Realism and Epistemology

DAVID PAPINEAU

1. *The question of realism*

'Realism' means different things to different people.

For most philosophers of science this century 'realism' has been the view that theories about scientific unobservables are to be taken at face value, as telling us the truth about an independent, albeit unobservable, reality. The opponents of such realists were the 'instrumentalists', who maintained instead that the worth of such theories about unobservables consisted in their usefulness as tools for organizing observations, or some such.

An analogous debate, in some ways more serious and in some ways less, is the traditional debate between 'realists about the external world' and their 'idealist' opponents. Here the idealists denied that *any* judgements about a world beyond our immediate experience should be taken at face value, including judgements about such mundane medium-sized physical objects as sticks and stones. The realist, siding with common sense and Dr Johnson, saw no reason to deny that such judgements answer to a non-mental reality.

One lesson of the past thirty years in philosophy has been to show us that these debates are misconceived. Both debates presuppose a level of judgement which is not itself problematic, a level of judgement the realist credentials of which do not get called in question. In the realism-instrumentalism debate this unproblematic level consisted of observational judgements. In the debate about the external world it consisted of judgements about sense-data, about the givens of sensory experience. And in both cases it was by contrast with such unproblematic 'givens' that the dubious judgements (about unobservables, about physical objects in general) were called in question.

We now know, following Wittgenstein, Sellars, and others, that this contrast between 'givens' and the rest of judgement is incoherent. All judgements are fallible. No judgement can gain its authority just by being made. I shall not argue against 'givens' in this paper. I shall take these arguments for granted and consider what happens to the question of realism once we give up the distinction between what is given and what is not.

One natural response is to conclude that once the given has gone the realist wins by default. If judgements about unobservables or physical objects are as good as any others, and in particular as good as those previously considered privileged, then surely it makes no sense to continue denying that they should be construed realistically. And indeed this was

the general consensus twenty years ago. It was widely accepted that both instrumentalism and idealism were premissed on an illegitimate distinction, and that realism should therefore be returned unopposed.

It is true that there are still some philosophers, like Bas van Fraassen, who do not take givens for granted, but nevertheless share the instrumentalist disbelief in theories about scientific unobservables.[1] I do not want to deny that one might have good reasons for distrusting certain kinds of scientific judgements. Indeed it will be a corollary of this paper, albeit one I shall not be able to explore at any length, that we ought to be more cautious that we actually are in trusting scientific claims. But this will not be *just* because such claims are about unobservables. And in any case I am not interested in the judgements we do not trust, but in those that we *do* trust, and for now I shall take it that these ought to include at least some part of what is normally considered 'scientific theory'.

I am interested in the judgements we do trust because even in this area realism does not nowadays go unopposed. Consider an alternative response to the rejection of the given. Instead of thinking of ourselves as having discovered that all our accepted beliefs are as good as the old 'givens', why not conclude instead that the judgements previously supposed to be privileged are as bad as the rest? That is, instead of extending our realism 'up' to the old suspect class, why not extend our anti-realism down to our most basic views about the world?

It might seem doubtful that there is much at issue here. If all accepted judgements are in the same boat, what does it matter if we call ourselves 'realists' or 'anti-realists'? But there certainly are contemporary philosophers[2] who find the idea of realism objectionable, who seem to find it important to deny that belief answers to an independent reality. These philosophers do not of course want us to *drop* our beliefs. They are not, like the instrumentalists or the idealists, recommending that certain of our beliefs should be rejected. It is just that they are unhappy with the idea that the aim of belief is the representation of an independent reality.

But, still, what does this come to, given that the contemporary anti-realists are not recommending that we change what we believe? It is true that there are also contemporary *realists*, people who find it important to insist that the aim of belief is truth, and that truth means correspondence to an independent reality. But is there really a substantial disagreement between the two sides? Do the realists and the anti-realists differ on how we should treat our beliefs, on what we should do with them, or is it just a difference in philosophical style? The realist wants to say a belief is true if it corresponds to reality. The anti-realist wants to say that we count something as real if it is

[1] B. van Fraassen, *The Scientific Image* (Oxford, Clarendon Press, 1980).

[2] I have in mind here particularly Richard Rorty (*Philosophy and the Mirror of Nature*, Oxford, Blackwell, 1980) and the Hilary Putnam of *Reason, Truth and History* (Cambridge, Cambridge University Press, 1981).

part of the view of the world that we accept. Aren't they just saying the same thing, but putting the emphasis in different places?

I think that the real difference between the realist and the anti-realist is located in their differing views about the justification of belief. Anti-realists will hold that at some point belief ceases to require justification. They will hold that an analysis of the notion of belief, or more generally of the notion of representation, will show that at some point belief cannot help but fit reality. And so they will hold that at that point questions as to what entitles us to adopt our beliefs cease to be appropriate.

The realist sees things differently. The import of the realist's talk about 'truth', 'correspondence', 'independent reality', etc., is precisely that the realist holds that for any belief there is a substantial question as to whether that belief gets reality right. And so the realist recognizes a need to justify our beliefs as accurate representations of reality. For the realist, the adoption of a belief calls for some assurance that it corresponds to how things are.

I think that here we have a real issue. In what follows I shall use the terms 'realist' and 'anti-realist' as implicitly defined in the last two paragraphs. So I take the dispute between realism and anti-realism to be a substantial disagreement as to how to *treat* our beliefs, not just a stylistic disagreement over what to *say* about them. Moreover, given that there is a substantial dispute here, I shall be siding with the realists. I shall be arguing that their views about how to treat beliefs are the right ones.

(A brief proleptic remark. Many readers, and this will include some who count themselves realists, will no doubt suspect that by committing myself to the thesis that beliefs require justification I must be backing a loser. Their suspicion would go as follows. To justify a belief is to show that it follows from other beliefs. So either I am going to be stuck with a vicious regress, or I will be forced to return to givens, to beliefs that somehow do not themselves require justification. But I shall not be taking it that a belief's being justified requires that it be shown to follow from other beliefs. So while I shall want justificatory epistemology, I shall not be returning to givens. But more of this later.)

