheritance of surnames would have deep structural effects, though I myself doubt whether any linguistic reforms will even ameliorate states of affairs, let alone transform structures — more likely they will preserve them by obscuring the fact that nothing has changed.

Nevertheless, there is a certain kind of reformist politics which does presuppose that whatever social transformations are required can be made without at any stage implementing 'structural reforms'. In the ironic words of Leon Rosselson's song, 'We'll change the country bit by bit/ So nobody will notice it/ Then ever after, never fear/ We'll sing The Red Flag once a year'. Depth realism, by contrast, draws attention to the same facts as Tawney's remark that you can peel an onion leaf by leaf, but you can't skin a live tiger claw by claw. If some changes can only come gradually, there are others that can only come all of a sudden. Hence, 'emancipatory politics or practice is necessarily ... revolutionary in objective and intent'. 'Revolution' here refers to the necessarily deep and sudden changes; it does not necessarily imply violence (except in the sense of the ancient distinction between natural and violent motion), though no one but a pacifist or a Hobbesian can doubt that violent revolutions are sometimes necessary. But it is clear that this notion of structural transformation sits easier with Marxist than with Fabian politics.

One reservation needs to be made here though. There is a certain kind of Marxist politics which sees emancipation as an all-or-nothing thing; it is assumed that nothing short of socialism is any sort of emancipation worth having, while the achievement of international socialism would emancipate all and completely, so that thereafter only ameliorations of states of affairs would be required. Bhaskar's definition of emancipation cannot be tied to any such all-or-nothing conception. One can transform *some* of the many unwanted and constraining structures, without transforming them all; and this can still be distinguished from mere amelioration of states of affairs. We have many instances of such partial emancipations: the great bourgeois revolutions which emancipated Europe from feudalism, but delivered it over to capitalism; the national liberations of the twentieth century, which ousted colonial rule, yet often replaced it by military regimes or corrupt bureaucracies; the overthrow of fascism, which everywhere replaced it either by bourgeois democracy or bureaucratic 'state socialism'; the political emancipation of Eastern Europe in 1989–90, which has for the most part led to economic and social developments which are the opposite of emancipatory. As yet we have *no* instance of 'total emancipation', and it would be utopian to predict its possibility. Most likely, emancipation will always occur as a multiplicity of partial emancipations. This does not preclude the possibility that some repressive mechanisms may be explained in terms of other, more basic ones: imperialism, and modern forms of sexism, may be explained in terms of capitalism, for example. But this is a substantive issue for social science, and cannot be resolved by philosophy. At most, Bhaskar's theory may suggest a framework into which we can fit the Marxist notion that the economic structures are 'determinant in the last instance', though not necessarily 'dominant', should concrete research justify it. I mean the notion that generative mechanisms are stratified, so that, on the one hand, they *conjointly* determine events, in no fixed proportion; yet on the other, one of these mechanisms may be rooted in, emergent from, and explained by another.

4. In the passage quoted from SRHE, p. 171, Bhaskar also says that emancipatory politics is necessarily 'grounded in

scientific theory'. Why should this be? The argument so far has shown such grounding to be possible, rather than necessary. But if emancipatory politics means transforming structures, it must be based on knowledge of those structures. It is such knowledge that transforms the will to ameliorate states of affairs — which is after all the necessary *motive* of emancipatory politics — into the project of transforming those structures which generate the unwanted states of affairs. Social evils may stare one in the face, but social structures don't. One can see people sleeping on the streets, and listen to their complaints; but one has to do research to understand the market mechanisms which cause this tragedy, and how they can be changed. William Morris is reported to have said that he didn't need Marx's *Capital* to tell him that the rich robbed the poor; if he nevertheless read his copy of *Capital* until it fell apart, perhaps that was because the *mechanisms by which* the rich rob the poor need to be analysed and understood if we are to abolish them.

If it is clear by now that Bhaskar's conception of human emancipation is not a cognitive one, I think it is also becoming clear that his paradigm of practical reasoning — of ethics — *is* a cognitive one. For while 'there are other good things in life apart from explanatory knowledge' (SRHE, p. 171), and most of ethics will be talking about those good things, not *about* explanatory knowledge, its kind of talking will be describing and explaining, not simply prescribing or evaluating. We talk most usefully about values when we do so by talking about facts. Unless people had values already, no amount of 'edifying discourses' could induce them, but given that people unavoidably have values, the way to change those values for the better is by increasing knowledge, both descriptive (e.g. what it is like to be a forest-dweller turned out of one's home and livelihood by a rancher), and explanatory (e.g. how come ranchers have the motive and the power to turn the forest-dwellers out?). Bhaskar has not elaborated this idea of a cognitive paradigm of ethics, which, so far as personal ethics is concerned, might look very like Spinoza's. But the possibility of such an ethics is implicit in his thought.

Some of the essential points of Roy Bhaskar's view of emancipation are summed up in his list of five conditions of the possibility of emancipatory practices (SRHE, pp. 210–11).

First, *reasons must be causes*, or discourse is ontologically redundant (and scientifically inexplicable).

As we have seen, our reasons for acting must have real effects through our action, co-determining events in the open systems of the world with divers other causes which pre-exist them and operate alongside them.

Second, *values must be immanent* (as latent or partially manifested tendencies) in the practices in which we engage, or normative discourse is utopian and idle.

This precludes the 'theoreticism' or 'scientism' criticized above, according to which theory can conjure values out of its own hat, where none existed before. We are all engaged in practices prior to the initiation of theory, and all practices necessarily involve and secrete values; the initial motive both for theory and for the transformations of practices that it effects must lie in those values. This also precludes the Platonist or Kantian location of values in an ideal or noumenal world distinct from the world in which we live, along with 'Cheshire Kantian' views such as emotivism or prescriptivism. And at the political level, it precludes the utopian project of basing programmes on how people might be in the future, rather than on what they need now.

Third, *critique must be internal to* (and conditioned by) *its objects*, or else it will lack both epistemic grounding and causal force.

That is to say, *if it is to have emancipatory effects*, an explanatory critique must be part of the society of which it is a critique. An explanatory critique of the institutions of ancient Babylon will hardly be emancipatory in modern England, or even modern Iraq; a Martian sociologist could report back on the state of the modern world without it having any effect on the world at all. And if the critique must be made from within, it is subject to all the same pressures that distorted the ideas that are the object of its critique. Hence it must always be ready for *self*-critique, and consequent self-revision. The point about 'epistemic grounding' is more contentious, and extraneous to the issue of

emancipation. It suggests that the explanatory critique of Babylon by a modern or Earth by a Martian are not just ineffective, but impossible.

Fourth, at the emancipatory moment, there must be a *coincidence of subjective needs . . . and . . . objective possibilities*, already at or close to their historical conditions of realization, as the articulated and achievable goals of groups, rather than merely the abstract properties of structures.

This specifies one of the non-cognitive, or only partly cognitive, conditions of emancipation. People must actually *feel* the need for change — and for just that change that is a real emancipatory potentiality of the time. Only then can an emancipatory programme that is at once realistic and popular — and hence actualizable — be projected.

Finally, for emancipation to be possible, *knowable emergent laws must operate*.

This is perhaps the most surprising claim, for it amounts to saying that idealist and reductive materialist philosophies are incompatible with human emancipation, in that, if they were true, that emancipation would not be a possibility. Let us take reductive materialism first.

Suppose that, while everything is governed by physical laws, there are no laws at the level of social existence, i.e. that there are no irreducibly social mechanisms; what would be physically possible would be socially possible, and the only way to apply knowledge in transforming social institutions would be by redescribing them as physical entities, explaining them physically and acting upon their physical structure. But (a) for most examples one can think of, such a manner of transforming social structures is inconceivable; (b) even if possible, it would presuppose a prior identification of the entities to be transformed under a social description, and a decision to transform them because of what they are under that description — without the aid of any explanatory theory of them under that description; (c) such a transformative practice, even if possible, would be systematically indifferent to the *social* properties of the

entities affected by the transformation process, and hence manipulative rather than self-emancipatory.

For the most part, the political effect of denying emergent social laws is to uncritically use pre-scientific theories full of unexamined assumptions about social causality, and at the same time assume that anything that is physically possible is socially possible. Thus the crucial fact that some physically possible and humanly desirable outcomes (e.g. the bringing together of unused resources, unemployed workers and unmet needs) may be impossible within a given social structure (e.g. a market economy) is obscured.

Idealism is, on the one hand, theoretically, unable to explain the constraints which make emancipation necessary, and, on the other, practically, destined to preserve real constraints from which we could have emancipated ourselves, by proclaiming an emancipation entirely internal to 'the mind' or 'discourse'. In times of difficulty for liberation movements, there will always occur a secession of erstwhile partisans of emancipation into such movements for 'inner' liberation, and this was noticeably the case in the 1980s. It is for this reason both that realist philosophy has been very much against the stream in 'radical' circles in this period, and that it has itself been a major political intervention as an antidote to this 'retreat to the inner citadel'.

Philosophy and Socialism

While Roy Bhaskar makes no secret of his socialist beliefs, his account of human emancipation is in very general terms, not specifically socialist ones. The question has often been posed, what is the relation between 'critical realism' and socialist politics? In this section I try to answer this question. The first thing to say is that the relation is not one of entailment. It is perfectly logically possible to combine such a realism with right-wing or middle-of-the-road politics. Indeed, no philosophical position — according to the conception of philosophy in question — entails any specific political position. Political positions, if rational, are arrived at by means of explanatory critiques of the societies they pertain to; these are the work of empirical social sciences. Marx was right to think that the

grounds for socialist politics were in the 'critical analysis of capitalist production'. Whether or not the content of his politics was correct depends on whether that analysis was correct. This is a substantive social-scientific issue, which cannot be resolved by philosophical argument.

However, there are two ways in which the realism and the socialism are linked. First, there are a number of arguments commonly used for certain non-socialist positions, or for versions of socialism which hope to avoid confrontation with and transformation of existing structures, which arguments are undermined by transcendental realism. We have already seen two of them: the gradualist argument that states of affairs can be ameliorated in all requisite ways without transforming any structures; and the idealist 'radicalism' which seeks to liberate the world by changing the colour of our discursive spectacles. I will mention one more here: certain sections of the political right, sometimes called the 'libertarian right', also claim to be working for human emancipation. There is another kind of rightism, which appeals not to liberty but to law and order, the national interest, traditional values, and so on. Since this kind of rightism — which may very well be realist — does not use the language of emancipation, I do not need to discuss it here. The 'libertarian right', however, would find it very difficult to make a plausible claim to be on the side of emancipation without presupposing a specific theory of human nature and social structure: that people are autonomous individuals, and society exists only by virtue of their voluntary or compelled relations (i.e. relations that are in each case the expression of *someone's* will, so that one person's unfreedom always results from another's bullying). Now the transformational model of social activity refutes this position, while taking into account the facts that lend it plausibility vis-à-vis holistic conceptions. It thus leaves the libertarian rightist without any ontological ground to stand on; an alternative defence of libertarian rightism would have to be found if that position were to remain in the field, and it is difficult to imagine what such a defence might be.

In addition to these refutations of alternative political positions, there is another relation between Bhaskar's philosophy and a certain kind of socialist politics. I am referring to the *homology* which exists between the transcendental realist

world-view and a certain political model. I am mindful of the fact that homologies can be misleading, and we do well to treat them warily. Some homologies have been very important in the history of ideas, yet of no philosophical importance; the fact that one set of ideas is homologous with another, *true*, set of ideas is no evidence for the truth of the former set. For instance there is a homology between Newton's atomist mechanics and the 'abstract individualist' conception of society; yet Newton's mechanics was an excellent scientific theory, which enabled much new knowledge to be discovered, even though it finally turned out to be inadequate, and in some respects was even contradictory; abstract individualism, on the other hand, has generated nothing but intellectually infertile and humanly destructive errors. Moreover, Newton's justified prestige has lent credence to these errors.

However, in the present case, I shall suggest that, while there is certainly a homology, there may be more than that in the offing. But first, the homology: according to transcendental realism, there are hierarchies of structures in the world, e.g. molecules are composed of atoms, cells of molecules, organisms of cells, societies of people — and in no case are these 'wholes' reducible to their parts, or the parts to their wholes. There are irreducible mechanisms existing at each level, which could not for the most part be predicted from knowledge of the higher- or lower-level mechanisms. This view contrasts with a number of one-level ontologies, which claim either that parts are mere aspects of some whole, so that ultimately there is only the Absolute, of which everything is an aspect; or that wholes are mere collections of parts, understood only when broken down into their components, which alone are ultimately real; or that some intermediate level of entity (e.g. 'selves') are the only reality, their parts being mere aspects, and the larger entities which they make up being mere collections. The common assumption of these three ontologies — that there must be one and only one ultimately real level — is homologous with a common assumption in political philosophy, namely the idea of *sovereignty*.

It is assumed by many writers — Hobbes and Rousseau, Hegel, but also modern political commentators discussing such issues as Britain's place in Europe, or home rule for Scotland and Wales — that there must be sovereignty at some one political level, and that if, for instance, it is located in the nation-state, neither smaller local units nor international organizations can have any but a derivative and retractable power. Likewise, in debates about the politics of economics (public versus private ownership, centralization versus de-centralization, market versus planning, etc.) it tends to be assumed that there must be one level of *units*: that while the 'firm' may be a multinational corporation, a government department or a backyard workshop, there must be some one level at which power is located, outside which there are relations of the market, and inside which there are relations of management. This assumption sets the agenda for debates about possible variants of socialism: it generates the dilemma 'either a command economy, or market relations between separate co-operatives'.

Yet it is no more obvious that such managemental monism is necessary than that some one-level ontology must be true. Even the corporate structure of monopoly capitalism includes relatively autonomous subsidiaries, and models such as guild socialism, though untried, are not obviously impracticable. There may be an alternative to market and command economies alike, in genuinely multi-levelled democratic structures, with real powers located at each level, adequate to deal with the problems of that level.

Likewise with regard to political structures: federal systems in which powers are really located at more than one level (not just devolved from one level to another) have long existed. In other words, even now 'sovereignty' is not in reality absolute. If it is necessary for world peace and ecologically sound planning that, on the one hand, international agencies with real powers be set up and that, on the other, units much smaller than most nation-states (in United Kingdom terms, cities and counties) take over wide fiscal, legislative and economic planning responsibilities, then the illusion of sovereignty as an absolute is a pernicious one.

The homology between such multi-levelled structures of economic and political power, and Bhaskar's conception of a real plurality of causal mechanisms, scientific strata, enduring structures, must be obvious. Is it more than a homology? If we understand political and economic agencies not as mere

repositories of legal legitimacy but as enduring structured entities (government departments, firms, trade unions, political movements, armies), with real powers and tendencies generated by their internal structures and their places in wider structures, then it is plausible to suggest that multi-levelled social organization is an *instance* of multi-levelled causal power. The myth of sovereignty — of the nation-state or of the economic firm — may (over and above its obvious apologetic function on behalf of nation-states and firms) be no more than an instance of the same epistemic 'idol' (in Bacon's sense) as the discredited one-level metaphysical systems: Hobbes's particles, Leibniz's monads, phenomenalism's sense-data, Bradley's Absolute. And the vision of a pyramid of democratic loci of political and economic power, from the street and shopfloor meeting to the planetary plan, may have no inherent impracticability — only the uphill task of overturning the vested interests that oppose it.

Notes

1. I am not claiming that Roy Bhaskar is the only philosopher to have shown how we can argue from facts to values. On the one hand, his arguments vindicate the practice of many philosophers before Hume and Kant who argued validly from facts to values without having to defend this against anti-naturalist critics — as indeed non-philosophers do all the time. On the other hand, there have been a number of defences of fact-to-value argument in recent philosophy; most, I think, rely on some notion of specifically moral facts, and hence are not really naturalistic. One, however, anticipates some of Bhaskar's arguments: Roy Edgley, in his book *Reason in Theory and Practice* and his article 'Science, Social Science and Socialist Science: Reason as Dialectic'.