2. *Two kinds of anti-realism*

I shall lead round to my defence of a realist justificatory epistemology by first discussing the opposition at some length. I characterized anti-realism as the view that at some point requests for justification ought to stop. At some point belief cannot help but get reality right (that is the reason for denying the 'independence' of reality) and so at that point further worries about the warrant for our beliefs become pointless.

Let me fill out the anti-realist position by considering what the anti-realist will say about relativism, about the fact that different people in different times and places adopt incompatible beliefs about the natural

world. The issue of 'relativism' tends to trigger much argumentative excitement. But I do not intend anything particularly contentious by the term. I mean simply the historical-anthropological fact that different people in different societies have adopted incompatible beliefs about the world. The implications of relativism are part of what is at issue in this paper, but I do not want to build any such implications into the term.

But, even in this innocuous sense, relativism raises difficulties for the anti-realist. What is the anti-realist to say about the existence of different peoples with incompatible beliefs? If 'reality' is just the picture presented by belief, are we to conclude that those with alternative beliefs live in a different world from us? This makes no good sense. Perhaps the anti-realist should drop talk of 'reality' altogether, including such dicta as 'reality is the picture presented by belief'. But a problem remains. Remember that modern anti-realists, unlike their instrumentalist and idealist predecessors, do not want to block belief. They intend to leave everything as it is (lest, being anti-realist across the board, they leave us with no beliefs at all). So the possibility of relativism raises at least this question: why *should* we believe our beliefs rather than those of other peoples? If alternative beliefs are indeed possible, surely it is indefensible philosophical complacency to continue endorsing our view of the world, without explaining why it is better than the alternatives.

Perhaps this complacent stance can be coherently maintained. But it is at best an uncomfortable position to end up in. And not surprisingly most anti-realists adopt a different tack. They deny that relativism is possible.

One can see why this is a natural response for the anti-realist. The anti-realist wants to be excused from the need to justify our beliefs on the grounds that at a certain point sceptical doubts cease to make sense, at a certain point belief cannot 'help but fit reality'. But if that goes for our beliefs, then presumably it goes for those of all belief-formers. And so, given that we are not going in for different realities for different people, it follows that different people just cannot have incompatible beliefs.

Anti-realists hold that *at some point* belief and reality cannot help but fit each other. This point, then, is the point at which they hold that relativism becomes impossible. But there are different kinds of anti-realism. One can classify them in terms of where they put the point at which the need for justification stops and relativism become impossible.

I want to distinguish two forms of anti-realism: a stronger form, which I shall call *anti-realism of current belief*, and a weaker form, which I shall call *anti-realism of method*.

Anti-realism of current belief puts the point at which justification stops right at the beginning. It maintains that just about all our beliefs—just about all anybody's beliefs—cannot help but be right. The idea that different people can have incompatible beliefs is simply incoherent.

It might seem surprising that anybody should defend such a view. But something very close to it is maintained by Donald Davidson in his 'On the

Very Idea of a Conceptual Scheme'. And Davidson's argument has been endorsed by Richard Rorty in *Philosophy and the Mirror of Nature* and elsewhere. (Indeed Rorty has described Davidson's argument as 'the transcendental argument to end all transcendental arguments'.)[3]

In essence Davidson's argument goes as follows. If a people are making claims about the world, their statements must have content. If their statements have content, it must in principle be possible that we should be able to specify what those contents are. That is, it must be possible for us to specify the truth-conditions of their various assertion-types. But how do we discover the truth-condition of a given alien assertion-type? Presumably by noting what circumstances characteristically obtain when they utter it. Which, if you think about it, implies that it cannot help but be the case that pretty much everything they assert must be true. (I oversimplify grossly. But I would defend this summary as faithful to the spirit of Davidson's argument.)

Davidson's position is encapsulated in the familiar 'principle of charity': a precondition of correct translation is that the translation represent everything (or at least as much as possible of what) the aliens say as true. It should be clear why this position supports a strong form of anti-realism. Davidson's ideas about representation imply that a people who make judgements cannot possibly have beliefs incompatible with ours. Or, to put it in a way that does rather better justice to the thrust of Davidson's argument, neither they nor we can possibly have beliefs incompatible with the world. And so, if Davidson were right, there would obviously be no point to epistemological anxieties as to whether our beliefs are accurate representations of reality.

However I suspect that few people nowadays would accept Davidson's position in the form just presented. The 'principle of charity' is much too strong. It is surely possible for a community to have a wide range of systematically false beliefs. It there is a general moral to be drawn from the analysis of representation, it will be at most the 'principle of humanity', the principle that the beliefs attributed to a community must be understandable ones for humans in the relevant circumstances to adopt.

Thus consider a translation which attributed to an alien people the belief that the earth is flat. In 'Thought and Talk'[4] Davidson suggests, heroically, that since the earth is not flat, this attribution ought to count against the translation. Perhaps we should translate the subject term in question as 'the-land-round-here' rather than 'the earth', or some such. But surely, *contra* Davidson, the falsity of the judgement in question does nothing in itself to

[3] R. Rorty, 'Transcendental Arguments, Self-Reference, and Pragmatism', in P. Bieri *et al.*, eds., *Transcendental Arguments and Science* (Dordrecht, D. Riedel, 1979), p. 78. Davidson's 'On the Very Idea of a Conceptual Scheme' is reprinted in his *Inquiries into Truth and Interpretation* (Oxford, Clarendon Press, 1984).

[4] *Inquiries into Truth and Interpretation*, p. 168.

undermine the translation. Provided we can explain why the people in question have the view (it is part of their cosmology, they have never seen ships coming over the horizon) it clearly should not count against the translation that the view is false. Translations can make people out to be mistaken, as long as they do not render them incomprehensible as human beings.

But now the 'principle of humanity' gives rise to a different kind of anti-realism, the kind I shall call 'anti-realism of method'. In shifting from charity to humanity we have allowed that it is possible for different people to have incompatible beliefs, and thus that there is room for a question as to whether *our* beliefs are true. But note that there is still no real question as to whether a people's beliefs will be explicable. On analysis there is no sense to the idea of a people believing (of their judgements being in principle interpretable) and yet it not being understandable that they should believe what they do. Suppose now we construe explicability as a matter of rationality. To explain why somebody believes p is to show that it was rational given their circumstances for them to form that belief. This then leads to the conclusion that, while different people might have incompatible beliefs, it is inevitable that they will conform to similar rational patterns of belief-formation. And, correspondingly, while anybody's beliefs might be false, there is no possibility of the standards by which they arrive at them being irrational.

This then is anti-realism of method. Anti-realism of method allows that beliefs might be false, and that therefore it makes sense to seek assurance that our beliefs are true. But it leaves no room for similar questions at the level of method. The methods by which we (and everybody else) arrive at beliefs cannot help but be rational, and so there is no need to raise questions about the warrant for adopting those methods.[5]

Rationality allows error at the level of belief because reasons are standardly non-conclusive. Even if we equate the *right* beliefs with those that would be validated by continued application of rational standards, it remains the case that in practice the beliefs that reason gives us might be wrong, if for no other reason than that our evidence is always less than complete.

One interesting question which arises at this point is how far the notion of 'rightness' which we get out of anti-realism of method can recapture the standard features of the familiar 'realist' notion of truth. Another interesting question is how far the anti-realist of method *wants* to recapture those features. But let us bypass these issues and consider instead the strength of the arguments for accepting anti-realism of method in the first place.

To do this we need to be more specific about the level of methods at issue when we talk about principles of 'rationality'. Suppose we take it, say, that the methods at issue include (i) the processes by which people form observational beliefs, (ii) the making of deductive inferences, and (iii) basic induc-

[5] Perhaps the clearest expression of this view is Johnathan Lear's version of Wittgenstein in 'Leaving the World Alone', *Journal of Philosophy* 79 (1982), pp. 382–403.

tive logic. The trouble then is that it seems fairly obvious that it is possible for different people to adopt different 'principles of rationality'. Surely different people can be better, or worse, or rather different from us, in respect of their deductive and inductive reasoning, or in respect of their observational dispositions, without this implying that they do not have beliefs at all? (Remember that it is what alien communities *do*, rather than what they *ought* to do, that is at immediate issue. The anti-realist of method wants to argue that it is *impossible* for people to adopt different standards.)

Perhaps this seems a bit quick. Isn't there *some* notion of 'rationality' such that alternative rationalities are impossible? Isn't there *some* sense of 'standards of rationality' such that the idea of a people having standards incompatible with ours undermines the possibility of our knowing what they believe, or indeed of there *being* something that they believe? But now the difficulty is that a notion of rationality that allows these questions to be answered in the affirmative is unlikely to do the job the anti-realist wants of it.

If all 'rational' means is 'ways of reasoning the transgression of which makes incoherent the idea that the people in question have real thoughts at all', then it is inevitable that all thinkers are going to be rational. But the anti-realist's transcendental argument needs more than that. In particular 'rational' needs in some sense to imply *justification*. Not only that, if the anti-realist wants to leave everything as it is, 'rationality' needs to show that our first-order beliefs, our theories of the world, are ones we are justified in holding, and in particular that they are more belief-worthy than the historical and anthropological competition (no doubt because we have more evidence). There does not seem to me any positive reason for supposing that the 'preconditions of thinking at all' are enough to do this rather demanding job. And I suspect that in the end much of the appeal of anti-realism of method derives from an equivocation between a weak (anormative) sense of 'rationality', in which it is analytic that all believers must conform to a common pattern of rationality, and a stronger (normative) sense, which would help with epistemological worries, but in which it is just not plausible that everybody has the same standards.

The real strength of anti-realism of method (and the reason, no doubt, that it has as many defenders as it does) is not the positive case for accepting it, but the difficulty of seeing what we can do when we reject it. Compare the rejection of anti-realism of method with the rejection of anti-realism of current belief. If we allow that our beliefs might be wrong, and wonder what will show us that they are not, it seems natural enough to appeal to authorized, rational principles of belief-evaluation to show us that our beliefs are right and the alternatives are wrong. But what are we to do if we allow that our standards might be wrong? Who is going to judge the judges?

So the real challenge raised by anti-realism of method is that of showing

how our accepted practices of belief-formation can possibly be evaluated. In the next section I attempt to answer this challenge.

3. *Naturalized realism*

I shall begin the attempt to get beyond anti-realism of method by concentrating on observational practices. At first sight observation might seem to be the stronghold of the anti-realist. Our observational judgements are the starting-point for our theorizing about the natural world. If we do not take our observations for granted, if we cannot be assured that they at least must be right, then how are we ever to get started in finding out about the natural world? One feels that the anti-realist had better be right about observations, for if we have to begin by justifying them, then we shall never get anywhere.

But, even so, it is not difficult to show that the anti-realist is wrong about observation. Different people do form observational judgements in incompatible ways. But at the same time it is (fortunately) perfectly possible to justify observational practices, to show that certain ways of forming observational judgements (our ways) are the right ways.

To illustrate how alternative observational practices are possible, and how they can be evaluated, let us turn to the history of science, and in particular to examples of the 'theory-dependence of observation'. These are cases where theoretical developments lead to changes in what were hitherto naïvely accepted as reliable ways of forming observational judgements. Thus, to take a familiar example, prior to Copernicus it was supposed that one could tell whether a body was falling straight down by looking and seeing. (This supposition gave rise to the 'tower argument': if the earth were moving, then a body falling straight down would land some distance to the west of where it was dropped; but such bodies land directly below where they are dropped; so the earth cannot be moving.) But then the adoption of the heliocentric view of the solar system showed us that the bodies we previously took to be falling straight down are in fact moving sideways at a great speed with respect to the fixed stars (so as to be able to 'keep up' with the moving earth).

I take such examples to illustrate how alternative observational practices are possible. Where Ptolemaists non-inferentially judged that dropped bodies were falling straight down, Copernicans arrived at different beliefs. But such examples do not of course themselves explain how alternative practices might be evaluated. Indeed the implicit threat of relativism makes many people suspicious of the 'theory-dependence of observation'. Examples like the above seem to suggest that when the adherents of a theory (the Copernican theory, say) run into difficulty with certain observations (bodies *falling straight down* do not land to the west), they are free to opt for some alternative practice with respect to those observations which avoids the

difficulty. But if that were all there were to it, then clearly the defenders of any theory whatsoever could tailor their observational practices to fit their theory, and general theoretical relativism would not only be possible, but absolutely rampant.