I concentrate on Bhaskar's version of the argument because of the purpose of this book; I also think it is the fullest and most fruitful version.

2. we have little faith in the 'average sensual man', we do not believe he can do much more than describe his grievances, we do not think he can prescribe his remedies. (B. Webb, *Our Partnership*, entry for 24 December 1894, quoted in SHRE, p. 170n)

=======7=======

Interventions

Following Locke, Bhaskar has described the work of philosophy in relation to science as that of an underlabourer. This may be a humble role compared with the claims of rationalist philosophy to be a master-science, whether (as in classical rationalism) as the method or magic key which unlocks all the doors to knowledge, or (as in the Hegelian tradition) as the summation and apotheosis of all knowledge. But the job of underlabourer is also a useful, indeed essential, one. So this conception of philosophy is not that of a 'purely academic' (in the popular sense) discipline, which 'leaves everything as it is', as Wittgenstein required.

In fact it is clear from the whole tone of Bhaskar's writing that he believes his philosophy can make a difference, can do something valuable for the sciences, and in particular for the human sciences. And as we shall see in this chapter, a number of people working in the human sciences share this view. How is a theory such as critical realism to be used in these sciences? First, there is a common way of posing this question which can only give rise to misconceptions. It is sometimes asked: how can critical realism be *applied* to (or in) the human sciences? The problem with this it that it suggests precisely the classical rationalist notion of a master-method: as if we could first sit down and study critical realism in our armchairs, and then go out into the world or the laboratory and apply it to our chosen subject-matter. A comparison with the way Bhaskar treats the natural sciences should show the error of this conception. He looks at their actual practice, and asks questions about the conditions of its possibility. He argues *from* ongoing scientific practices, not *to* some norms which those practices are required to conform to. And since he holds that the human sciences are in principle sciences in the same sense as the natural ones, this ought to be the relation of philosophy to them too.

Nevertheless, there seems in practice to be a difference between philosophy's relation to the human and to the natural sciences. It may be expressed in this way: while in both natural and human sciences critical realism is of use in helping to answer questions already posed by them, rather than prescribed for them in advance, it is nevertheless the case that the natural sciences do not for the most part need, for their everyday work, to ask these questions. It is rather the 'nocturnal' questions and reflections of the scientists into which critical realism can intervene. It may certainly show some accounts by scientists of the implications of their own practice to be false. But if so, those accounts are false *of* the scientists' practice. The premises of the critical realist critique of science's 'nocturnal' philosophies (for instance, positivism) are provided by the 'diurnal' practice of those same sciences. Into that diurnal practice, the underlabourer's interventions are rarely required. His or her task is, so to speak, to sweep up after the laboratory is closed for the night. Two exceptions may be made to this division of labour. First, on the frontiers of theoretical science, the nocturnal/diurnal distinction sometimes breaks down. We enter a twilight world of highly technical thought-experiments and metaphysical speculation. Here, realist or subjectivist or positivist or operationalist assumptions may affect the outcome. Second, mistaken nocturnal reflections on science may affect the ways in which science is *applied*. Not a mistaken science, but a mistaken conception of that science, may be held partly responsible for our present ecological crisis (though I would argue that in the last analysis its causes are socio-economic).

In the human sciences, on the other hand, the picture is notoriously different. Their best practitioners encounter urgent philosophical questions right in the heart of their work. Their worst practitioners make unquestioned philosophical assumptions at every step, often unaware of the philosophical origins of those assumptions.

A word on the history of the relations between philosophy and the sciences is in order here. It is often remarked that philosophy once covered the subject-matter of all the sciences, and that as the sciences became truly scientific, they declared independence and went their own way, rapidly losing the marks of their origin. In the case of the human sciences,

however, while they may have *declared* their independence in the loudest tones, they continue to be guided (one might say 'trapped') in each case by some approach borrowed from philosophy at the moment of their birth, and thereafter taken for granted. The so-called 'immaturity' of the human sciences is not a matter of their youth — they have been around for about as long as the natural sciences. But they have seemed unable to sever their umbilical cords, substituting an unconsciousness of their philosophical assumptions for an independence of their philosophical origins. The history of economic theory, for instance, while it is marked by a number of theoretical breaks, is also marked by the philosophical character of those breaks. Without its philosophically trained or philosophically inclined recruits — Adam Smith, the Mills, Marx, Jevons, Keynes, Sraffa — where would it be? 'Experimental psychology' is virtually defined by its imitation of the positivist picture of natural science.

In this situation, a philosophical intervention in a human science need not be an unwarranted interference in the affairs of a sovereign state; it may be welcome assistance in the struggle against the 'neo-colonial' dominance of another philosophy.

The persistence of philosophical constraints on work in the human sciences partly explains — given the plurality of philosophies — the pluralism of contesting theories that prevails in these disciplines. For instance, not only positivism but also existential phenomenology has set up colonies on the terrain of psychology. The critical realist interventions that I shall be discussing in this chapter take place on such disputed ground. However, it is no part of the project of critical realism to set up its own colonies. That would be to fall back into the rationalist search for 'applications' (rationalist in origin, though it has often been positivist in content). I know of no such critical realist theories in the human sciences, and I hope for no such thing as a critical realist psychology or economics or linguistics. Where a critical realist intervention has been made in defence of one human-scientific theory against another, the theory is one that has already established itself, prior to or independently of the critical realist intervention (as with Chomsky's linguistics or Freud's psychoanalysis). By the same token, critical realism is not in itself committed to any one theory in any given human-

scientific area. Philosophical criticisms of Freud or Chomsky from rival theories in their respective fields may have been refuted using critical realist arguments; it does not follow that these theories are true, or the best theories in their fields. And there could very well be more than one theory in a given field, mutually contradictory, yet each able to use critical realist arguments against other contestants and critics.

Nevertheless, the use of critical realist arguments to defend a theory from attacks based on (for example) positivist or subjectivist assumptions is one legitimate and important form of intervention. Another is that in which critical realism is not used to sponsor any particular theory, but takes on a purely critical function, attacking the (overt or covert) philosophical premisses of an existing theory. In this case the benefit might be to facilitate the development of work in the discipline concerned away from its philosophical strait-jacket in positivism or whatever, breaking the ground for the emergence of a new autonomous science, without prescribing its content. Such cases are perhaps what Bhaskar means when he says that philosophy can occasionally be not only the underlabourer but also the midwife of a science: cutting the umbilical cord. We should be cautious, however, about extending this metaphor; it could easily lead us to exaggerate what philosophy can contribute to the birth of a science.

In what follows, I shall discuss examples of intervention made in various disciplines in the general area of the human sciences, either by practitioners of the sciences with an interest in critical realist philosophy, or by critical realist philosophers who have a special interest in a particular scientific discipline. In restricting myself to disciplines with a putatively scientific character, and predominantly *human*-scientific in reference, I leave out some interesting uses to which critical realism can be put in other contentious theoretical areas, such as politics, biology, ecology and feminism (see below). No section of this chapter is meant to be a full account of critical realist contributions to the discipline in question. In each case I have taken a single writer as an *example* of how critical realist interventions can be made in a given discipline.

Before moving on to the particular disciplines, however, I would like to mention two critical realist texts which address the

social sciences more generally: Peter Manicas's *A History and Philosophy of the Social Sciences*, and William Outhwaite's *New Philosophies of Social Science: Realism, Hermeneutics and Critical Theory*.

Linguistics: Trevor Pateman

Trevor Pateman is a philosopher by training (and a critical realist), with a wide knowledge of modern linguistics. The theoretical position that he defends in his book *Language in Mind and Language in Society* is that of Chomsky's school. While critical realism has no unique or unbreakable relationship with any substantive human-scientific theory, it has, I think, a fairly obvious natural affinity with Chomsky's theories, which is reflected even in the terminology both approaches use. For critical realism, science discovers generative mechanisms underlying and explaining the phenomena. Chomsky describes his work as 'generative grammar', and aims to discover 'underlying mechanisms' of speech. Of course, Chomsky's early work predates Bhaskar's, and has perhaps influenced the critical realist terminology. But the agreement is not merely verbal. Chomskyan linguists take speakers' intuitions as their data to be explained, exemplifying the 'hermeneutic moment' that critical realism leads us to expect in every human science; and they then *explain* those data in terms of mechanisms that are no mere 'constructs' from the data, but transphenomenal causes of them, opaque to the speakers themselves — just what critical realism says a human science should do.

The main explicit references to critical realism in *Language in Mind and Language in Society* (LMLS) are in defence of Chomsky's practice of science against the attacks of positivism on the one hand, and hermeneutically oriented philosophies, on the other. Nineteenth-century linguists looked for exceptionless regularities and were disappointed to find that there were none. 'Then as now, acceptance of this assessment was seen as a threat to the very existence of linguistics' (LMLS, p. 6). But the threat is unreal since this situation only reflects the openness of the systems linguistics studies, and an explanatory science of

open systems is possible. Like critical realists, Chomsky argues that

> just because many faculties and competences interact in the production and understanding of speech, which is consequently the joint product of an (open) set of complicated mechanisms, that does not mean that as scientists we should not have as a primary goal the isolation and description of the powers and liabilities of individual mechanisms (see Chomsky 1980 [*Rules and Representations*]). The fact that we have to work largely non-experimentally just makes our tasks that much more difficult.
>
>   This all amounts to saying that realism breaks the link positivism insists upon between science and prediction. The two central tasks of science are now seen to be these: (1) isolating and describing the real causal mechanisms at work in producing the world of events; (2) reconstructively explaining past events in terms of the conjunctural operation of particular mechanisms. (LMLS, p. 8)

How this works out is explored in Pateman's chapter 'A Realist Theory of Linguistics' (LMLS, pp. 18–42). This takes the form of a critique of the philosophical parts of an important book on historical linguistics, *On Explaining Language Change* by Roger Lass. Pateman's claim is that however committed they are to positivism as a meta-theory, linguists in practice assume something like critical realism, on two counts:

> First, most linguists both assume [that] and seek to show how *languages* (in the Saussurean sense) or *grammars* (in the Chomskyan sense) are structures, systems or causal mechanisms neither reducible to nor inductively inferrable from the speech events or system sentences which realize them, or are their effects. (LMLS, p. 20)

Second, he argues that linguistic events quite obviously occur in open systems where counter-instances do not disprove the operation of a tendency, since counter-tendencies exist and it cannot be predicted which will prevail.

Pateman then considers an argument of Lass's which illustrates 'the havoc positivism can wreak on science, havoc from which Lass never extricates himself despite his own later

disavowal of positivism — in the end, he is left not with an alternative to positivist metascience, but only the discontents of a failed would-be positivist' (LMLS, p. 23). The gist of the argument is that laws must be deductive-nomological, that is, universal laws from which predictions of instances can be deduced. But laws of linguistic change admit many counter-examples. So their claim has to be merely probabilistic; but since probabilistic laws are not falsified by counter-instances, they cannot really be explanations at all.

Pateman's point-by-point reply to this (LMLS, pp. 24–9) is an account of how the critical realist analysis of laws applies in the case at issue — that is, the law that speakers tend to prefer combinations of sounds that are easy to articulate. This law is quite compatible with instances of actual historical changes occurring in the opposite direction, since other tendencies may also operate. (Indeed, one of the most noticeable recent changes in British English pronunciation must surely be the move towards pronouncing every letter — for example one increasingly hears the 'a' in 'holiday', the 't' in 'often' and even the 'd' in 'sandwiches'. Presumably there is some such explanation as: the petty bourgeoisie favours 'spelling pronunciation', and other classes tend to copy petty bourgeois speech — for which in turn there are doubtless sociological explanations.)

Pateman points out that ease of articulation is definable independently of what speakers do, and that a tendency-statement (such as 'speakers tend to prefer ease of articulation') is not the same as a probability-statement (such as 'more often than not change is in the direction of ease of articulation').

Having used a positivist account of explanatory science to argue that historical linguistics is not one, Lass goes on to reject positivism as inapplicable to historical linguistics. This does not strike me as it does Pateman as 'biting the hand that has fed you for half your book' (LMLS, p. 39). We have already seen that it is a common anti-naturalistic move to accept a positivist account of the natural sciences the more plausibly to plead that the human world is altogether insusceptible to scientific explanation. Lass wants to make way not for hermeneutics, as is so often the case with such anti-naturalist positions, but for 'free will'. Pateman comments:

Lass derives the false conclusion that linguistic change is uncaused from the true claim that human choice operates in the domain of linguistic change. For if human choice operates, it is itself a (or the) cause of change, and that choice is in turn causally explicable by the conscious or unconscious reasons agents have for the choice they make, except in the limit — and rare — case where they 'just choose'. (LMLS, p. 40)

Apart from the avoidance of the positivist science/no science dilemma, another way in which critical realism is useful for work in linguistics is that its doctrines of the stratification of nature and the multiple determination of events in open systems justifies the coexistence of several distinct sciences studying the same phenomena and explaining them through different mechanisms, such that the adequate explanation of what happens must be interdisciplinary.

One virtue in Bhaskar's arguments is that if they can be sustained, then the prospects for a unified linguistics are rather brighter than they must appear at the moment, when natural science oriented linguists are inclined to go off to study the brain; social science oriented linguists end up doing sociology or politics rather than linguistics; and those who want nothing to do with either—or choices turn linguistics into a mathematical science of Platonic entities. For Bhaskar wants us to think of the world as complex and stratified but essentially unified, not bifurcated into nature and culture, brain and mind, mind and society. Hence, our explanatory structures must (for example) allow for interaction between body and mind, nature and nurture, mechanism and rationality. (LMLS, pp. 13–14).

This is a salutary corrective to the departmental arrogance that is a besetting sin of theorists in the human (and, to an extent, in the natural) sciences. This arrogance expresses itself in that series of errors often designated by appending '-ism' to the name of the offending discipline: psychologism, sociologism, biologism and so on. We encounter it in 'applied psycho-analysis' with its accounts of social institutions in terms of unconscious mechanisms, and in that common error of Marxists which Lenin dubbed *Komchvanstvo*:

Whenever any Marxist attempted to transmute the theory of Marx into a universal master key and ignore all other spheres of learning, Vladimir Ilyich would rebuke him with the expressive phrase 'Komchvanstvo' ('communist swagger'). (Trotsky, *Problems of Everyday Life*, p. 221).

All these errors — psychologism, sociologism, biologism, applied psychoanalysis, *Komchvanstvo* — are alive and well in the neighbourhood of linguistics.

Now it might be thought that these inter-disciplinary aggressions and colonialisms were nothing but personal failings in their perpetrators: exaggerated claims for the importance of one's own work such as is found not only in the ranks of professional theorists but also among fishmongers and politicians and dentists and signwriters. The corrective, then, would not be critical realism so much as common humility.

However, the assumptions of actualism make it very easy to slip into this attitude, and indeed difficult not to. For to assume that the actual world is a closed system (which is what the expectation of spontaneously occurring constant conjunctions amounts to) is tacitly to deny the multiplicity of strata. And one will then naturally tend to think that the stratum which is one's own field of expertise is *the* explanatory stratum. If, for instance, sociology were to be a predictive science, it would have to deny that the mechanisms of biology, linguistics, geography or psychoanalysis could interfere with that course of events which would be the site of its predictions, since if they did interfere, they could deflect that course of events from its sociologically predicted outcome. For a sociologist to abandon 'sociologism' is to abandon the claim to discover Humean laws or to make historical predictions. And if the sociologist is also an actualist, he or she would see such a move as giving up all pretensions to science. And what is true for sociology is also true for each of the other disciplines mentioned — as indeed we have seen in the case of historical linguistics.

If, on the other hand, critical realism is right in claiming that since open systems are multiply determined, one can explain them in terms of many mechanisms belonging to different scientific strata, and cannot predict them at all, then sociology, linguistics and all the others can safely allow each other

autonomy and their explanatory power in human history, without giving up their own claims to be sciences.