But that is not all there is to it. There is a natural enough way to understand the evaluation of observational practices. The key is to recognize that we are not (as the anti-realist picture seems to suggest) locked inside our heads, able to note what observations come in, but with no real idea of where they come from. On the contrary, we have in general fairly detailed views about the kind of processes that give rise to our observational judgements, and this gives us a way of evaluating our perceptual habits which does not simply collapse into adopting those habits which do not contradict our favoured theories.

In effect we are able to think of ourselves as 'observational instruments', as beings whose interaction with the natural world gives rise to perceptual beliefs, or 'readings'. Our understanding of how we work as instruments allows us to raise the question of how reliable we are at generating the relevant readings. Thus we can recognize that an instrument that detects the direction of fall of moving objects by comparing their motion with that of the background is not going to be very reliable in circumstances where the background itself is moving. If such reflection shows us that the processes giving rise to our observational judgements are unreliable, then we have every reason (with Copernicus and Galileo) to change our observational habits. If, on the other hand, our view of the world and how we react to it gives us no reason to distrust the observational judgements, then we would do better to rethink any theories whose predictions are contradicted by those judgements.

This then is my model of the evaluation of perceptual practice, and more generally of how to judge the judges. But of course there is more to be said. Let us start by considering two immediately obvious objections.

Firstly, there is the point that in evaluating our reliability as 'instruments' we will be taking for granted certain aspects of our current theory of the world. Thus in the 'tower argument' example I was implicitly presupposing that we were entitled to certain substantial empirical assumptions about eyes, light rays, the geometry of space, etc. Clearly we would not have any hold on our reliability as instruments unless we could presuppose something about how such instruments work.

These presuppositions do not involve us in the immediate circularity of simply adopting the Copernican heliocentric view and then deciding that all observations that contradict it *must* be unreliable. The relevant assumptions about eyes etc. are pretty much independent of astronomical theory. But, still, they are worrying. After all, if I am proposing a general account of how we can assure ourselves that our beliefs about the natural world are true, surely I am precluded from presupposing anything. Surely assumptions

about how our sense organs work, the theoretical claims of perceptual psychology, are as much part of what I am trying to justify as anything else?

Clearly more needs to be said on this topic. But it will be helpful to shelve it till the next section, and turn instead to a rather different worry about my strategy.

It might seem that, in addition to presupposing generalizations from perceptual psychology, my strategy also needs to presuppose givens. Am I not suggesting something along the following lines? When one is unsure about an observation, one should look into one's mind to note how things appear (it *looks* as if it is falling straight down); one should then appeal to the generalizations of perceptual psychology which say how it must be 'out there' for things to be appearing as they are; and one should then infer from these two premisses how it in fact is 'out there' (the body in question is actually moving rapidly to the east). But then am I not taking for granted our ability to know how things *seem* to us?

It is important that this is not my model of the evaluation of observation. If it were, then I would indeed be presupposing givens. And, as I said at the beginning, my concern is to see what happens without them.

I attach no special role to our ability to introspect our own mental states. In so far as we have such an ability (as of course we do), it is another mode of observation (a sixth sense—'insight'),[6] which is fallible, and which is itself susceptible of the kind of justification I am currently explaining.

There are things to be said about the relation between our visual phenomenology, our introspective judgements of how things seem to us, and our further judgements of how they are. But for present purposes this merely complicates the picture. Let us simplify by thinking of observational habits of any kind (both introspective and others) as *pre-conscious* processes, the *outputs* of which are our having (conscious) beliefs as to how things are (both in our minds and elsewhere). We should think of them as entirely dumb habits. (The phenomenon of 'blind sight' embodies the simplification I want—we just *end up* believing that bodies are falling straight down.)

My schema for the evaluation of observation does not require us to *use* the observational habit at issue. Rather we need to 'stand back' from it and reflect on its reliability, in just the way we would reflect on the reliability of a real instrument.

An important corollary follows. The inference we are supposed to draw from the appeal to perceptual psychology is not a move *within* our corpus of extant beliefs. (Such as: 'Since (i) it appears to fall straight down, and since (ii) bodies which appear so are moving to the east, then (iii) it is moving to the east.') Rather it is the *practical* inference that, since we have an unreliable belief-forming habit, we should *stop* doing it. The conclusion is that we should, so to speak, rebuild ourselves as instruments. In general this will not

[6] The term is due to D. H. Mellor. See his 'Conscious Belief', *Proceedings of the Aristotelian Society* 78 (1977/8), pp. 87–101.

be something we can do at will. It will require retraining and discipline. But, as Pascal did after reflecting on his wager, one can perfectly well set about getting oneself to believe differently. (In our case it would be a matter of monitoring ourselves and correcting the judgements we are naturally inclined to make. I take it that enough such correcting will stop one being so naturally inclined. It might also alter one's visual phenomenology—it might lead to our 'seeing things differently', as when people 'learn to see' X-ray tubes, electron-positron pairs, etc. But, again, that is not the crucial point—what matters is that we can change our natural inclinations as to what observational *beliefs* we *end up* with in given situations.)

So the model is this. We embody a set of belief-forming habits. These are unconscious processes which deliver beliefs to us. But we can reflect, via perceptual psychology and other relevant bits of science, on what is going on in us when these processes operate, and we can thereby reach conclusions on whether these processes are unreliable or not. And if we decide that such a process is unreliable, we should reconstruct ourselves as belief-formers.

It might be helpful to relate this approach to the 'reliability theory of knowledge'. According to this theory a true belief counts as knowledge if it is produced by a reliable process (and its truth is not accidental relative to the facts that make the process reliable). This reliability approach to knowledge is in many ways intuitively appealing. But many people find it difficult to see how it can be a contribution to *epistemology*. Whether a belief is produced by a reliable process or not depends on circumstances outside the consciousness of, and possibly entirely unbeknownst to, the believer. So it seems as if the reliability approach must lack normative import, must fail to contribute anything to the question of what believers *ought* to believe.

I take my argument in this paper to show how the reliability theory of knowledge can be relevant to normative issues. Suppose we read the reliability theory as saying believers ought to believe beliefs produced by reliable processes. We should not, however, interpret this as recommending that at the moment of believing, so to speak, believers ought somehow to assure themselves that their belief is issuing from a reliable process. That recommendation would be either empty or regressive. Rather the reliability theory should be understood simply as saying that people ought to *embody* reliable processes (and *thereby* get to be believers whose beliefs are produced by reliable processes). In so far as we do have views about what belief-forming processes are reliable, and about what kinds of retraining (if necessary) will bring it about that the processes we engage in are so reliable, there is no reason to suppose that this recommendation is any more problematic than any other prescription for action. And, more generally, once we see epistemological recommendations as suggesting that we should *act* in order to adjust the way we work as belief-formers, rather than as authorizing certain moves within our existing corpus of beliefs, it becomes clear how the reliability approach to knowledge can have normative import.

Let me take it that I have shown we ought to resist anti-realism of method in the case of observation—that is, that I have shown that alternatives to any given observational practice are possible, and moreover that we can justify the observational practices that we actually adopt.

The next task is to generalize the strategy from our observational practices to our ampliative methods for forming beliefs, that is, to our deductive and inductive practices. In these cases also I want to show that alternatives are possible, and that therefore we need to justify our adopting the practices we actually do adopt.

Let me take deduction first. Here I simply appeal to the writings of Michael Dummett, and in particular to his 'The Justification of Deduction'.[7] Dummett has shown how the familiar proofs of soundness for various principles of deductive inference can be taken seriously as justifications of those principles. Given that one wants to get from truths to truths, one will do well to use those principles, for we can show, in terms of our understanding of how the truth-conditions of complex sentences depend on the semantic values of their components, that those principles will lead you reliably (infallibly) from true premisses to true conclusions.

It is true that one of Dummett's main reasons for discussing this issue is to contrast the way classical logic is justified by 'realist' semantics with the way that intuitionistic logic can be justified by 'anti-realist' semantics. But I am not interested in that. As will become clear shortly, I think there are good independent reasons for adopting what Dummett calls 'realist' semantics, at least in connection with discourse about the natural world. All I want to borrow from Dummett is the point that, given such a 'realist' semantics of sentential structure, we can *show* what is *wrong* with the people who adopt invalid habits of reasoning. (There are plenty of them—*pace* the anti-realist of method.) And correlatively we can *justify* the practice of people who do adopt valid habits of reasoning.

What I have to say about induction will also be brief. For here I have only promissory notes to offer. The task is to assure ourselves that our methods of inductive inference are reliable. (Not infallibly so, but at least probabilistically reliable.) If we cannot do this then we will have no way of showing that our beliefs are better than those derived from alternative inductive methods, and so no uncomplacent, honourable way of continuing to uphold our own beliefs. Unfortunately (indeed scandalously) we cannot currently show that our inductive methods are reliable.

It is worth being specific about what is scandalous here. The scandal is not that we are unable to convince somebody who has no truck with inductive inferences, or any conclusions derived therefrom, that they ought to start making them. That seems a hopeless task (perhaps a demonstrably hopeless one). Rather the scandal is that we cannot even explain to ourselves, who do

[7] 'The Justification of Deduction', in M. Dummett, *Truth and Other Enigmas* (London, Duckworth, 1978).

make inductive inferences, and enjoy their products, why these inferences are a good thing. Indeed we do not even have any good ideas about how to start on this lesser task. We are not even clear about the content of inductive conclusions—are they Humean regularities, or are they relationships between universals?—let alone about a general metaphysics that will show why such conclusions are likely to be true when there is inductive evidence for them.

Perhaps new work on universals and laws of nature will enable us to make some headway in this area. (And then we will be in a position to judge whether our inductive methods really are a good thing after all. Perhaps they are not. Perhaps we should adopt rather different methods—be more cautious, project rather different predicates . . .) But, as I said, all this is promissory, and I have nothing definite to offer here.

Let me press on. I have argued that for observation and deduction we can, and that for induction we ought to be able to, evaluate different belief-forming practices, and show why we should engage in ours. I take this to discredit anti-realism of method, and to indicate that we ought instead to adopt a realist attitude to our belief-forming practices.

It might be unclear what is so distinctively *realist* about the position I am advocating. Maybe I can show that observation, deduction, and induction, are in need of, and ought to be susceptible of, justification. But what is so realist about that? It is not as if the idea of justification is in itself alien to anti-realism. After all, anti-realism of method was brought in to replace anti-realism of current belief precisely because it was clear that consensus was not enough, and that beliefs needed justification by standards of rationality. I have now shown that we should admit even more justification than that, that standards of rationality themselves need justifying. But why should this refute anti-realism as such, as opposed to some specific version of 'anti-realism of method'?

The distinctively realist aspect of my position can be brought out by unpacking the sense in which belief-forming practices need to be justified as *reliable*. 'Reliable' here means reliable for truth. And truth in this context needs to be thought of in terms of a belief corresponding to how things are, as opposed to a belief having been arrived at in the right way. For the very idea of *evaluating* belief-forming methods for reliability requires a notion of truth which distinguishes clearly between the truth of a belief and its intellectual provenance.[8]

These last remarks indicate something of a lacuna in the argument of this paper, namely, the absence of any analysis of the notion of truth as correspondence. It is true that there is a sense in which the analysis of

[8] Perhaps one could have *some* provenance-transcendent notion of truth which fell short of a realist correspondence notion. Cf., for example, S. Blackburn, *Spreading the Word* (Oxford, Clarendon Press, 1984), Chs. 6 and 7. But even so I would argue that my notion of *evaluating* belief-forming methods *for reliability* demands a realist notion of truth.

representation is not as pressing a problem for the realist as it is for the anti-realist. Most people find the idea of correspondence intuitively natural enough, even if difficult to articulate. The urgent problem for the realist is to show how we can possibly manage to justify our judgements, *given* that we accept the idea of correspondence, and therewith the possbility of error. And this is the problem I have been trying to get to grips with in this section. (The realist and anti-realist argumentative strategies contrast sharply here. Anti-realists do need to start with an analysis of representation, for their argumentative strategy is precisely to block difficulties about justification by defending some non-intuitive analysis of representation.)

Still, even if these considerations of argumentative strategy account for the lacuna, they hardly excuse it. The idea of correspondence might be natural, but it is scarcely philosophically unproblematic. Certainly I owe some explication of the correspondence notion of truth. But this is not the place to tell that story. Here I can only refer the reader to another paper[9] in which I do address the question of representation as such. In that paper I argue that the primary representers are functional states of the brain, and that the representational powers of such states should be analysed *teleologically*, in terms of the fact that their biological purpose is to occur in the presence of the circumstances they represent.