In fact one of the most common and helpful of the interventions of critical realism in the everyday work of the human sciences is precisely this insistence that the republic of knowledge has a federal constitution. It is particularly valuable in concrete, practical and interdisciplinary studies — in areas such as human geography, sociobiology, ecology and feminist studies.

The unravelling of social and properly linguistic mechanisms is essential to Pateman's programme of defending the Chomskyan thesis that there is an innate and hence socially invariant tendency in all human beings to learn language — a tendency which also constrains and influences the structure of the language learnt. At the same time, he defends the view that *a* language — something like English or Welsh or Urdu — is a social fact, not a linguistic fact. This means, among other things, that the demarcation between two languages is neither a matter of arbitrary 'nominalist' choice, nor something inherent in the nature of the languages, in the way that the distinction between two natural kinds (such as chemical elements or zoological species) is inherent in their nature. Whether Hugh MacDiarmid's early poems are written in English or in Scots could depend on the success or failure of Scottish nationalism.

The main criticisms of Chomskyan linguistics from within Anglophone philosophy come from the followers of Wittgenstein, and Pateman devotes a chapter, 'Wittgensteinians and Chomskyans' (LMLS, pp. 120–46) to this debate. He formulates and replies to five Wittgensteinian allegations against Chomskyans:

'*Chomskyans treat something essentially social as if it were essentially individual.*' (p. 122)

'*Chomskyans treat something essentially public (outer) as if it were essentially private (inner).*' (p. 132)

'*Chomskyans treat something we ascribe on the basis of successful practice as the cause of that success.*' (p. 136)

'*Chomskyans treat something rule-like and normative as something law-like and predictive.*' (p. 140)

'*Chomskyans treat something open-ended (creative) as something closed.*' (p. 144).

The first two of these objections are based on versions of the private language argument and are the most essentially Wittgensteinian. The last three can all be called anti-causal objections, and are based on the common hermeneutic view that the positivist account of cause works for the natural sciences, but that all causal talk is out of place in the human sciences. The first of these three also rests on anti-realist, verificationist assumptions which exclude underlying mechanisms in general. Pateman suggests a natural science parallel:

> suppose someone said, 'We say someone is suffering from consumption when they grow pale, weak, thin, clear headed, cough blood and die — and that's what consumption is.' This attitude would rightly be regarded as anti-scientific; had people thought to take it seriously, we should all be dead of tuberculosis. (LMLS, pp. 137–8n)

The last two belong with the Wittgensteinian denial that reasons can be causes, to which we have already seen some objections. But they also involve the a priori rejection of anything mental but not conscious; as Pateman puts it:

> Under some descriptions [speech] is properly explained . . . in causal-mechanical terms — yet the nature of what is being explained will dictate that the causal-mechanical explanation include reference to inaccessible rules or representations — exactly the position for which Baker and Hacker [Wittgensteinian critics] reserve the most delicate weapons in their armoury of adjectives: 'nonsensical' and 'absurd'. (LMLS, p. 141)

Realist positions already argued for undermine these views, though Pateman's arguments are also partly of an empirical nature. In the case of the private language argument, on the other hand, the issue is not realism v. non-realism but the share of explanation of language-learning which falls to social mechanisms and that which falls to innate mechanisms. Indeed the data which Pateman presents against the Wittgensteinians

may also require some amendment to be made to the account of language reproduction/transformation that I outlined on p. 146 (though it does not undermine the illustrative function of that account, for reasons which will appear shortly).

The point of contention is this: according to the model of language-learning generally derived from the private language argument, the *input* to the language-learning process must be a language. And obviously enough, in order to learn an existing language like English or Welsh or Urdu one must be exposed to that language in adequate measure. But there are also cases, Pateman claims, in which a language is acquired (*produced*) without adequate linguistic input. The first case is that of creolization. When a linguistically fragmented community comes to use a pidgin for communication between people for whom it is not their mother-tongue, the coming generation will sometimes be exposed only to this pidgin — a simplified and impoverished language. However, they will not become speakers of such a truncated language, but will transform it into a much richer, syntactically more complex, flexible and creative language. (This transformation process is known as creolization, since linguists use the term 'creole' for a language created in this way, 'pidgin' for the truncated language from which it arose.) This case, of course, while it suggests some innate tendency to acquire a language in the full sense (the sense in which a pidgin is not a language and a creole is), does not involve the language being in any sense *private*; it is the product of a generation, not an individual. But the other case takes the matter further: deaf children of hearing parents who do not teach them sign-language will create their own sign-language; they do this with *no linguistic input*, though (importantly) not without *social* input; wolf-children are in quite a different case. (It is for this reason that I say that my example on p. 146 still serves its illustrative function.) These examples do not show the possibility of a private language if that is taken to mean a language that is *in principle* private. Others may learn the deaf child's private language. Hence the use of the private language argument against (for example) sense-datum theories is not impugned. Nevertheless, some Wittgensteinians do read the private language argument as ruling out as impossible a priori these cases which are a posteriori actual.

As I have said, realist philosophy as such does not give reasons for taking one side or the other in this dispute. The Chomskyan case is empirical — though the Wittgensteinian counter-case is not. The relevance of critical realism here is that the Wittgensteinians, in this case as in the anti-causal arguments, are placing a priori limits on the number of possible strata. They are insisting that a phenomenon be explained in terms of an already familiar stratum (the social), though there are empirical grounds for postulating a distinct stratum consisting of innate, rule-following but unconscious mechanisms.

Psychoanalysis: David Will

The Scottish psychiatrist and psychoanalyst Dr David Will has published several papers in professional journals using critical realism to defend the scientific status of psychoanalysis, to criticize developments within psychoanalysis which he sees as retreating in the face of fundamentally misconceived criticisms from outside, and to make some suggestions about which directions in psychoanalytic theorizing are likely to prove fruitful and which not.

In 'Psychoanalysis as a Human Science', David Will starts by noting that many claims that psychoanalysis is unscientific are inspired by a Popperian account of science. Psychoanalysis is said to be unfalsifiable, yet, if Bhaskar is right, strict falsification cannot be had except in closed systems, and these cannot be had in the human world. David Will outlines the critical realist case against the empirical realism to which Popper adheres (a case by now familiar to readers), and thus defends the *possible* scientificity of psychoanalysis; of course, it has still got to prove itself by its explanatory power:

> any adequate critiques of psychoanalytical hypotheses must demon-
> strate empirically a greater explanatory power. That is they must
> render intelligible all the significant phenomena that are rendered
> intelligible by the psychoanalytic hypothesis and in addition must
> render intelligible significant phenomena that are not so rendered by
> it. (p. 210)

However, he makes a number of points which make psycho-analysis look at least a likely candidate for a science. First of all, there are parallels between the way it proceeds and the way critical realism describes natural science as proceeding. He summarizes Bhaskar's account of this as follows:

> *Phase One.* The first empirical stage. A phenomenon is empirically identified utilizing antecedently existing cognitive materials. . . .
> *Phase Two.* A plausible explanation for this invariance is invented. Plausible generative mechanisms are imagined which could account for the phenomena in question.
> *Phase Three.* The reality of the generative mechanisms imagined in phase two is subject to empirical testing. (The natural and human sciences differ in the nature of the empirical testing. . . .) (pp. 205–6)

David Will suggests that Freud's postulation and later rejection of his 'seduction theory' instantiates this; the ready-to-hand explanation 'memory of a real event' turns out not to explain his patients' reports of their childhood seduction by their fathers, so another mechanism is postulated (infantile fantasy explained by the Oedipus Complex).[1]

Furthermore, the schema for the deepening of scientific explanation quoted on p. 49 above is paralleled by the following schema suggested by David Will (p. 208):

Stratum 1	Seductive behaviour of a female hysteric in relation to a male analyst.	
	explained by	
Stratum 2	Repetition of previous experiences (theory of transference).	*Mechanism 1*
	explained by	
Stratum 3	Projection of unconscious material.	*Mechanism 2*
	explained by	
Stratum 4	(conflicting theories of the nature of conscious)	*[Mechanism 3]*

It should be noted that this example also indicates that psycho-analysis is not a closed system of fixed beliefs and that indeed it is at this most basic level that major conflicting theories (e.g.

Freudian, Kleinian, Fairbairnian and Lacanian) do disagree most fundamentally.

Second, Will suggests that mistaken criteria of scientificity have led to the separation of research from the practice of psychoanalysis.

> All too often it has been assumed that 'research' must be a distinct field of activity and that 'scientifically meaningful' work cannot take place within the clinical setting. I have argued that such views are incorrect since they imply that the scientific study of open systems is impossible. (p. 208)

The objection to clinical practice as the site of discovery is only that it does not establish closure — but neither do the supposedly experimental practices which are therefore quite wrongly privileged as 'more scientific'. Often such strategies are blind alleys in which 'a great deal of effort is necessary to achieve very little' (p. 209). Anyone who has ploughed through such painstaking but peculiarly unfruitful work as that contained in Fisher and Greenberg's monumental volumes *The Scientific Credibility of Freud's Theories and Therapy* and their edited collection *The Scientific Evaluation of Freud's Theories and Therapy* will surely agree.

In 'The Progeny of Positivism: The Maudsley School and Anti-Psychiatry', David Will traces the effects of the positivist misunderstanding of natural science on psychiatric practice. The Maudsley School, which has been deeply influential on British psychiatry, has quite explicitly appealed to Popperian concep-tions of science, and this has given it a certain direction both in theory and in practice — a direction which therefore has a philosophical rather than a psychiatric rationale; and the philosophy is a mistaken one. Thus Sir Aubrey Lewis, a central figure of the School, states that objectivity in psychiatry is to be obtained by 'measurement, systematic observation and experi-ment, as in the natural sciences', that unless 'a generalization is stated in such terms that it can be tested and possibly falsified it may serve pragmatic ends but is hardly a scientific hypothesis', and that psychoanalysis fails this test (on which he cites Popper). He claims '"that a process of methodological cleansing

is in train", whereby scientific psychiatry will be cleansed of psychoanalytic residues'. (All quotes cited by David Will on p. 52 of this paper.) This surely shows that philosophy is not innocuous: the near-absence of psychoanalysis from English psychiatry is partly to be explained by the prevalence of mistaken philosophical views.

David Will further argues that one important reaction against the Maudsley School, the 'Anti-Psychiatry' movement, (Laing, Cooper, Ingleby and others) shares many of its assumptions; it

> has abandoned the field of natural science to positivism, and has then argued that the human sciences, like psychology, are radically distinct from the natural sciences and must be based on a distinct hermeneutic epistemological framework. (p. 50)

It also has its own versions of three aspects of the Maudsley School's outlook that Will describes as the price that has been paid for the Maudsley School's undoubted contribution in certain areas of psychiatry. This price comprises

1. 'uncertainty about the validity of applied science';
2. 'the proscription of certain areas of knowledge and investigation as forbidden'; and
3. 'the denial that mental events can have causal efficacy' (p. 53).

1. Positivism has problems accounting for applied science at all, since the rare conditions under which positivist assumptions seem to obtain (that is, in closed systems) are not the conditions in which the applied scientist is working. Will claims that this has practical effects in an 'uncomfortable dichotomy between the psychiatrist-as-scientist and the psychiatrist-as-clinician' (p. 54). If one is a positivist under the former description one will be a sceptic under the latter. Clinical research itself becomes 'forbidden knowledge'.

2. As an example of forbidden knowledge, Will refers to the conclusions drawn from studies of asthma in children, under the psychiatric aspect. The statistical correlation between asthma and 'maladjustment' was small, and comparable with that in other physical disabilities. But a psychoanalyst might well argue that 'asthma is *itself* a psychiatric disorder, a suggestion which explains why asthmatic children show no significant excess of other psychiatric symptoms' (p. 56). Yet the conclusion drawn was the Popperian one, that the question 'is asthma a psychiatric disorder?' gave rise to 'no unambiguous testable predictions'; so the question is effectively dismissed. Will comments:

> The peculiar poignancy that this example carries is related to the quite massive gap between research and practice that it implies. For clinical child psychiatrists, the forbidden question of whether or not asthma is a psychiatric disorder in its own right, is a vital one. Certainly no child psychiatrist could possibly afford to overlook this possibility in his clinical work. Yet, when Graham, Rutter *et al.* [the authors of the paper on asthma] became men of science, 'hard' researchers, they must dismiss this question as unanswerable. To paraphrase Wittgenstein, it is as if they are saying 'Whereof we cannot make falsifiable hypotheses, thereof we must remain silent'. Rather than admit that the reality of asthma is perhaps too complicated to be fully studied by positivistic research protocols, they prefer to proscribe that reality in all its complexity, rather than alter their research protocols. (p. 56)

3. For the 'strong form of Maudsley School positivism', *reasons* for actions, whether conscious or unconscious, would be just one more forbidden subject, since they are mental rather than physical or behavioural. It is concerned only with causes which are not reasons. But there is also a 'weak form' which is prepared to allow talk of reasons, as long as they are not said to be causes. Adherents of this view

> make a distinction between mechanisms and meanings, or between reasons and causes, and divide up the field of scientific inquiry accordingly. For such writers, the natural sciences, which are conceived of in a positivistic way, are applicable to the study of mechanisms and causes. Human sciences, such as psychoanalysis, are seen as being concerned with 'meanings' or 'reasons', which are held to exist in a logically separate domain from 'mechanisms' or 'causes', and the human sciences are seen as being based on a hermeneutic epistemology. (p. 58)

This creates a space for psychoanalysis, but it is not the space of an explanatory science. What sort of psychoanalysis is left? The answer is perhaps provided by the Anti-Psychiatry movement, which accepts the 'weak Maudsley' account, with an inversion of value-judgements: while the latter relegates non-causal studies to a lower, non-scientific league, the anti-psychiatrists revel in their abstention from causal explanation, as more 'humanistic'. But this makes it impossible for them to distinguish between a reason and a rationalization (for a rationalization is precisely an agent's reason for an action which is not the reason that caused the action); and this leads 'to the assumption that when the experience of a patient has been made intelligible it has been shown to be rational and not at all sick' (p. 62).

The segregation of reasons from causes also disables the anti-psychiatrists from describing the real multiplicity of causes in psychiatric matters, for some of these causes are reasons while some (for example, genetic factors in schizophrenia) are not. Once the evidence for a genetic factor is admitted to be conclusive, the anti-psychiatrist cannot say that genetic and social factors *conjointly* cause schizophrenia. So they end up consigning the whole causal explanation to the genetic reductionists, in order to preserve a social phenomenology which disowns its own explanatory power, or even postulating two types of schizophrenia, one which has causes and one which has reasons (see David Will's elucidation of 'The Progeny of Positivism' in his short paper 'Science, Psychotherapy and Anti-Psychiatry').

This denial of multi-level causality is Anti-Psychiatry's own version of 'forbidden knowledge'. Will also claims that the anti-psychiatrists share the Maudsley School's conflation of epistemology and ontology; in the Maudsley School this conflation is the familiar empiricist form of the epistemic fallacy, while for David Cooper 'analytic rationality must be replaced in the human sciences by Sartrean 'dialectical rationality'. He claims that, in such a dialectical rationality, epistemology (the act of knowing) and ontology (the existence of the object known) are one and the same' ('The Progeny of Positivism', p. 59).

I now turn to a rather different use which David Will makes of critical realism, in his paper 'Psychoanalysis and the New

Philosophy of Science'. Here he is primarily concerned not to defend psychoanalysis against external critics, but to find a way of arbitrating some of the controversies between different schools within the psychoanalytical movement, using the notion of the stratification of reality. These rival theories, he suggests, 'are not homogeneous' (p. 164). They are not all conflicting theories about the same subject-matter; some of them may be theories about different strata, which could in principle be fitted together into an integrated theory of psychoanalysis. He mentions two attempts to fit them together which are not satisfactory: the idea that classical Freudian theory is about three-person relationships while object-relations theory is about two-person relationships; and the assignment of different theories to different chronological stages of development (for example, Kleinian to pre-oedipal, classical Freudian to oedipal). But if there is a genuine distinction of strata within the subject-matter of psychoanalysis, different theories might be referred to different strata and thereby integrated into a consistent general theory.