Let me conclude this section with some brief comments about morality and mathematics. In these areas I think that we do lack a notion of truth as correspondence. We do not naturally think of correctness, in moral or mathematical judgement, as a matter of a correspondence between a belief and some further state of affairs. And, accordingly, there is no sensible project of justifying such judgements by showing that they issue from belief-forming practices which are reliable for producing beliefs that so correspond. The most appropriate model for correctness of moral and mathematical judgement does seem to be one of intellectual provenance, of beliefs being reachable by certain authorized methods, which of course blocks any question of *showing* that those methods produce correct beliefs. That is, anti-realism of method does seem to be the appropriate stance to adopt on matters of mathematics and morality. But of course this issue is complicated by both Platonist and reductionist accounts of the contents of mathematical and moral judgements.

4. *Is naturalized realism enough?*

Let me return to the question of relativism. Anti-realists needed to deny the possibility of relativism: anti-realists of current belief needed to deny that different beliefs are possible, and anti-realists of method needed to deny that different methods are possible. Realists, on the other hand, are happy to allow that different people can have different beliefs, and even different

[9] 'Representation and Explanation', *Philosophy of Science*, forthcoming.

standards of rationality. Their task is to show that certain standards (ours), and the beliefs they give rise to, are better than the historical and anthropological alternatives.

The question I now want to address is—will the realist succeed in this task? We have already seen how the realist can evaluate standards of rationality for their reliability. But we still need to consider whether such reliability-evaluations have the power to vindicate our own ways of thinking over the historical and anthropological alternatives.

Let us begin here by considering what the realist will say to the sceptical 'meta-induction' from the history of science. This argument, which threatens any attempt to vindicate our current beliefs, goes as follows. Past theories have characteristically turned out to be false. So no doubt our theories will turn out to be false too. So surely we ought not to believe our theories after all.

The first thing for the realist to say to this argument is that, in so far as past theories have characteristically turned out to be false, they could not have been justified, in the sense of having been produced by reliable belief-forming methods. It is possible for a reliable belief-forming method *sometimes* to give rise to a false theory. Reliability does not have to amount to infallibility.[10] But a method that *characteristically* gave rise to false theories would thereby be shown to be unreliable.

One can usefully distinguish three different ways of defusing the meta-induction, of responding to the fact that certain past claims, such as, say, Galileo's law of circular inertia, or the phlogiston theory of combustion, have turned out to be false. In the first place we could conclude that certain apparently trustworthy methods of belief-formation are less reliable than hitherto supposed. Thus we might, for instance, become more cautious about the extent to which we can inductively extrapolate beyond the limits of experimental test. Secondly, we might be able to show that the claims in question did not even have prima facie trustworthy methods behind them. That is, we might be able to show that certain putatively 'scientific' claims were generated by religious or other such 'external' causes. And then, thirdly, there is always the option of concluding that the scientist in question was just unlucky—that although the scientist's methods were genuinely reliable and actually operative, this was just one of those cases where the right method gives the wrong results.

[10] If we recognize that a certain method is less than fully reliable, oughtn't we to attach a degree of belief less than one to its products? But if we do that, won't there then be a sense in which all methods will end up reliable, in that the degree of belief attached to their outcomes will be proportional to the frequency with which they produce true results? This issue calls for far fuller discussion than can be given here. A first thought is that for beliefs about non-chance matters we should aim for methods that will warrant *high* degrees of belief. But this then runs into conflict with the desire that our methods should be informative, that they should deliver results as often as possible, and should not buy a high success-rate just by staying silent when answers are not obvious. For this play-off between reliability and informativeness, see H. Field, 'Realism and Relativism', *Journal of Philosophy* 79 (1982), pp. 553–67.

However, this initial realist response to the meta-induction is obviously bypassing an important issue. Certainly, given the standpoint of our present theories, we cannot both accept that a certain method characteristically produces false beliefs, and that it is a reliable method of belief-formation. That is, our present theories, when applied to the history of science, have certain natural implications as to which belief-forming methods are reliable. But, still, won't different people, with different theories, be led by similar reasoning to *different* views about what methods are reliable?

This is a special case of the general worry I left up in the air in the middle of the last section. I pointed out there that what gets accepted as a reliable method of belief-formation depends upon certain parts of currently accepted science. There I had in mind certain parts of perceptual psychology, and associated subjects like optics, etc. Now we see that the same worry arises in connection with views about the history of science and how past theories stand with respect to current accepted theory.

The worry, to spell it out, is that people who have different views of the world will regard different belief-forming processes as reliable. And so we seem to face the possibility that their beliefs will seem to them as justified as ours do to us. Their beliefs will be generated by belief-forming processes that *they* regard as reliable. No doubt we, from the standpoint of our theories, will regard their belief-forming processes as unreliable, and their beliefs as therefore unjustified. But while this gives us something sensible to say *about* them, it does not give us anything sensible to say *to* them. If we are really to vindicate our views over the historical and anthropological alternatives, surely we ought to be able to say something *to* our intellectual competitors which is genuinely persuasive, and does not merely beg the question at issue.

At this point I think the realist ought to appeal to a general epistemological principle implicit in the argument of the last section. In effect I there defended the following practical recommendation: *one's beliefs ought to be generated by methods that one can show (assuming current science) to be reliable.* Let us call this the 'meta-principle'. I want to suggest that the meta-principle gives the realist a non-question-begging way of saying something persuasive to defenders of other belief systems.

Many people seem to assume that the meta-principle lacks bite, that it is an empty dictum that leaves everything as it was.[11] But this is a mistake. Just because a people *engage* in a certain belief-forming process, it does not automatically follow that their current science will be such as on reflection to *justify* that process. Just because one has a certain habit for coming to believe, say, that bodies are red, or square, or falling straight down, it does not follow that one will have further general assumptions, in perceptual psychology etc., that will show this habit to be reliable. In general the beliefs

[11] Thus see Putnam, *Realism and Reason* (Cambridge, Cambridge University Press, 1983), pp. 233–4.

that are the *output* of a certain belief-forming process (that a particular body is red, or square, or falling straight down) will be quite different from the beliefs required to *justify* that process (general assumptions about the connection between the relevant 'external' circumstances and our arriving at such particular conclusions).

So the meta-principle is not an empty dictum. It leads people to change their belief-forming methods and therewith their beliefs. Moreover, it is, in itself, independent of any commitment to our first-order theories about the natural world. So perhaps the realist can use it to show, in a properly impartial manner, why our views are superior to the alternatives. The strategy would be to show that by their *own* lights the adherents of the relativist alternatives are in trouble, in that *their* theories are incapable of showing that the processes generating their beliefs are reliable. And then the hope would be that by continued application of the meta-principle the proponents of alternative views of the natural world could be brought round to our way of thinking, could be led to converge on our view of the world. Indeed it is arguable that this is just what has happened to a particular group of thinkers—namely, our ancestors in our own intellectual tradition.

Perhaps it is worth digressing briefly at this point, in order to say a bit more about the significance of theoretical *convergence*. One might be tempted to see my overall argument in the following terms:

(1) Particular *beliefs* are justified in so far as they emerge from authorized belief-forming *processes*.
(2) Such *processes* can be justified by appeal to our *current scientific* view of how we work as belief-formers.
(3) *Current science* is then justified in so far as the continued application of the *meta-principle* has moved us from previous views to it.
(4) And then the *meta-principle* will be justified in so far as we can show that, from a God's-eye point of view, it will eventually lead us and current science to converge on the *Truth with a capital T*.

This is not my line of thought. My interest in convergence is nothing to do with any need to justify the meta-principle. I do not think the meta-principle needs justifying. The meta-principle says: adopt those belief-forming practices you believe to be reliable, and shun those that you do not. It seems to me that one would be a fool to do otherwise. My concern is rather to justify current science. Far from wanting to appeal to *eventual* convergence to justify the meta-principle, I want to appeal to the meta-principle in the face of relativist alternatives to back up the claim that we have *already* converged—on current science.

Some readers might feel uneasy here about this unqualified desire to validate current science. For it is fairly deeply built into much modern philosophy of science that nothing of what we currently believe is at all likely to be the last word on anything, that indeed the current science of any

period, past, present, or future, is going to be superseded by a new alternative, which will in turn be superseded . . .

I am against this 'jam tomorrow' theory of science. This is not, in the present context of argument, because I want to assume parts of current science in order to reflect on the reliability of belief-forming processes, but simply because I accept the principle that thinking a certain belief is incorrect ought to (indeed inevitably will) stop one believing it. I have taken this to be a ground rule of the whole debate, applicable to the anti-realist as much as to the realist. The anti-realist, remember, did not want us to stop believing so much as to adopt a different attitude towards truth and justification, etc. The only person who does want us to stop believing, and so would be happy with the idea that our current beliefs are in error, is the sceptic, who thinks we ought indeed to have no beliefs. But I suspect that few of the modern advocates of the 'jam tomorrow' view of science would be happy with the consequence that we ought to have no beliefs.

The 'jam tomorrow' view seems to me to have been foisted on philosophers of science by an unholy alliance of Popperian irrationalism and closet Cartesianism. Popper at least is consistent—he does not care about belief and as far as justification goes would be as happy that we believe anything as that we should believe modern science. But it is unlikely that many others would have been prepared to follow him in this nihilistic line unless they had at the back of their minds the thought that there was something at least (knowledge of directly observable matters, knowledge of our own mental states) about which we can have fully justified beliefs. The moral is that those of us who are not prepared to take givens for granted, but would like to be authorized in at least some beliefs, need to resist the wholesale scepticism implicit in the idea that every theory will in time turn out to be incorrect.

Let me return to the main line of argument. I suggested that the meta-principle might provide an impartial basis for persuading our intellectual opponents to our view of the natural world. The idea was that *they* would be led to admit that, by *their* lights, their belief-forming methods were unreliable. But some readers may remain unconvinced. Even if the meta-principle has *some* bite, what guarantee is there that it will have *enough* bite to persuade all possible opponents? What is to guarantee that our intellectual opponents will not sometimes conclude on reflection that, by their lights, their belief-forming methods *are* reliable?

Thus consider once more the 'tower argument' case. Of course, from our neo-Copernican point of view, the observational habit which uses movement with respect to the earth to judge direction of fall is unreliable and should therefore be abandoned, for we Copernicans believe the earth is moving. But consider the situation from the Ptolemaists' point of view. Reflection gives them no reason to distrust the habit, for they do not think the earth is moving.

I accept that if we focus on this particular aspect of the difference between the Ptolemaists and their opponents, the meta-principle leads to a stand-off. But this leaves it open that the Ptolemaists might be in trouble further down the line. What, for instance, about their belief that the earth is not moving? Have they taken pains to ensure that the process behind that belief is reliable? And then what about the beliefs they appealed to in deciding that . . .? I would suggest that if we continue in this way it will turn out that at some point the Ptolemaists will be forced to admit that, by their own lights, certain of their beliefs are unjustified.

It might seem that this is a game nobody can win. Will not continual demands for justification simply lead to a bottomless regress? But this does not follow. Certainly there is an element of boot-strappery about the appeal to the meta-principle. But remember the point that the meta-principle gives rise to *practical* conclusions, not theoretical ones. This means that the regress can well stop. Suppose, for the sake of illustration, that a certain community embodies a finite number of belief-forming processes. They can then, so to speak, check through those processes, in the light of their beliefs, to see if they are all reliable. If they conclude that some of their habits are not reliable, then they would be rational to change their habits. And then they would be led to a new set of beliefs, and would have to check again. But if, at any stage, they conclude, in the light of their beliefs, that their habits *are* all reliable, they can stop—at least until new evidence comes along which disrupts the stability. The crucial point, as before, is that the aim is not to have beliefs which are justified by other beliefs (which are justified . . .), but simply to *embody* processes which one believes to be reliable.

Still, even if there is no regress, what about the boot-strappery? Isn't that worrying in itself? To come to the point, what have I said to rule out the possibility of another community who are quite as well 'strapped-up' as we are, who have also checked through *all* their belief-forming methods in the light of *all* their beliefs and have found, to the same extent as we have, that the latter indicate the former to be reliable? For instance, might not the Aristotelian-Ptolemaic world system have contained a comprehensive account of perception, the history of thought, man's place in nature, etc., which made their overall belief system just as stable, and apparently well-founded, as ours?