> Lacan's distinction between the Imaginary Order and the Symbolic Order first provides a means of situating different theories: classical Freudian theory takes the Symbolic as its main frame of reference while the various object-relations theories are primarily concerned with the vicissitudes of the Imaginary Order. (p. 165)

Will goes on to discuss how certain schools of psychoanalysis, in trying to assimilate philosophical conceptions of science, have distorted the content of psychoanalysis, with reference to the respective places of reality and fantasy in mental life. On the one hand, Bowlby is overly influenced by positivism:

> Thus, as a result of a process of epistemological identification with the aggressor Bowlby progressively abandons the inner world and fantasy, in favour of observable behaviour. This abandonment of the inner world and fantasy necessitates an ultra-realist theory of psychopathology which then becomes explicable in terms of real traumata (e.g. all neurotics are victims who were really subject to an experience of abandonment). (p. 167)

On the other hand, hermeneutics errs in the opposite direction. It 'paves the way for solipsism, for the return of the transcendental subject and for a universal theory of symbolism' (p. 167). For it treats a person as 'ultimate guarantor of the meaning of his own reality' (p. 168), an infallible authority on the truth about himself. This eliminates the psychoanalytic notion of resistance, and the distinction between rational and irrational, sanity and madness. It presupposes just that Cartesian view of the subject whose death-knell Freud's discovery of the unconscious had struck. Furthermore, it makes it difficult to apply general concepts like the Oedipus Complex or the (Kleinian) paranoid-schizoid and depressive positions, which are clearly explanatory concepts going beyond and often contradicting what the subject will avow. If it generalizes at all, hermeneutics can only do so in terms of a theory of universal symbols, as did Jung (presumably in his theory of archetypes and the collective unconscious) and Stekel (with his general theory of dream-symbolism).

The pressure of positivism is to attribute all the individual's vicissitudes to the environment, while hermeneutics (like biological reductionism) attributes everything to the individual's inner endowment. Psychoanalysis must avoid both these pitfalls and propound an epigenetic theory of development, that is, one which recognizes the interaction of a real environment with a really existing inner world of fantasy, including misperceptions of the environment. Here, as in the case of socio-linguistics, critical realism's notion of stratification and multiple determination serves as a corrective to one-sided explanations.

Economics: Tony Lawson

In a number of papers, the Cambridge economist Tony Lawson has used critical realism to expose the weakness of orthodox ('neo-classical' or 'marginalist') economics. As in the case of Will's critique of psychiatry, the argument is based partly on the fact that orthodox economists are, either consciously or unconsciously, founding their conception of what economics ought to be able to achieve on a misconception of the practice of

the natural sciences; in particular, a Humean or positivist misconception. But whereas psychiatry at least achieves some success in its own terms — it cures or relieves the symptoms of some patients — economics manifestly fails in the task it sets itself: it aims at prediction, and fails.

Lawson's starting point is that orthodox economists assume a conception of causality as constant conjunction. Often, this assumption is taken for granted and not recognized as a legacy from a questionable philosophical account of natural science. Thus in a short discussion titled 'Methodology: Non-optional and Consequential', he criticizes the view of Frank Hahn that economics can dispense with methodology (here meaning concern with the philosophical foundations of its claim to be scientific), and just get on with the job. The rejection of methodology is really a refusal to examine an implicit positivistic methodology; this positivism is not a simple commitment to unprejudiced scientific work, it has assumptions both about the nature of the object studied and about the nature of the human agents of the study: they are 'passive sensors of atomistic events and recorders of their constant conjunctions' (p. 1). But if we are not passive sensors, and if economies are open systems in which constant conjunctions do not occur, such a methodology will get us nowhere. Refusal to own up to having a methodology means refusal to re-examine one's assumptions.

> Positivism, however, is not mandatory. Indeed, it has long ago, and justifiably, been rejected in the discipline of philosophy, and it really only lives on in unreflective subjects such as economics. (p. 3)

Hence Lawson's critique of orthodox economics, while it takes its rise from the internal anomalies of economics — from the fact often repeated by Lawson that 'in the field of economics significant invariant empirical regularities are yet to be observed', appeals to critical realism in a more direct way than we have seen in the cases of linguistics and psychoanalysis. In the latter cases, it was a matter of defending certain work from philosophical critics; economics, on the other hand, will get nowhere if it does not change the sort of question it asks. So long as its philosophical assumptions are Humean, it will go on looking for such regularities, and failing to find them. They are

not there to be found, since the systems studied are open, both externally and internally.

Closure — tacitly assumed to occur spontaneously by all those who seek invariant empirical regularities — would require both an extrinsic and an intrinsic condition to be met (see 'Realism, Closed Systems and Expectations' pp. 7–8). The extrinsic condition is met either by isolation of a system from outside influences, or the constancy or known and calculable nature of any such influences. Something approximating to this might occasionally occur — Lawson does not rule out such local and temporary closures. But they are certainly the exception. The intrinsic condition of closure is constancy of input from the individuals which make up the system (economic individuals: not only people but 'households, consumers, firms, banks, trade unions or whatever', 'Economics and Expectations', p. 12). But these are complex individuals, and have the power to respond differently to the same situation on different occasions; likewise different individuals of the same type may respond differently on the same occasion. Economists are likely to respond to the absence of these two conditions of closure, as Bhaskar has suggested, by a double regress: to ever larger systems to eliminate external effects; and to ever smaller atoms. But if the world is an endless complex of complexes, as transcendental realism seems to suggest, neither regress is terminable. (Or if there are termini, they are certainly not within the economic field.)

Orthodox economics begins to look like 'hunting for a black cat in a dark room where there is no cat', as Bergson said of metaphysics; hunting, that is, for constant conjunctions in an open system where there are no constant conjunctions. Does this mean that orthodox economics is a total non-starter? Why then do hard-headed businesspeople pay for the advice of economists? Are not such anomalies recognized within orthodox economics? And would an alternative, non-predictive economics be any use in guiding decisions?

So far as the expertise of economists is concerned, Tony Lawson does not deny that it exists, only that it has anything to do with their theories: it is a function of the care with which they study the economic news (oral communication). As to internal critiques of orthodox economics, they form a considerable part

of the subject-matter of Lawson's papers. But in order to understand the force of such critiques, he argues, they must be retrieved into a critical realist context, and then they will be far more far-reaching than they would otherwise seem.

An economic theory which took critical realism seriously would be an explanatory but non-predictive study of relatively enduring structures and institutions, their powers and tendencies. Would such a theory be useful to policy-makers, in the way that orthodox economics claims to be? One might of course justifiably say: with successes like those of (for example) monetarism, who's afraid of failures? But the question is not just whether a given theory gives good or bad advice, but whether a non-predictive theory is any practical use at all. This is a question that has recently been raised about critical realism as a whole by Wal Suchting, in his article 'Reflections upon Roy Bhaskar's "Critical Realism"':

> What is to be said about a doctrine that on the one hand claims to be devoted to 'projects of human emancipation' . . . and on the other denies that social theory can be predictive, that is, aimed at the future, rather than explanatory, that is, aimed at [what] has been or is the case? Is there anyone who needs to have spelled out the premise or two that would permit the deduction of a formal contradiction here? (p. 30)

Aside from the ambiguities of 'aimed at' in this context, the idea that the difference between explanation and prediction is just one of tense is a positivist chestnut that one would not expect Suchting to pull out of the fire. The transcendental realist account of explanation — tendencies of things explained by their structures — is untensed: if things with a given structure exist in the future, they will have the same tendencies as they did in the past. If that is prediction, so be it! The critical realist slogan 'not prediction, explanation' means, on the one hand, that there are no reliable predictions of constant conjunctions, and, on the other, that statements about past conjunctions are not in themselves explanations — more is needed. However, one can imagine how someone might think non-predictive theory would have no practical implications, so it is interesting to look at a counter-example shown by Lawson to be present in Keynes.

According to Keynes, long-term economic prospects are radically *uncertain*. In one sense of the word, most economists would agree, and say that economic predictions had degrees of probability rather than certainty. But Keynes — no novice when it comes to probability theory — did not mean that:

By 'uncertain knowledge' ... I do not mean merely to distinguish what is known for certain from what is only probable. The game of roulette is not subject, in this sense, to uncertainty. ... Even the weather is only moderately uncertain. The sense in which I am using the term is that in which the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence, or the obsolescence of a new invention or the position of private wealth owners in the social system in 1970 [Written in 1937 — A.C.]. About these matters there is no scientific basis on which to form any calculable probability whatever. We simply do not know.' (*Collected Works* XIV, p. 114, quoted by Lawson in 'Economics and Expectations', p. 19).

Lawson points out that this is not just a throw-away remark, but a pervasive theme of Keynes's thinking throughout his life. And it means that a good deal of the 'economic rationality' that is supposed to explain people's actions is just whistling in the dark. Investors make decisions on the basis of short-term expectations, which are themselves largely a function of 'mass psychology'. A small minority of clever investors make money by speculating on short-term predictions, arrived at through knowledge not of the economy but of the mass psychology of other investors.

But economic institutions constrain or facilitate the activities of investors, and different institutions do so in different ways. 'Liquidity', i.e. relative ease in buying and selling shares, converting investments into money and vice versa, tends to separate investment from enterprise and deliver it over to speculation.

Speculators may do no harm as bubbles on a steady stream of enterprise. But the position is serious when enterprise becomes a bubble on a whirlpool of speculation. When the capital development of a country becomes a by-product of the activities of a casino, the

job is likely to be ill-done. (*Collected Works* VII, p. 159, quoted by Lawson, ibid., p. 31).

Some policy recommendations in the public interest follow from this account: measures to make the investment market less liquid, such as a transfer tax on transactions, or more radically, making 'the purchase of an investment permanent and indissoluble, like marriage' (*Collected Works* VII, p. 166, quoted by Lawson, ibid., p. 31).

Here we have proposals generated by a theory, not because that theory forewarns us against a long-term future that the proposals can then forearm us against, but because it reveals the tendency of a particular economic substructure (a liquid investment market) to generate a particular anti-social outcome (the dominance of speculation over enterprise), and so points out the public interest in altering that structure. While of course such a proposal is wholly within the capitalist framework, it illustrates how a non-predictive structural realism can guide policy.

In two of Tony Lawson's papers he discusses instrumentalism as an influential non-realist methodology in economics. In 'Realism and Instrumentalism in the Development of Econometrics' he contrasts the two positions in terms of Bhaskar's account of the use of models in science: according to Bhaskar, some regularity is first identified, a model is postulated which would explain it, and the model is then tested with a view to seeing if it matches some real structure. Instrumentalism leaves out the last stage: so long as the initial data are 'as if' the model had generated them, the model is satisfactory. Indeed, 'for the instrumentalist there is no necessary requirement that the model even be plausible', and in his paper 'Realism, Closed Systems and Friedman' Lawson discusses some manifestly implausible models proposed by Friedman as examples:

i) Under a wide range of circumstances, bodies that fall in the actual atmosphere behave *as if* they were falling in a vacuum.
ii) [Leaves on a tree] are positioned as if each leaf deliberately sought to maximize the amount of sunlight it receives, given the position of its neighbours, as if it knew the physical laws determining the amount of sunlight that would be received in

various positions and could move rapidly or instantaneously from any one position to any other desired and unoccupied position. (Friedman, *Essays in Positive Economics*, pp. 18–19, quoted in 'Realism, Closed Systems and Friedman', p. 9)

The striking thing about the second example — though Lawson does not spell this out — is that had such an account been regarded as acceptable in botany, genuine scientific explanation could not have got off the ground in that science, since that explanation aims precisely to show how such an appearance could be generated by a plausible hypothesis which did not need to be quarantined in an 'as if' clause; and a methodology which would have prevented a real science from emerging can hardly be regarded as a promising one for an aspiring science. Lawson concentrates on showing that the former example allows no principle for delimiting the 'wide range of circumstances'; granted that a leaf falling from a tree does not normally behave in this way, but a compact ball falling from a roof may — though not in a hurricane, and so on. Whereas if the law of falling bodies is read not as an exception-prone generalization but as a tendency-statement, it can take its place alongside other tendency-statements in explaining falling leaves and balls in hurricanes.

To return to 'Realism and Instrumentalism in the Development of Econometrics', Lawson claims that econometricians, in their search for measurable statistical regularities, are forced to take refuge in instrumentalism: whereas the realist 'may just rest content with the ability to explain and predict the tendencies of identified causal structures' (p. 243), econometrics, if it is to propound measurable probabilities which will not be submerged in Keynesian uncertainties, must presuppose closure — the absence of extrinsic factors. But this is unobtainable:

the econometrician will usually have positive knowledge of numerous ... potentially relevant causal factors that it is not possible to explicitly consider. In such situations, I want to suggest, econometricians who have acknowledged this problem ... appear to have been unable, or unprepared, to develop any option other than that of introducing, or recommending others to introduce, some

acknowledged convenient fiction ... in the hope that the model so constructed turns out to be data consistent ... it is at this point that econometric analyses appear to become, in effect, instrumentalist. (p. 243)

This is contrasted with Keynes's realist understanding of probability, which of course circumscribes the conditions of its applicability.

Keynes's main conclusion is that the use of statistical induction is only justified when there are grounds available for supposing that the nature and conditions of things under consideration are of a particular type — specifically when their conditions can be likened to a game of chance. (p. 246)

The conclusion is highly sceptical about the prospects of econometrics. It should be added this is not in a situation

in which the status of some empirically successful model is being disputed. Rather it is one in which successful empirical generalizations have yet to be discovered in economics, while there exist good theoretical reasons to suppose that simple empirical relationships may not be there to be found. (pp. 252–3)

We have seen in other social-scientific disciplines that positivism (of which instrumentalism can be considered one avatar) is often challenged by a hermeneutic rival, which often at least tacitly shares its account of the natural sciences. Economics is no exception.

Tony Lawson discusses the hermeneutic method in economics with reference to the Austrian economist Hayek (in 'Critical Realism and the Analysis of Choice, Explanation and Change', and in 'Realism and Hayek: A Case of Continuous Transformation'). Hayek's position in most of his writings is that social sciences are not about objective reality but about our conceptions. Insofar as these conceptions are *about* the objective world they may be true or false, but the social scientist need not ask about their truth, for false ideas will be just as effective in governing behaviour. The social scientist then studies a pre-interpreted world and his or her subject-matter is that interpretation — the set of concepts by which ordinary, non-

scientific agents theorize their lives. The economist studies the self-understanding of economic agents. This goes beyond the claim — endorsed by critical naturalism — that social realities are concept-inclusive and in part concept-dependent. For Hayek, social realities are *constituted* by concepts. Of course, material entities may be social, but only by virtue of the concepts we have of them: it is our ideas about hammers that make them hammers, as distinct from pieces of wood with metal ends.

Lawson raises three objections to this account:

1. The social realm is not exhausted by its conceptual aspects.

   > To an individual agent, being unemployed, fighting a war, living in poverty, etc. is not just (and sometimes perhaps not at all) possessing a particular idea of what one is doing: it involves being physically separated from the means of 'earning a living'; being party to armed conflict, and being separated from (adequate) forms of health-care, shelter and nutrition, etc., with all the material problems which that involves. ('Critical Realism and the Analysis of Choice, Explanation and Change', p. 20)

2. The relevant conceptual aspects of society are not necessarily conscious, and so not always accessible to the social scientist; they include tacit conceptual skills and tacitly followed rules (as studied by linguists), unrecognized beliefs and needs, unconscious motives and attitudes, and so on.
3. Agents' conceptions may be false — not just false beliefs about the physical world, which Hayek's view explicitly accommodates, but false beliefs about their own actions, the social relations in which they stand, and so on.

All these points indicate that a social scientific inquiry cannot limit its subject-matter to concepts, and must be able to unearth hidden, and correct mistaken, ideas.

Lawson further argues that Hayek's errors stem from a mistaken ontology, which he takes over in part from his positivistic opponents. The 'external world' is seen as 'the empirical world': the causal criterion for existence is ignored and unperceivable entities treated as mere constructs. Apart from this flattened natural world, the only realities admitted are concepts: social structures, wholes, etc. are treated as having

reality only in our minds. Once this ontology is accepted, it is difficult to avoid Hayek's methodology. But as we have seen, a far richer ontology of stratified nature and structured wholes is well-founded.

Finally, in 'Realism and Hayek: A Case of Continuous Transformation', Lawson traces the development of Hayek's thought, claiming that it eventually moves away from this ontology, particularly by its foregrounding of the place of tacit rules in social life and its acceptance of a sort of hypothesis-testing which seems compatible with Roy Bhaskar's three-stage account. This opens the way to recognizing an intransitive dimension to social science, and the possibility of counter-phenomenality. While Lawson describes this as a move towards transcendental realism, he recognizes that Hayek's conservative political judgements, which he does not endorse, remain intact. This illustrates the logical independence of transcendental realist philosophy and the socialist politics with which it is so naturally combined, as stated in chapter 6 of the present book.

Some Other Critical Naturalist Interventions

While the standing of all human sciences is contentious, the last three disciplines considered, linguistics, psychoanalysis and economics, at least have the appearance of sciences under their transitive aspect: there are organized bodies of trained researchers and teachers working in them with some shared principles and approaches and the intention (even if tacit or denied at the meta-level, as in the case of instrumentalists) of deepening knowledge of some fairly clearly defined intransitive object (signifying practices, unconscious processes and the production and exchange of commodities, respectively). But there are also aspects of human life so politically or ethically contentious that there is no science-like discipline in the study of them; or about which several human and perhaps also natural sciences contend. Such areas are often a happy hunting ground for philosophers in much the way that areas now regarded as within the competence of some science once were. It is all the more important that such philosophical interventions should avoid the characteristic mistakes of empiricism or idealism

which have damaged even more established research programmes.

Perhaps I may be allowed the liberty of introducing first my own contribution in this area. I refer to my argument in *Scientific Realism and Socialist Thought* and *Socialist Reasoning*. The former of these two books is about the relation of critical realism to Althusser's version of Marxism, and more specifically the way in which dilemmas unresolvable within Althusser's system could be resolved by critical realism: how to avoid empiricism without severing theory from practice; how to avoid economism while retaining the primacy of base over superstructure (briefly, by reading the levels as rooted and emergent strata in the transcendental realist sense, between which one-way relations of vertical explanation hold, while mechanisms at all strata conjointly generate the phenomena of history, in no invariant proportions); how to reconcile structural causality with effective human agency; and how to give a philosophical account of scientificity without making philosophy into a legislator for the sciences. On the basis of the critical realist solutions to these questions, I suggest that Marxian social science is about *constraints on the reproduction and transformation of social structures*. The knowledge of these constraints is the ground for political judgements: constraints on the reproduction of a society show how it cannot reproduce itself without developing certain destructive and even self-destructive features; constraints on transformation show which putative solutions of these problems are blind alleys.

In *Socialist Reasoning* I try to show how a socialist political philosophy — in the sense of a theory of the reasons for socialism — can be based on knowledge of such constraints (prominent among which are those constraints which Marx and Engels dubbed 'contradictions of capitalism'). The strength of a style of political advocacy which starts from contradictions in some existing social order is that it can avoid, on the one hand, the errors of utopianism or abstract ideals which ignore the historical specificity of human needs; and, on the other, the entanglement with particular traditions by which historicist and Hegelian approaches are mortgaged to conservatism or at best to moderate reformism. It is interesting that Ian Shapiro, also writing political philosophy in a critical naturalist context in his

Political Criticism, should be concerned to avoid a very similar Scylla and Charybdis: the communitarian successors of Hegel, and the successors of the unhistorical foundationalism of the Enlightenment, Rawls and Nozick. It is interesting because Shapiro's political stance, while left rather than right, is certainly not in the Marxist tradition.

Four issues on which overlapping work has been done from a critical realist perspective are feminism, human needs, ecology, and the relevance to the human world of its closest natural science: biology. Contentious issues in these areas on which critical realism might throw some light include: what sort of 'vertical explanation' holds between biological mechanisms and human needs, social possibilities, differentiation of sex roles and so on? Do biological mechanisms generate identifiable tendencies which operate alongside social and psychological ones; or do they codetermine all human activities and institutions equally with and perhaps inextricably from social and psychological mechanisms; or should biological mechanisms rather be seen as vertically explaining social and psychological ones, but so to speak disappearing behind them, so that they make no contribution of their own to horizontal explanation? Does the 'nature' that preceded society and human agency live on alongside them, or through them, or both? Can we distinguish the 'natural' and 'unnatural' within human society? Can the concept of human nature help us distinguish real needs from false needs, or must we accept the liberal view that all wants are equally good 'needs' (so long as they are backed by purchasing power)? What can we mean by 'nature' when we treat it as a value? Given that nature as we know it has been transformed by us, in what senses is it independent of us?

While some of these questions may look abstract and technical, a great many questions in politics and social criticism turn on them, as do questions about the relations between different disciplines in the human sciences. I believe that they can only be resolved in a critical realist framework, and I refer the reader to the work of Kate Soper, Maureen Ramsay and Ted Benton (see Bibliography for relevant sources). A critical realist intervention in an even more concrete aspect of social research in this area is Sue Clegg's paper 'Studying Child Sexual Abuse'.

It will be clear from these texts, as well as those I have discussed in this chapter, that critical realism is an ongoing research programme within the human sciences, and particularly in their theoretically and politically contentious border areas. It is certainly not a completed system which can simply be applied in these fields, to solve all problems: on the contrary, by treating scientific projects as explorations of realities with inexhaustible depths, it helps to keep those projects open for self-criticism and development.

Notes

1. A word is perhaps needed to rebut the modern fable that Freud rejected the Seduction Theory for no better reason than that he couldn't bring himself to believe that respectable bourgeois fathers would commit incest. He *did* of course believe for a while that all hysterics has been victims of seduction, and even when the evidence forced him to revise this belief (a difficult enough confession of error), he never doubted that in some cases real incestuous advances had occurred, as in the case of Katharina (see Freud and Breuer, *Studies in Hysteria*). He also knew of 'respectable' men who interfered with young girls (including one prospective patient), not to speak of the childhood 'seduction' of the Wolf-man by his sister (though I would have thought that 'seduction' was hardly the right word for genital play between consenting infants). In general, he had a fairly balanced view of the interaction and joint effects of reality and fantasy: an epigenetic view of development, as David Will puts it.

===========8===========

Why Critical? How Naturalist?

So far as possible hitherto I have kept to exposition and defence of Roy Bhaskar's ideas. In this chapter I shall look at some criticisms which have been levelled at his work, and also make a few of my own; and I shall air what I see as some outstanding problems for critical naturalism.

. Critical naturalism enters the philosophical battleground fighting on two fronts: against the reductive naturalism that is rooted in empirical realism; and against anti-naturalistic positions rooted in idealism. It is not surprising that adherents of each of these tendencies assimilate critical naturalism to the other one. I have heard hermeneutic philosophers call it 'just another positivistic scientism'; on the other hand, not only positivists but even Ted Benton — a philosopher whose position is arguably within critical realism — wonders whether the views defended in PN are really naturalist at all (in 'Realism and Social Science'). The phrase 'critical naturalism' contrasts implicitly with uncritical naturalism: uncritical presumably in the senses of naïve, dogmatic, extreme; naturalism that is positivistic in method and reductive in ontology; or alternatively naturalism which is, as has been said of Lenin's *Materialism and Empirio-Criticism*, situated in pre-Kantian theoretical space.

I hope that I have already shown that whatever critical naturalism is, it is not such uncritical naturalism. While rejecting Kant and Hegel's idealism, it has taken on board their critiques of empiricism; it sees knowledge not as impressed by nature on the wax tablets of our minds but as a product of skilled mental labour; it treats the various strata of nature as irreducibly emergent.

Yet I have the impression that Bhaskar is much more worried about being confused with uncritical naturalism than about the opposite error, and so rather too willing to make concessions to

hermeneutic and relativist positions. If this is on the principle that you straighten a bent stick by bending it in the other direction, I think Bhaskar has misjudged the philosophical state of play. For instance, the idea that science (or any particular science) either has reached, or someday will reach, complete truth about (the relevant aspect of) nature — that scientific progress has a terminal destination — is an idea that has hardly any serious twentieth-century adherents. No doubt it was once held by outside admirers of science.

> Nature and Nature's laws lay hid in Night.
> God said, 'Let *Newton be!*' and All was *Light*.
>   (Pope, 'Epitaph. Intended for Sir Isaac Newton,
>   In Westminster Abbey' in *Poems*, p. 808)

But to attack this view today is to tilt at windmills. The idea of such a 'final truth' is now found only as a caricature, used in anti-science rhetoric. For example, the Marxist conception of 'scientific socialism' or of historical materialism as a science is often portrayed as a ploy to claim infallibility, even though Marxism has always been explicitly fallibilist in its account of science and hence of itself:

> science . . . mounts from lower to ever higher levels of knowledge without ever reaching, by discovering so-called absolute truth, a point at which it can proceed no further, where it would have nothing more to do than to fold its hands and gaze with wonder at the absolute truth to which it had attained. And what holds good for the realm of philosophical knowledge holds good also for that of every other kind of knowledge and also for practical action. Just as knowledge is unable to reach a complete conclusion in a perfect, ideal condition of humanity, so is history unable to do so; a perfect society, a perfect 'state', are things which can only exist in imagination. (Engels 'Ludwig Feuerbach', in Marx and Engels *Selected Works in One Volume*, p. 598)

On the other hand, extreme forms of relativism, denying that one theory can be rationally preferred to another on the grounds of its relation to reality, are so widespread that some such view is often simply assumed as axiomatic in arts and social science circles. This surely is the stick that needs to be straightened.

In passing, the following should be noted: as we have seen, Roy Bhaskar ejects epistemology from the central place it has

had in philosophy from Descartes onwards, but does not reject it altogether, as is fashionable among some mock-radicals. I say 'mock' since all such rejections of epistemology leave a theoretical vacuum which is inevitably filled by some new epistemology, unrecognized as such.

However, Bhaskar has not written about epistemology in general, but only about a *regional* epistemology — the epistemology of the sciences. This is an important region, but not the only one. The epistemology of everyday pre-scientific knowledge cannot be read off from it. Since the most serious form of relativism — the sort which stems from Kuhn's work — is also concerned with the sciences, the whole debate over incommensurability takes place on the ground of *scientific* epistemology. A full confrontation between realism and relativism awaits the clarification of the grounds for everyday knowledge as well. And that clarification must, I would argue, start from an examination of *hearsay* and the grounds on which we assess it. If I say 'I know that the artesian well on Southampton Common is 1,317 feet deep', that is not because I have climbed down it with a tape measure, or made any other kind of experiment, but because a plaque placed there by the City Council tells me so, and I have no reason to suspect them of lying or ignorance about the matter. It will be clear that the problem of relativism would be posed in quite a different way for an epistemology oriented towards hearsay and our grounds for accepting or rejecting it than it is for an epistemology oriented towards the experimental testing of explanatory models.

Now to return to the question whether Bhaskar concedes too much to relativism: if this were only a question of emphasis it would not matter very much, but I think it leads Roy Bhaskar to throw out one idea which a consistent realism ought to retain: the correspondence theory of truth. It is important first of all to make clear what this theory is and what it is not. It gives a definition of truth, *not* a criterion of truth. Kant was right to accept correspondence ('agreement of knowledge with its object') as defining truth yet say that the search for a general *criterion* of truth was like 'one man milking a he-goat and the other holding a sieve underneath' (*Critique of Pure Reason*, p. 97).

The two main alternatives to correspondence theories — coherence theories and pragmatic theories — gain their

plausibility from the importance of both coherence and practice as *criteria* of truth. If these theories are used to *define* truth, we get only circular definitions, since coherence theories presuppose that we understand the concepts of consistency and contradiction, which can only be defined in terms of truth; and pragmatic theories tacitly work with the notion that 'everything is as if' the pragmatically verified statement were true (in the correspondence sense).

The second thing that the correspondence theory is not is a resemblance theory. Several philosophers, including both Frege and Bhaskar, reject the correspondence theory on grounds that only hold good if that theory is taken to involve some notion of the true statement resembling what it is about. And of course a statement (or sentence or proposition or judgement or whatever it is that can be true or false) is not very much like anything else except another statement. The sentence 'all cows eat grass' is much more like the sentence 'all cows eat glass' than it is like the contented beasts munching away out there in the meadow. But correspondence does not necessarily involve resemblance. Everyone understands that if the inspector says 'your inventory did not correspond to what was really in the warehouse', she is not complaining that a sheet of paper did not resemble a stack of tinned fruit. 'Correspond' here is specially chosen to pick out the relation that holds when *as it is said, so it is*. This may look a bit thin — not a thousand miles from the so-called redundancy theory (that sentence 'P' is true if and only if P, for example 'the Minister of Education has fallen down the artesian well' is true if and only if the Minister of Education has fallen down the artesian well: and that that is all there is to truth). But it may be filled out, in scientific contexts, by some such account as Bhaskar's, of the deepening of knowledge of the intransitive object as explanatory models are tested; or in non-scientific, commonsense contexts, by some such account as Heidegger's in chapter 6 of *Being and Time*. These should be seen as amplifications of what correspondence means in different contexts, not alternative theories of truth.

Now let us tie this in to a specific text of Bhaskar's.

> Epistemological relativism . . . is the handmaid of ontological realism and must be accepted. Now this does not mean that it is impossible

> to communicate between different theoretical or conceptual schemes or that a scientist cannot know the same object under two or more descriptions. To show the difference between say Newtonian and Einsteinian dynamics and that the latter is in advance of the former a scientist must be capable of doing so. Similarly though there is no guarantee of successful communication between the adherents of two different conceptual schema, there is no inevitability about failure. (It is difficult to understand the concept of total failure.) Epistemological relativism insists only upon the impossibility of knowing objects except under particular descriptions. And it entails the rejection of any correspondence theory of truth. A proposition is true if and only if the state of affairs that it expresses (describes) is real. But propositions cannot be compared with states of affairs; their relationship cannot be described as one of correspondence. Philosophers have wanted a theory of truth to provide a criterion or stamp of knowledge. But no such stamp is possible. For the judgement of the truth of a proposition is necessarily intrinsic to the science concerned. There is no way in which we can look at the world and then at a sentence and ask whether they fit. There is just the expression (of the world) in speech (or thought). (RTS, p. 249)

Several points can be made about this in the light of what has already been said. The first sentence could be accepted by many fully fledged relativists — by adherents of the 'strong programme' in the sociology of knowledge, for instance. The remarks that follow that sentence show Bhaskar's difference from that position, in terms filled out by his critique of incommensurability, which I have discussed in chapter 3. But none of this is shown to 'entail' rejection of the correspondence theory. Two quite distinct points against the correspondence theory are then introduced as if they were somehow connected: the Kantian point against correspondence as a *criterion* of truth, which I have already accepted; and the confusion of correspondence with resemblance, implicit in the statement that *comparison* of a proposition with a state of affairs is impossible. In one sense of course that is so: we do not compare them as we compare two pictures in a game of 'spot the difference'. But we do 'look at the world and then at a sentence' (or vice versa) 'and ask whether they fit'. We say things like: 'this place is just as (or not at all how) it was described'; we look to see if the bottle opener is really in the knife drawer as we were told, and so on.

It is for just such relations of comparison without resemblance that we need the word 'correspondence'.