It is not at all clear to me that the realist needs to say anything further at this stage. Why should it always be possible to say something persuasive *to* alternative intellectual communities, as opposed to something explanatory *about* them? Why should we suppose that the human mind as such (in combination with the world, and the meta-principle, and enough time) will suffice to lead any thinkers, whatever their starting point, to the truth?

But at the same time there does remain something very worrying about the idea of a community with different beliefs, but as well 'strapped-up' as us. We would, it is true, be able to *account* for their having different beliefs,

in terms of their adopting (what we take to be) unreliable processes. But they would be able to do the same to us. Shouldn't that in itself somehow undermine our faith in our own beliefs?

But suppose the realist does recognize this worry, and does concede that the possibility of an equally well 'strapped-up' community would be a good argument for scepticism. There is still room for the realist to deny that this putative possibility is a real one.

We need to be careful here. We need to distinguish the possibility that there could actually be another equally well strapped-up community, from two closely related possibilities.

In the first place, the realist does of course need to admit that our current beliefs might be false. After all, our current view of the world only implies that our belief-forming methods are reliable, not that they are infallible. So, even by our current lights, it is perfectly *possible* that much (most, all) of what we believe might be wrong. And if that were the case, then of course there would be a different but equally well strapped-up view of the world—namely, the one that got things right.

But this is not the worrying possibility. That we *might* be wrong does not force us to conclude that we *are*.[12] And, correspondingly, that the world might, epistemologically speaking, be different from what we take it to be, does not force us to accept that (given the way the world is) a community with different beliefs to ours could be as well strapped-up as we are.

Secondly, the realist should be prepared to admit that (as the world is) certain questions will be evidentially inaccessible to human beings, in the sense that whenever an answer T1 has a certain amount of evidential support, there will be another incompatible answer T2 which is equally well supported.

This is not a worrying possibility either. For the right response, in so far as we do recognize such underdetermination of theories by evidence, is simply to withhold belief. After all, any principle of ampliative inference which yielded T1 would necessarily be an unreliable belief-forming method, for by hypothesis it would also yield T2, and T1 and T2 cannot both be true.

It is true that, if we accept that there are underdetermined alternatives to a *great deal* of what we believe, then we will be in trouble, for we will then have to give up a lot of our beliefs. But I do not see any reason to suppose that underdetermination is that rife. It seems to me that most of the arguments for chronic underdetermination trade on an equivocation between 'being equally consistent with the evidence' and 'being equally well supported by it'. One can allow that, for any answer T1 to some question, there is another answer T2 (say, one which multiplies unnecessary entities), which is also *consistent* with the data. But only somebody with Popperian

[12] Exactly how realists should accommodate the possibility of error without giving in to scepticism depends on how they deal with the issues raised in note 9.

prejudices about inductive support will immediately conclude that such a T2 will be equally *well supported* by the data.

So we should allow that we might be wrong in believing what we believe. And we should allow that we should withhold judgement on certain questions. But having allowed that much, is there any remaining reason for admitting the possibility of an 'equally well strapped-up' community? Once we put to one side the thought that the world might be different from what we take it to be, and once we leave undecidable questions out of it, what *arguments* are there for the possibility of such an alternative community? Realists are forced, by their own beliefs about the world, and in particular by their beliefs about how we find out about the world, to concede the possibility of error, and of localized ignorance. But what, in their view of the world, forces them to concede that a community committed to beliefs incompatible with ours could be as epistemologically well-placed as we are?

There is of course a question of argumentative onus here. The realist may have succeeded in resisting certain arguments for the possibility of an equally well strapped-up community. But that is scarcely a positive demonstration of the impossibility of such a community. Still, I feel that it is the sceptic who bears the argumentative obligation here. We should not be unduly swayed by the force of our modern tradition. Descartes demonstrated that, if one is after certainty, then there is a sceptical problem about getting beyond the given. But we, I take it, neither have the given, nor seek after certainty. If modern-day sceptics are to worry the contemporary realist, they need new arguments, not mere appeals to Cartesian authority.

Let me conclude by returning to something I said earlier about the meta-principle. Namely, that it did not stand in need of justification. The meta-principle, remember, tells one to adopt those belief-forming processes one believes to be reliable and shun those one believes to be unreliable. And, as I said, it seems to me that one would be a fool to do otherwise.

But doesn't this create an opening for a super-sophisticated Mk III anti-realist? Anti-realists hold that in some sense reality cannot help but fit thought. Which cashed out as the view that at some point justification had to stop, that at some point one should stop asking why a certain way of thinking should be supposed to get reality right.

But haven't I now allowed, indeed insisted, that the meta-principle is non-optional, that it would be inappropriate to try to justify the meta-principle by showing that it is well suited to getting us to Truth with a capital T? So doesn't this make me an anti-realist of meta-method?

Perhaps. I do indeed think the meta-principle is non-optional. And if that makes me an anti-realist, then so be it. But this would be a cheap victory for the opponents of realism. By now an awful lot of anti-realism has fallen by the wayside. *Contra* Davidson and the anti-realists of current belief, our beliefs need justifying. Any, or indeed all, of the beliefs we currently accept might be in error, and our acceptance of them is answerable to their having

been produced by authorized methods. *Contra* the anti-realists of method, standards of rationality need justifying. Any, or indeed all, of the habits of thought we engage in might be unreliable, and our adopting them calls for assurance that they are indeed good for producing true beliefs. True, the meta-principle does not need justifying. But consider what the meta-principle *says*: we should only have beliefs which are generated by processes which we can show to be suitable for producing beliefs that correspond to reality. That is realism, as I characterized it. Somebody who now wants to turn round and say that this is just a super-sophisticated form of anti-realism, ought in all honesty to have objected at the beginning, when I first explained what I was taking to be the difference between realism and anti-realism. And if they did not object then, they ought now to allow that there are good reasons for being a realist.

Department of History and Philosophy of Science DAVID PAPINEAU
University of Cambridge
Free School Lane
Cambridge CB2 3RH