Does this tendency to 'bend the stick' towards the critics of realism and naturalism affect Roy Bhaskar's view of social being and social knowledge in any way? I think that it does on the one hand lead him to assert differences between natural and social being that are unfounded, and as a result of this (but one might nevertheless say 'on the other hand') to exaggerate the degree to which the social and more generally the human sciences can achieve a genuinely scientific status, comparable with the natural sciences. I shall defend a view of which is more 'naturalistic' at the level of being, denying some of the contrasts Bhaskar makes, but less 'naturalistic' at the level of knowledge, in that it is more pessimistic about the prospects of the human sciences. This critique is an immanent one, in that all its premises are within critical realism. I shall conclude with some suggestions about what studies of the human world ('human sciences' if you like) ought to look like, if these critical realist premises are true. In doing so I do not intend to go back on my claim in the previous chapter that critical realism does not generate or authorize particular theories within the human sciences. I shall be concerned rather with the relations between abstract and concrete in studies of the human world.

The Ontological Divide

In this section I am going to argue that some of the distinctions that Roy Bhaskar draws between the natural world and the social or human world are unreal. I hope that it should by now be obvious that this argument is not motivated by any reductive programme. I am convinced by Bhaskar's arguments about stratification and emergence, and regard it as impossible to reduce social to natural, or indeed social to psychological or psychological to social, or either to biological, or biological to physical, and so on. There is not one Great Divide here, but many divisions between mutually irreducible strata.

There are indeed special differences at the dividing line between the natural and human worlds. All the strata on the 'human' side of it are marked by the presence of those

phenomena and mechanisms which the characterization of humans as 'rational animals' is meant to denote: language and meaning generally, intentions, self-consciousness, consciousness of universals, moral judgements, and so on. But while the human/natural divide is unique in these respects, it is not unique in this uniqueness: the divide between living and inorganic matter has probably as many and as important features unique to it.

But Bhaskar is particularly concerned with certain differences which have consequences for the nature of the sciences on the respective sides of this divide. He tells us that

> it will be shown that *ontological*, *epistemological* and *relational* considerations all place limits on the possibility of naturalism (or rather, qualify the form it must take); and that these considerations all carry methodological import. However, it will transpire that it is not in spite of, but rather just *in virtue of*, these differences that social science is possible; that here, as elsewhere, it is the nature of the object that determines the form of its possible science. (PN, p. 3)

If the 'as elsewhere' in this passage were always kept in mind, there would be no objection. For in the case of *every* stratum studied by a science, it is the real differences from other strata characterizing that stratum which make it the possible object of a separate science. There could be no science of biology if life-forms were not really governed by mechanisms not found in inorganic matter, and so on. But this passage easily suggests special differences between the natural and social worlds which give the social sciences some advantages, which might perhaps compensate for the absence of experiments in them. And I think that in fact, underlying Bhaskar's (qualified) optimism about the prospects of the social and human sciences generally, there is an echo of the thought that society is (potentially) transparent to us since 'we made it'. Let us now consider the ontological limits to naturalism, which I have already quoted without comment:

1. Social structures, unlike natural structures, do not exist independently of the activities they govern.

2. Social structures, unlike natural structures, do not exist independently of the agents' conceptions of what they are doing in their activity.
3. Social structures, unlike natural structures, may be only relatively enduring (so that the tendencies they ground may not be universal in the sense of space-time invariant). (PN, p. 38)

Let us take this third point first. Social structures are certainly only relatively enduring; the laws governing capitalist economies did not operate in the high Middle Ages or earlier, and I hope there will come a time when they will cease to operate. They did not operate because they are the tendencies of a certain kind of structured entity (a capitalist economy), and such entities did not exist at that time. If such entities cease to exist, these tendencies will cease to operate. In *that* sense, they are not space-time invariant. But in another sense they are: that *whenever* economies with the relevant structure exist, these tendencies operate. So these laws can be formulated in terms which are universal, by virtue of being *conditional*: 'if the ownership of productive wealth is separated from the direct producers and divided between competing sellers of the products, then tendencies x, y, z will operate.'

But the universality and space-time invariance of *natural* laws is of exactly the same form. If life becomes extinct, biological tendencies will cease to operate; if our universe bumps into a universe of anti-matter and returns into nothingness, the law of inertia will cease to operate. But like the laws of capitalist economies, they will still be true in their conditional form: if bodies exist, they will tend to persist in a state of rest or uniform motion in a straight line, and so on. Of course, there is a vast difference in time scale between the 'relative endurance' of social structures and of natural ones. 'Nature is ever green, or rather goes/ by such long paths/ that she seems still' says Leopardi (quoted by Timpanaro in *On Materialism*, p. 43). But the natural sciences, even the physical sciences, even cosmology, do recognize the emergence of newly operating tendencies as the structure of the cosmos changes. For example 'big bang' theories about the origin of our universe postulate quite different laws operating immediately after the big bang than those that have operated in subsequent ages. For everyday

purposes the contrast between unchanging nature and changing society works well enough. But it is not a difference of principle, only of degree. And this is worth pointing out, because it helps to undermine point (1) as well.

What can it mean to say that natural structures exist independently of the activities that they govern? Do molecules exist independently of the activities of their component atoms, or living organisms independently of the activities of their organs, or the solar system independently of the movements and gravitational pulls of the planets? The statement is only plausible if 'structure' is read as meaning 'type of structure' rather than 'structured entity'. Then the statement would refer to the time-space invariant 'structures' which may or may not be instantiated, but about which conditional statements may be true. But as we have seen, there is no difference between natural and social structures in this respect.

If we leave out these questions about timeless essences and look at actually existing structured entities, another reading of this supposed ontological difference may be suggested. It could be a denial that unactualized powers exist at the social level. If so, one might say that a motor bike had the power to go at 100 m.p.h. though it never had, but not that the proletariat had the power to emancipate itself though it never had. Ted Benton has pointed out ('Realism and Social Science') that this difference too is unreal:

> An organism may, for example, never engage in reproductive activity, yet retain its reproductive system and powers. However, *some* activities of the organism (such as nutrition) would be necessary to the retention of these powers, but not the ones directly governed by the reproductive system itself. (p. 17)

And the case is just the same in the social sciences: powers are dependent on some activities, yes; but not necessarily those activities which are the exercise of those powers. The powers of the state, for example, could not exist without some activities of its agents; but the state has powers which it never exercises, such as to suppress a threat to its authority; it may not have to exercise this power precisely because everybody knows that it

has this power. One might even say paradoxically that the power is realized though unexercised. Benton refers in this connection to Steven Lukes's *Power: A Radical View*, which shows from empirical studies how the power of large capitalist corporations can work in this way, preventing protests about the environmental damage inflicted by a firm which is the major local employer, for instance.

In the Postscript to the second edition of PN, Bhaskar replies to some of Ted Benton's criticisms. As we would expect, given the general nature of the distinction between powers and their exercise in Bhaskar's thought, he accepts the point about unexercised social powers. His sticking point is that 'nothing happens in society save in or in virtue of something human beings do or have done' (PN, p. 174). But if we substitute the relevant kind of entity in each case for 'human beings', exactly similar principles would hold in the natural sciences.

Point (2) remains. We need to be clear what is being said here, and what depends on it. It is certainly true that social activities, and therefore (by (1), properly construed) social structures, necessarily involve the agents' conceptions of what they are doing. And this is a difference from the objects of natural sciences, since atoms, amoebas, light-sources and (arguably) dogs and cats do not have any conception of what they are doing. But what follows from this? It has often been said that it establishes a 'partial identity between subject and object' in the social sciences. But if it is put this way, it might be replied that there is a partial identity of subject and object in the natural sciences too, since we ('subjects') are not only social agents but also living beings (part of the object of biology), bodies with a chemical composition, and so on. But Bhaskar's point is rather different. When he comes to base a 'relational limit' to naturalism in the human sciences on this ontological difference (PN, p. 47ff), he is working with the fact that, not just the human 'subjects', but the sciences themselves as theoretical formations are (a) potentially part of their own object, and (b) able to enter into logical relations (confirmation, implication, contradiction) with other parts of their objects (agents' conceptions). It is this latter point that makes explanatory critiques in general possible, while the former gives rise to the special sort of explanatory critique in which a social science

turns reflexively on itself: sociology of sociological knowledge, and so on. This is a limit to naturalism only in the sense that it is a possibility to which there is no parallel in the natural sciences. It does not either constrain or facilitate the work of social science in general.

But does the presence of ideas in the object of the human sciences provide a compensator for the absence of experiments? We have seen in chapter 5 that Bhaskar thinks it does. I have accepted that the agents' conception of what they are doing must be the starting point, in that an account of a social activity which excluded the agents' own description of it would be radically misleading, even if the agents' description includes errors that the social science can eventually correct. And granted that these agents' descriptions give us some data without which we cannot proceed, these data may form the premises of transcendental arguments, as suggested in chapter 5. But a comparison with our knowledge of the natural world will help to put these data in their place.

We know quite a lot about the social world simply from being agents within it. But we also know quite a lot about the natural world simply from being agents within it. The practices of agriculture and stockbreeding, manufacture and cookery, navigation and building, all of them conceptualized in the minds and discourse of their agents, provide the original raw materials for the scientific production of knowledge. These conceptualized practices are no less informative about the natural world than the practices of commerce and statecraft, conflict and cooperation, 'knowing oneself' and 'knowing others' are about the human world. There is not only the connatural knowledge which we have of the human world by being human, but also that of the 'biosphere' which we have by being alive, of the world of physics by being beings which can push and pull, jump and fall, feel heat and cold and so on. The difference in the significance of these sources of information between the natural and the social sciences is not that the social sciences have more input from them, but that the natural sciences have gone so much further beyond what these sources can give us. For in the natural and social sciences alike, these sources are fallible and corrigible. The human sciences as much as the natural can, as we have seen, be counter-phenomenal, they can expose

ideological illusion, unconscious motivation, and so on. They
would lack much of their interest if they could not surprise us,
and they could not be emancipatory if they could not undeceive
us. Their input from agents' conceptions is no more
authoritative than such input is in regard to the natural world;
but our capacity to correct, revise and add to the knowledge
derived from agents' conceptions is immeasurably more
advanced in those sciences where experiments are possible. The
teachability even of an experimental natural science doubtless
presupposes our initial familiarity with (for example) heat, light
and sound, push and pull, speed and weight; but before we
have gone very far, we have redefined such concepts and left
our homely understanding of them far behind. The hermeneutic
moment is so prominent in the human sciences not because it is
a more essential stage or a more reliable or informative source
than in the natural sciences, but because, in the absence of
experiments, we have so little else. As a result, we are also much
more likely to get things wrong and much less likely to correct
them in the human than in the experimental natural sciences.
The plurality of theories in the field at any time in the human
sciences is partly due to this. Of course it is also partly because
conflicts of interest affect work in some human sciences. But
such conflicts hardly affect linguistics, yet it is as controversial as
any human science. And surely everyone must have been struck
by the incongruity of the fact that humankind can solve abstruse
problems of theoretical physics so elegantly and so empower-
ingly, yet flounders in the dark when it comes to running an
economy, or even a love affair.

My conclusion from what we know about the ontology of the
human world is that it gives grounds for scepticism about the
prospects of the human sciences. In the next section I will spell
this conclusion out, and in the final one suggest what does
remain for science-like knowledge of the human world.

The Epistemological Thicket

It is generally agreed that the human sciences are in a much
more controversial state, and also a much less advanced state,
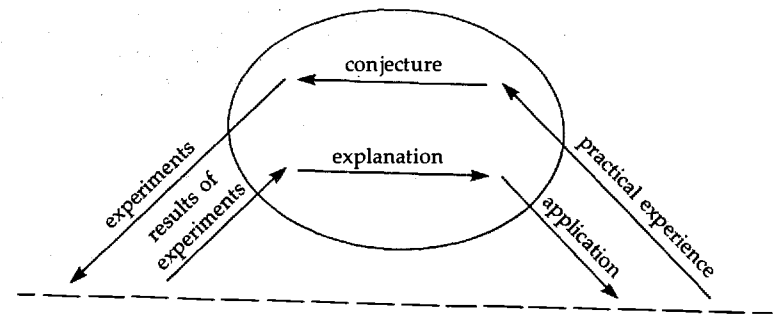
Figure 8.1 Relation between a Science and Its Real Object (Source:
'Scientific Realism in the Human World: The Case of
Psychoanalysis' p. 17)

than the natural sciences. Sometimes it is thought that they are
less advanced because they are younger, and will be proper
sciences when they grow up. But in terms of duration of
existence, they are not younger: Hobbes was fifty-four years
older than Newton, Adam Smith twenty years older than
Lavoisier, Marx only nine years younger than Darwin, Freud
twenty-three years older than Einstein. Of course, in each pair
the scientific status of the social researcher's work is disputed
while that of the natural scientist is not. But in intention all were
scientists, and if the success of the inquiries into the natural
world is certain and that of those into the human world in
doubt, some other explanation is required than dates of origin.

If I am right that we have no special insight into the human/
social world on the basis that 'we made it', we are returned to
the question: what are the prospects for the human sciences
given that they have no experiments in the relevant sense —
they study systems which are not only further removed from
natural closure than those studied by the theoretical natural
sciences, but also insusceptible to artificial closure? Elsewhere
(in my paper 'Scientific Realism in the Human World: the Case
of Psychoanalysis'), I have described their situation as 'concrete-
bound', as expressed by Figure 8.1 and the following comments
on it.

This figure represents the situation in the experimental
sciences. Practical experience leads to a degree of concrete

knowledge of the object (the farmer's knowledge of the soil, the navigator's knowledge of the stars); this suggests explanatory conjectures which produce abstract models which can then be tested by experiments; the results of the experiments lead to the confirmation, refutation or revision of this part of the science; the resulting body of tested abstract knowledge can be used to explain the concrete object more accurately, and this explanation used to generate new and more effective kinds of practical interaction with the object, which in turn will yield new practical experience of it, and so on. In the non-experimental sciences, the process represented by the arrows at the left of the figure are absent. Hence the abstract models that are conjectured cannot be tested before the explanations they generate are used in practice.

To an extent, input from experimental sciences into the abstract part of non-experimental sciences can relieve the concrete-boundness of the latter; this is particularly important in non-experimental *natural* sciences like geomorphology and meteorology. It has a much more limited role in the human sciences, since in this case the input would only be from relatively distant regions of science, or perhaps more importantly, from philosophical conclusions derived from other sciences: for instance from the ontology of transcendental realism with its concepts of open and closed systems, structure and tendency, stratification and emergence, and so on. But this is a far cry from saying that, even with the best philosophical midwifery in the world, human sciences could be born with any expectation of reaching a maturity comparable with that achieved by the natural sciences. The human sciences are doomed to neoteny.

Elsewhere (*Scientific Realism and Socialist Thought*, chapter 4) I have used the analogy of the knowledge one might have of different types of terrain, if we had no helicopters or other vantage points above its surface, and were reliant only on what we saw while walking about. In open country, we might be able to make quite accurate maps. This corresponds to the position in the experimental sciences. The situation in sciences without experiments is more like mapping impenetrable forest transected by narrow, winding paths. Input from ontology derived philosophically from the practice of other sciences may

yield some background information (as of the geological formations present in a terrain), but we can never expect scale maps of the forest.

Keeping within this metaphor, I take it that Bhaskar's own view, in this respect like that of the hermeneutic theorists, is that the human world is actually open country too, but for different reasons. Experiments constitute the openness of the natural terrain, our self-understanding constitutes that of the human terrain. But Bhaskar himself has pointed out the limiting of our understanding even of our own actions by unacknowledged conditions, unintended consequences, unconscious motivation and tacit skills.

It remains to be asked whether there are not closer analogues to experiment in the human sciences than I have admitted. Some research in the human sciences certainly involves setting up artificial test situations, which are *called* experiments. Statistical controls are regarded as analogues of experiment, and the use of mathematics generally is treated (as it was even by Marx) as a sign that scientific maturity has been reached. The question here is whether we are dealing with genuinely measurable data, or whether quantification is imposed inappropriately, at the cost of suppressing crucial data.

The experimental sciences can *actualize* many of their abstractions. That is what experiments do when they establish closed systems. So they can not only test and establish the reality of the various tendencies that are at work in nature, but also *measure* them. This enables these sciences to be genuinely quantitative ones; they can use mathematics to calculate what happens under given conditions. This does not mean that the seventeenth-century ideal of a purely mathematical science of nature is realizable; the stratification of nature and the reality of emergence mean that even within nature irreducibly qualitative differences have real effects. Nevertheless the value of mathematics in the natural sciences is great and undeniable. Outcomes in closed systems can be calculated. One could not be a physicist without being a mathematician. In the human sciences on the other hand, outcomes cannot be calculated, and quantification is vague. We may know that the rate of profit is falling, but we could never have accurate knowledge of what it is. Where quantification is presented as if it were exact, it is

almost always a sign that some qualitative distinction has been misleadingly ignored. There are (questionable) commercial and managerial reasons why this occurs, but at the level of the philosophy of science it is merely an inappropriate aping of features of the experimental sciences which make no sense in the absence of experiments. Social science departments in universities often require their students to take courses in statistics, but the use to which these sciences put statistics tempt one to suspect that innumeracy is a positive virtue in a social scientist. It is not just a matter of *careless* use of statistics or quantification generally. As soon as mathematical calculation is taken as a desideratum, qualitative distinctions which are the crucial ones, causally and morally, are lost sight of. To start with a simple practical example from university life, consider virtually any form of assessment, whether of students by lecturers or lecturers by students; for instance, the marking of essays. On the one hand, the marker may make comments, for which there will be objective grounds in the essay (or if there are not, the comments can be contested and the mistake identified). These comments may point to specific ways in which the essay could be improved. On the other hand, a numerical mark may be given. Any marker who takes their work seriously must often have been struck and disturbed both by the crudity of the judgement expressed in the mark, in that it covers over the distinction between say a highly creative essay full of mistakes and an errorless but uninspired essay; and by the ultimate arbitrariness of the mark. Yet because numerical marks can be added and averaged, it is in these, in the end, that assessment consists. The mathematical processing of the marks gives an illusion of objectivity, like the 'fairness' of a game of cards in which everything depends on the luck of the initial deal. It is as if there were a magic formula for transmuting subjectivity into objectivity: just add mathematics. But there is not: if you average subjective judgements all you get is an average subjective judgement.

An example with more vital implications is that of economics. So long as economics is tied to a model of rational decision-making based on calculation of commensurable values, it is unable to take account of use-values (in Marxian terms) and hence systematically ignores environmental values and quality

of life generally (see my paper 'Value, Rationality and the Environment'). A genuinely rational procedure of economic decision-making would involve irreducibly qualitative judgements. The future of life on earth could depend on the recognition of this, and the future of a decent life for human beings certainly does. Of course, the tendency to quantitative calculation in economics is not just a philosophical error, it is deeply rooted in the market economy. But even a socialist economy would be no improvement in this respect unless it involved a rejection of calculation in favour of a more comprehensive form of practical rationality.

At the purely methodological level, the tendency to focus on what can be measured leads to systematic blindness to certain features of the human world.

An example of this is the chapter on 'The Dream Theory' in Fisher and Greenberg's *The Scientific Credibility of Freud's Theories and Therapy*. The conclusions of this chapter are presented as if they were the results of some kind of test between Freud's theory of dreams and an alternative account, according to which there is no latent content to the dream, and the manifest content is explained by various features of the dreamer's life situation, (age, sex, culture, health and so on). But it is easy enough to see that this result is generated by the method adopted, that is, the statistical analysis of large samples of dreams, counting the number of times particular elements occur, e.g. physical activity, friendly or aggressive interaction with members of one's own sex or the opposite sex, and so on. The data discovered by these methods may be interesting enough facts, if usually rather unsurprising (for example, that recently disabled people dream about physical activity more, and long disabled people less, than able-bodied people do). But the kind of facts that *can* be discovered this way is limited; and it is the predilection for the statistically analysable that determines that it is these and no other facts that are discovered. Thus Fisher and Greenberg claim (p. 66) that 'some of the best validated projective tests' involve 'no special search for "The Latent Content"'. So however well validated these tests are with respect to whatever it is that they are testing, it is not surprising if they don't find what they are not looking for. They don't find the fruit because they look in the bread bin not the fruit bowl.

But they then claim their discovery of bread to be a refutation of the existence of fruit:

> What reasonable conclusions about 'dream interpretation' can we now offer on the basis of the review unfolded in this chapter?
>
> First, there is no rationale for approaching a dream as if it were a container for a secret wish buried under layers of concealment. (pp. 67–8)

My point is not that these statistical analyses were looking for the wrong things, as if the same method could have been used to look for Freudian things. It is that the method inevitably passes over the Freudian things. For in order to do statistics one must abstract from the particularity of the dream elements; a dream act which is categorized for statistical purposes as *an act of friendly interaction with a member of the same sex* might also be an act of plotting to assassinate the emperor. Since acts under the former description are common and those under the latter rare, the latter will be of little use for statistical analysis. But it might be essential to the function of the act in the dream.

> The dream is a fantasy structure in its own right and susceptible to direct forms of inspection and partitioning. It should be cautioned, though, that enough dream data have to be available to enable a reliable sampling. ... In studying almost any form of human behaviour, it has been found that unless a reasonably representative sample of that form is secured, a reliable job cannot be done. There is no reason why dream behaviour should be treated as a unique exception to this rule. (p. 68)

It has not been *'found'* that such a sample is necessary; statistical research presupposes it. But for just this reason, statistical research is limited in its subject-matter to widespread phenomena. Hence Fisher and Greenberg are quite wrong to say that 'such an approach places no limits on possible "depth" interpretations of dream content' (p. 70). Indeed they go straight on to say that the scales used for such inquiry 'assume, too, that the symbols utilized to carry these meanings are widely shared and therefore do not need to be decoded in terms of the dreamer's private associations'. This is an assumption not a discovery, and it is an assumption that precludes discoveries of

the sort claimed by Freud. The fact that some discoveries can be made on this assumption does not cast any suspicion on discoveries which can only be made by different modes of inquiry, unrestricted by such an assumption.

In this section I have mainly been saying what I think we should *not* expect from a human science, given critical realist understandings of science and of the place of the human world in stratified nature. It behoves me to say something about what sort of study of the human world approximates closest to science in the *right* respects.

What should a Scientific Study of the Human World Be Like?

The following remarks should not be taken as any kind of dogmatic prescription. They are intended not to discourage any kind of social research (even, in its place, statistics), but rather to encourage kinds of research which have been undervalued in some circles.

I start from the concrete-boundness of the human sciences. I have no wish to revel in it. It means that our information about the human world is far more likely to be in error than is our information about the natural world. This is not to be rejoiced at — it can have quite disastrous consequences; but so can the refusal to admit it.

Concrete-boundness means that we can only directly study concrete entities, not the diverse mechanisms and tendencies which make them what they are. We can study the latter only through the former, not by isolating them in closed systems. The further our theory gets away from the concrete towards the abstract (which it must nevertheless do) the more prone to error it is.

According to Lenin, the concrete analysis of the concrete conjuncture is the heart of Marxism. It should also be the heart of all good practice in the human sciences. But note: the concrete *analysis* of the concrete conjuncture. In order to explain the concrete conjuncture we have to unravel by analysis (in thought) the multiple mechanisms and tendencies which make it what it is. 'The concrete is concrete because it is the

concentration of many determinations, hence unity of the diverse', said Marx (*Grundrisse*, p. 101). We have no explanation of any concrete conjuncture until we have identified the many determinations — yet we cannot isolate them in reality, as we can in experimental sciences. Once again Marx: 'in the analysis of economic forms neither microscopes nor chemical reagents are of assistance. The power of abstraction must replace both' (*Capital vol. 1*, p. 90).

The two opposite errors to be avoided here are the belief that we can somehow isolate the 'many determinations' as we can in the experimental sciences, and the belief that we should not be analysing out the many determinations at all, but somehow apprehending the concrete whole without analysis.

The former always has the effect that we move away from the concrete not towards genuinely explanatory abstractions (that is, the concepts of the many determinations that come together to form the concrete conjuncture) but towards mere generalizations, 'bad abstractions' since they are arrived at merely by ignoring certain features of the concrete conjuncture in order to group it with other conjunctures with which it has superficial similarities. This is what we have seen already going on in the statistical approach to the interpretation of dreams: the rich and convoluted complexity of the individual dream is bypassed in order to enter it into a statistical sample which can be sorted by various general features such as whether they include friendly or hostile interaction with the same sex and the opposite sex, and so on. There is no guarantee that these particular criteria of classification have any explanatory power in accounting for these dreams at all. Freud suggests, for example, that a dream about someone's death may in one case express an (unconscious) wish that they should die, while in another it may have a purely instrumental function, such as leading to a meeting (in the dream) with a mutual friend. An entry of such a dream into a statistical sample is likely to miss this difference, and if it does not, that can only be since some close analytic work has already been done on the dream, to find out how the death is related to other elements in the dream, to the dreamer's unconscious, and so on. The abstractions which we require can only be discovered by focusing on the 'concrete conjuncture' (in this case the dream) and analysing the interconnection of its

details. Likewise if the concrete conjuncture is, say, the Russian Revolution, we will unravel the 'many determinations' that are conjoined in it and explain it only by zooming in on it till we can see the interacting forces (the proletariat and peasantry, the army, the parties; and more abstractly: Russia's backwardness, its imperialism, its uneven development). If instead we move further into the distance (that is, neglect the minute particulars in favour of bringing in comparisons with other historical conjunctures, such as the French or Chinese Revolutions) we will easily arrive at generalizations without explanatory power — accidental similarities, contrived analogies and so on. Real structural similarities can only be arrived at after the depth analysis of the separate conjunctures has been done.

On the other hand, we must avoid what might be called the 'fallacy of analysis' fallacy: the idea that concrete particulars — whether historical situations, individual characters or whatever — are such integral wholes that to analyse them is to falsify them. Thus Sartre makes a fair point against generalizations masquerading as expanations when he says that to explain Pierre's fondness for rowing by his (somehow prior) fondness for sport in general is 'to assume the priority of the abstract over the concrete — as if the fondness for play existed first in general to be subsequently made specific . . . in the love of sport, the latter in the fondness for rowing' and so on (*Being and Nothingness*, p. 562). Yet a fondness for rowing may well be explained by the conjunction of several desires and beliefs. But he goes on to propose an existential psychoanalysis based on the principle that 'man is a totality not a collection. Consequently he expresses himself as a whole even in his most insignificant and his most superficial behaviour' (p. 568). So the fondness for rowing is explained by 'himself as a whole' — about which, presumably, nothing can be said, since to say anything is to pick out an aspect, to abstract, to analyse. Any attempt to unravel the 'many determinations' which are conjoined in us are rejected by Sartre, since he says that 'the *man* disappears . . . the being whom we seek vanishes in a dust of phenomena bound together by external connections' (p. 561). But dust bound together by external connections is no longer just dust; it is, in this case, a man. In fact, since we cannot say everything at once, to speak is to analyse, and Sartre himself cannot help but do so once he gets

going — as indeed the term 'existential psychoanalysis' concedes. And this analysis is not a falsification, not even an inevitable one. We really are constructed out of many determinations. At times, to a degree, they may come apart. Consider Freud's account of sexual perversion as the falling apart, or incomplete integration, of the *component instincts* out of which adult sexuality has been constructed (oral, anal, genital; active and passive scopophilia, tenderness, aggression and so on). To the extent to which we really are totalities, that is an achievement, not a given.

At times (as in the last quotation) Sartre, like other adherents of the 'fallacy of analysis' fallacy, speaks as if to analyse were literally to pull apart — an act of violence rather than of understanding. Melanie Klein tells how a child with a block against division sums turned out to associate them with the cutting up of a person's body. I suspect that some such unconscious 'thought' underlies much holistic antipathy to analysis.

It is beginning to look as if psychoanalysis — generally regarded as something of a maverick among the human sciences, should be treated as the paradigm of good practice in this area. It derives all its theories from the analysis in depth of particular individuals. Insofar as it generalizes, it generalizes about mechanisms and tendencies discovered in such analyses. It generates no statistical predictions whatever; if it corrects, revises and supplements its discoveries, it does so not on the basis of any statistical data (which it can always explain away), but on the basis of more depth inquiries into more symptoms of more people. The 'more' adds not statistical confirmation, but new data: it is *different* phenomena, not more of the same, that refine and complexify this science.

It is often said that psychoanalysis is unscientific because it generalizes from few cases, but this is to misunderstand the nature of psychoanalytical discoveries. If a botanist discovers a new species, no one accuses him or her of generalizing from few instances, however rare the species is. Psychoanalysis, in unravelling our many determinations, identifies mechanisms and tendencies present in some one or few individuals, and therefore possibly present (though possibly latent) in others. A mechanism may exist unoperative or operate unrealized or be

realized unperceived. So it may not show up in a statistical survey. On those occasions when Freud does take leave of existential quantifiers ('there is some x such that . . .') and speak in terms of universal quantifiers ('for all x . . .'), it is in general because he has tied in some psychoanalytical concept to some general biological fact about human beings. Hence his remark to Geza Roheim, when told that there was no anal sexuality among the Trobriand Islanders: 'Don't they have an anus then?' Since it has also been claimed that the Trobriand Islanders have no Oedipus Complex, he might have asked: 'Don't their mothers have partners, then?' But if psychoanalysis has some use for universal statements ('all') and much for existential ones ('some'), it has none for 'most'.

I do not wish to claim that psychoanalysis contains no errors. I have already given reasons why it must be expected, as a concrete-bound discipline, to be far from the exactness of the experimental sciences. And I do believe that Freud made some serious errors (for one: his tendency to describe instincts as if they pre-existed any inter-personal input — a tendency corrected somewhat by the Kleinians). My point here is mainly that the deep analysis of the minute particulars of some concrete conjuncture, rather than superficial knowledge of great statistical populations, should occupy the foreground of the picture of the human sciences. And if anyone objects that while such depth studies have their place, they are not science, I would only point out that there are areas in the natural sciences where research work takes this form too.

Concluding Remarks

If these comments of mine are correct, how much of Bhaskar's critical naturalism still stands? Most of it, I think. To parody the old election slogan of the German Social Democrats ('not different but better'), I would say that the situation of the human sciences is not different but worse than Bhaskar has portrayed it. I have marked this view elsewhere by calling them 'epistemoids' rather than sciences (in *Scientific Realism and Socialist Thought*). I do not expect social scientists ever to achieve the sort of consensus that well-established natural scientific

communities enjoy, even if all the distortions of such things as class interest were removed. And this is not because of 'essentially contested concepts'; nor is it the sort of diversity in which we should rejoice. We will not reach consensus because we will continue to make too many mistakes.

Nevertheless, work in the human sciences can go on, and they can be liberated from the constraints of positivist and non-realist philosophies. They can become at once non-reductive, qualitative, explanatory and counter-phenomenal; and insofar as they do, they can also make essential contributions to human emancipation. Their capacity for explanatory critiques can ground a 'scientific politics' that has nothing in common with the bureaucratic manipulation or triumphalist predictions that have often usurped that title. Scientific in the sense of giving more knowledge than untutored experience could about the causes of oppression and the conditions for our emancipation from them.

Finally, it may be useful to mention potentialities of critical realism that are relatively under-theorized by Roy Bhaskar — rather as Bhaskar himself has pointed out how Marx has left one side of several pairs of ideas undeveloped compared with the other.[1]

In the first place, Bhaskar has been primarily concerned with *scientific* forms of knowledge. However, since this does *not* in his case reflect either an empiricist denial of essential differences between science and everyday knowledge, or a positivist contempt for the latter, the whole question of the epistemology of everyday life, and its ontological foundations, is left open. Yet there can be transcendental arguments from the cognitive (and other) aspects of non-scientific practices — of work and play, conversation and mutual aid, love and strife, self-expression and aesthetic contemplation. Such arguments are often found in a tradition of philosophizing that has an ambiguous or perhaps ambivalent relation to realism, namely existential phenomenology. What would a realist version of Heidegger's existential analytic look like? Or a realist critique of Sartre on concrete relations with others (*Being and Nothingness*, part 3, chapter 3)? Is it not precisely the absence of a depth-realist notion of counter-phenomenal (and therefore potentially liberating) truths that vitiates those brilliant phenomenological inquiries?

Second, while explanatory critiques clearly have application in morality as well as politics, Bhaskar has not spelt this out. I have suggested that an ethics based on explanatory critiques would look rather like Spinoza's. However, there is another leaf that ethics might take out of the critical realist book. Philosophy from Descartes to logical positivism and 'postmodernism' has been thoroughly, and in many ways increasingly, anthropocentric. Critical realism has shown us a way out of this subjectivistic madness, so far as ontology is concerned. It may be asked how this is related to the argument of some ecologists that there are *values* independent of us? I suggest that the relation is parallel with that between critical realism and socialism, that is:

1. critical realism does not *imply* that there are values independent of us;
2. nevertheless it undermines several prevalent arguments against there being such values; and
3. it suggests a non-anthropocentric ethic by way of homology.

At any rate it is clear that critical realism is at least *compatible* with belief in non-anthropocentric values in a way in which empiricism and idealism are not.

Perhaps it is not only *human* emancipation for which critical realism can do the philosophical underlabouring.

Notes

1. Thus in RR, pp. 133–6, and again in Appendix 2 to PIF, Bhaskar argues that Marx criticized both empiricism and idealism but left his critique of empiricism under-theorized relative to his critique of idealism; that he gave an account of science as both objective and a work of cognitive labour, but left the former aspect undeveloped relative to the latter, and so on.

Biographical Note

Roy Bhaskar was born in London in 1944 to theosophist parents. He is the elder of two brothers. His interest in philosophical problems began early; apparently he was 'confounding the doctors of the law' on free will and determinism at the age of five! In his school days, his passion was cricket: he once amassed 400 runs in one match, and he corresponded with Len Hutton and Peter May. But cricket's loss was philosophy's gain when hay fever interrupted his cricketing career at the age of eleven. He first turned to music, and became, by his own admission, a bad drummer in a bad band. However, he retains his love of music, popular and classical, and of dancing.

In 1962 Roy obtained a scholarship to Balliol College, Oxford, and in the following year began his course in Philosophy, Politics and Economics. In the interim, he did various jobs, including bouncer in a Brighton night club.

After graduating with a First, he initially began research at Nuffield College on the relevance of economic theory for underdeveloped countries, while lecturing in economics at Pembroke College. But he soon turned to philosophy, in which Rom Harré supervised his research. They exchanged work and found that they had many shared concerns: indeed Roy describes it as intellectual love at first sight.

Roy was active in the 'events of 1968', for which he was on several occasions summoned before the Proctor. Ever since this time he has identified with the left; he was a founder-member of the Socialist Society, and through this has been involved with the Chesterfield Conferences and the Socialist Movement that arose from them, playing a central role in that movement's Philosophy Policy Group. It was in 1968, too, that he met Hilary Wainwright, who was studying sociology and was also a political activist, and who has since become well known for her books on socialist feminism and the labour movement. They married in 1971, and spent their honeymoon in the Liberated Zones of Mozambique and Angola, as guests of Frelimo and the MPLA.

In 1973 Roy began lecturing in philosophy at Edinburgh, and completed *A Realist Theory of Science* between November 1973 and February 1974. He has since held posts at the University of Sussex and City University, London, and from 1986 at Linacre College, Oxford. Since 1985 he has been an organizer of and regular participant in the annual 'Standing Conference on Realism in the Human Sciences'.

The reading and writing of work on philosophy and social theory is a passion with Roy, and since his student days he has penned hundreds of thousands of words on these subjects, most of them unpublished. At the time of writing (May 1993) he has just completed his long-awaited book, *Dialectic: The Pulse of Freedom*, rounding out critical realism with a theory of totality and negativity. Planned works include a critical history of western philosophy, a text on the problems of philosophy, a book entitled *Philosophical Ideologies* which will engage with the ideas of Nietzsche, Heidegger and Derrida, and a book on the politics of critical realism, to be written jointly with Hilary Wainwright.

Bibliography

Althusser, Louis, *Reading Capital*, NLB, London 1970
____, *Lenin and Philosophy*, NLB, London 1971
Ayer, A.J., *Language, Truth and Logic*, Gollancz, London 1964
Benton, Ted, 'Natural Science and Cultural Struggle: Engels and the Philosophy of the Natural Sciences' in J. Mepham and D.-H. Ruben, eds, *Issues in Marxist Philosophy* vol. 2, Harvester Press, Brighton 1979
____, 'Realism in Social Science', *Radical Philosophy* no. 27, 1981
____, 'Marx, Humanism and Speciesism', *Radical Philosophy* no. 50, 1988
____, 'Marxism and Natural Limits: An Ecological Critique and Reconstruction', *New Left Review* no. 178, 1989
____, 'Ecology, Socialism and the Mastery of Nature: a Reply to Reiner Grundmann', *New Left Review* no. 194, 1992
____, *Natural Relations. Ecology, Animal Rights and Social Justice*, Verso, London 1993
Berger, P. and Pullberg, S., 'Reification and the Sociological Critique of Consciousness', *New Left Review* no. 35, 1966
Bhaskar, Roy, 'Feyerabend and Bachelard: Two Philosophies of Science', *New Left Review* no. 94, 1975
____, *A Realist Theory of Science*, 2nd edn, Harvester Press, Brighton 1978
____, *The Possibility of Naturalism*, 2nd edn, Harvester Wheatsheaf, Hemel Hempstead 1989 (first published 1979)
____, *Scientific Realism and Human Emancipation*, Verso, London 1986
____, *Reclaiming Reality*, Verso, London 1989
____, *Philosophy and the Idea of Freedom*, Blackwell, Oxford 1991
____, *Dialectic: The Pulse of Freedom*, Verso, London 1994
Blake, William, *Complete Writings*, Oxford University Press, London 1966
Bradley, F.H. *Ethical Studies* (1876), Oxford University Press, Oxford 1962
Chomsky, Noam, *Rules and Representation*, Blackwell, Oxford 1982
Clegg, Sue, 'Studying Child Sexual Abuse', *Radical Philosophy* no. 66, 1994

Collier, Andrew, *R.D. Laing*, Harvester Press, Hassocks 1977
____, 'Scientific Realism in the Human World: The Case of Psychoanalysis', *Radical Philosophy* no. 29, 1981
____, *Scientific Realism and Socialist Thought*, Harvester Wheatsheaf, Hemel Hempstead 1989
____, *Socialist Reasoning*, Pluto Press, London 1990
____, 'The Materiality of Morals: Mind, Body and Interests in Spinoza's Ethics', *Studia Spinozana* no. 7, 1992
____, 'Value, Rationality and the Environment', *Radical Philosophy* no. 66, 1994
Davidson, Donald, 'Actions, Reasons and Causes', in Alan White, ed., *The Philosophy of Action*, Oxford University Press, Oxford 1968
Davies, Paul, *God and the New Physics*, Penguin, Harmondsworth 1990
Durkheim, Émile, *Rules of Sociological Method*, Macmillan, New York 1964
Edgley, Roy, *Reason in Theory and Practice*, Hutchinson, London 1969
____, 'Reason as Dialectic: Science, Social Science and Socialist Science', in Roy Edgley and Richard Osborne, eds, *Radical Philosophy Reader*, Verso, London 1985 (reprinted from *Radical Philosophy* no. 15, 1976
Farrington, Benjamin, *The Philosophy of Francis Bacon*, Phoenix Books, Chicago 1966
Feyerabend, Paul, *Against Method*, 3rd edn, Verso, London 1993
Fisher, Seymour and Greenberg, Roger, *The Scientific Credibility of Freud's Theories and Therapy*, Harvester Press, Hassocks 1977
____, eds, *The Scientific Evaluation of Freud's Theories and Therapy*, Harvester Press, Hassocks 1978
Fisk, Milton, *Ethics and Society*, Harvester Press, Brighton 1980
Flew, A.G.N., *Thinking About Social Thinking*, Blackwell, Oxford 1985
Freud, Sigmund, *Jokes and Their Relation to the Unconscious*, Routledge & Kegan Paul, London 1966
____, 'Beyond the Pleasure Principle', in *On Metapsychology* Penguin, Harmondsworth 1984
____, 'The Loss of Reality in Neurosis and Psychosis', in *Collected Papers* vol. 2, Hogarth Press, London 1957
____, 'A Case of Obsessional Neurosis', in Sigmund Freud's *Collected Papers vol. 3*, Hogarth Press, London 1969
Freud, Sigmund and Breuer, Joseph, *Studies in Hysteria*, Penguin, Harmondsworth 1974
Geach, Peter, 'Aquinas', in Elizabeth Anscombe and Peter Geach, *Three Philosophers*, Blackwell, Oxford 1967
Gramsci, Antonio, *Prison Notebooks*, Lawrence & Wishart, London 1971
Hampshire, Stuart, 'Disposition and Memory', in *Freedom of Mind*, Oxford University Press, London 1972

Harré, Rom, *Great Scientific Experiments*, Oxford University Press, Oxford 1983

Heidegger, Martin, *Being and Time*, Blackwell, Oxford 1967

____, *Basic Writings*, Routledge & Kegan Paul, London 1977

Hessing, Siegfried, ed., *Speculum Spinozanum*, Routledge & Kegan Paul, London 1977

Hindess, Barry and Hirst, Paul, *Mode of Production and Social Formation*, Macmillan, London 1977

Hume, David, *A Treatise of Human Nature*, Dent, London 1911

Kant, Immanuel, *Critique of Pure Reason*, Macmillan, London 1964

Kuhn, Thomas, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago 1970

Laing, R.D., *The Voice of Experience*, Penguin, Harmondsworth 1983

Lass, R., *On Explaining Language Change*, Cambridge University Press, Cambridge 1980

Lawson, Tony, 'Realism and Instrumentalism in the Development of Econometrics', in C. Gilbert and N. De Marchi, eds, *History and Methodology of Econometrics*, Oxford University Press 1989

____, 'Realism, Closed Systems and Friedman', in *Research in the History of Economic Thought and Methodology*, vol. 10, JAI Press, 1992

____, 'Methodology: Non-optional and Consequential', *Royal Economic Society Newsletter*, October 1992

____, 'The "Lucas Critique"': Inadequate diagnosis, wrong prescription', mimeograph, Cambridge 1993

____, 'Economics and Expectations', in S. Dow *et al.*, *Keynes, Knowledge and Uncertainty*, Edward Elgar, Aldershot 1994

____, 'Hayek and Realism: a Case of Continuous Transformation', in M. Colonna, H. Haggemann and O. Hamouda, eds, *Capitalism, Socialism and Knowledge; The Economics of F.A. Hayek*, Edward Elgar, Aldershot 1994

____, 'Critical Realism and the Analysis of Choice, Explanation and Change', *Advances in Austrian Economics* vol. 1, no. 1, 1993

Lévi-Strauss, Claude, *Tristes Tropiques*, Penguin, Harmondsworth 1976

Lewis, C.S., *Studies in Words*, Cambridge University Press, Cambridge 1960

____, *Surprised by Joy* Fount, London 1977

____, *First and Second Things*, Fount, London 1985

Lukes, Steven, *Power: A Radical View*, Macmillan, London 1974

MacIntyre, Alasdair, *After Virtue*, Duckworth, London 1985

MacMurray, John, *The Self as Agent*, Faber and Faber, London 1957

Magee, Bryan, *The Philosophy of Schopenhauer*, Oxford University Press, Oxford 1987

Manicas, Peter, *A History and Philosophy of the Social Sciences*, Blackwell,

Oxford and New York 1987

Marx, Karl, *Capital vol. 1*, Penguin, Harmondsworth 1976

____, *Grundrisse*, Penguin, Harmondsworth 1973

Marx, Karl and Engels, Frederick, *Selected Works in One Volume*, Lawrence and Wishart, London 1968

____, *The German Ideology*, Lawrence and Wishart, London 1970

Mill, John Stuart, 'On Liberty', in *Utilitarianism, Liberty and Representative Government*, Dent, London 1910

____, *System of Logic*, 8th edn, Longman Green, London 1861

Moore, G.E., *Principia Ethica* (1903), Cambridge University Press, Cambridge 1959

Morton, A.L., ed, *Freedom in Arms*, Lawrence and Wishart, London 1975

Mundle, C.W.K., *A Critique of Linguistic Philosophy*, Glover and Blair, London 1979

Nietzsche, Friedrich, *Genealogy of Morals and Ecce Homo*, ed. Walter Kaufman, Vintage Books, New York 1969.

____, *The Birth of Tragedy and the Genealogy of Morals*, Doubleday Anchor, Garden City N.Y. 1956

Outhwaite, William, *New Philosophies of Social Science: Realism, Hermeneutics and Critical Theory*, Macmillan, London 1987

Pateman, Trevor, *Language in Mind and Language in Society*, Clarendon Press, Oxford 1987

Pope, Alexander, *Poems*, Methuen, London 1963

Ramsay, Maureen, *Human Needs and the Market*, Avebury Books, Aldershot 1992

Rorty, Richard, *Philosophy and the Mirror of Nature*, Blackwell, Oxford 1980

____, *Contingency, Irony and Solidarity*, Cambridge University Press, Cambridge 1989

Ruben, David-Hillel, *Marxism and Materialism*, Harvester Press, Hassocks 1977

Ruse, Michael, 'Definitions of Species in Biology', *British Journal of the Philosophy of Science* no. 20, pp. 97–119, 1969

Russell, Bertrand, *A History of Western Philosophy*, Unwin, London 1984

Sartre, Jean-Paul, *Being and Nothingness*, Methuen, London 1957

Sayer, Andrew, *Method in Social Science: A Realist Approach*, Routledge & Kegan Paul, London 1992

Shapiro, Ian, *Political Criticism*, University of California Press, Berkeley 1990

Soper, Kate, 'Marxism, Materialism and Biology', in Mepham and Ruben, eds, *Issues in Marxist Philosophy vol. 2*, Harvester Press, Brighton 1979

_____, *On Human Needs*, Harvester Press, Brighton 1981

_____, *What is Nature?*, Blackwell, Oxford 1994

Spinoza, Baruch, 'Ethics', in Edwin Curley, ed., *Collected Works Vol. 1*, Princeton University Press, Princeton N.J. 1985

Scriven, M., 'Causes, Connections and Conditions in History', in W. Dray, ed., *Philosophical Analysis and History*, Greenwood Press, New York 1966

Sraffa, Piero, *Production of Commodities by Means of Commodities*, Cambridge University Press, Cambridge 1960

Suchting, Wal, 'Reflections upon Roy Bhaskar's "Critical Realism"', *Radical Philosophy* no. 61, 1992

Tester, Keith, *Animals and Society*, Routledge & Kegan Paul, London 1991

Timpanaro, Sebastiano, *On Materialism*, Verso, London 1980

Trotsky, Leon, *Problems of Everyday Life*, Monad Press, New York 1973

Westlake, Michael, *One Zero and the Night Controller*, Routledge & Kegan Paul, London 1980

Will, David, 'Psychoanalysis as a Human Science', *British Journal of Medical Psychology* no. 53, 1980

_____, 'The Progeny of Positivism: the Maudsley School and Anti-Psychiatry', *British Journal of Psychotherapy* vol. 1, no.1, 1984

_____, 'Science, Psychotherapy and Anti-Psychiatry', *British Journal of Psychotherapy* vol. 2, no. 3, 1986

_____, 'Psychoanalysis and the New Philosophy of Science', *International Review of Psychoanalysis*, no. 13, 1986

Wittgenstein, Ludwig, *Tractatus Logico-Philosophicus*, Routledge & Kegan Paul, London 1961

Name Index

Subject Index