5

Agrippa's Trilemma

Philosophical scepticism

Philosophical scepticism needs to be distinguished from other outlooks often thought of as 'sceptical'. The essential feature of philosophical scepticism is a *general view* about human knowledge: in the broadest terms, the view that knowledge is *impossible*. Let us call this view *theoretical scepticism*.

In ordinary parlance, scepticism is often associated with an attitude of incredulity. The sceptic is someone who practises suspension of judgement. If he is uncomfortable with this inability to make up his mind, he will suffer from doubt or uncertainty, often regarded as the sceptic's characteristic state of mind. People have also been called sceptics because they have been suspected of not just doubting but actually *denying* important truths: there have been times when the terms 'sceptic' and 'atheist' were more or less interchangeable. These various forms of incredulity are versions of *practical scepticism*.

Philosophical scepticism may involve an element of practical scepticism, but need not; and even when it does, the theoretical and practical components may not be coextensive. Indeed, one reason for being intrigued by philosophical scepticism is that, if the sceptic is right, our urge to believe far outstrips our capacity for knowledge.

Because it involves theoretical scepticism in an essential way, philosophical scepticism must be argumentatively based. Not that the mere existence of sceptical arguments amounts to a problem. One can argue—one way or another—for just about anything, but one's arguments might be unappealing. They might involve logical errors or depend on premises we see no reason to concede. However, the best sceptical arguments are not like that. Not only do they display no obvious logical flaws, they seem to involve only the simplest and most mundane considerations. They appear to be highly 'natural' or 'intuitive', to be arguments that almost any reflective person can understand and be moved by.

This apparent 'intuitiveness' is crucial to scepticism's being such an intractable problem. If sceptical arguments were obviously dependent on arcane theoretical ideas, there would be a quick way of dealing with them. We could

take the fact that certain ideas lead to scepticism as itself a decisive criticism of those ideas. Maybe this is the right approach in the end. But if so, it will take some showing.

Of course, not even the existence of powerful and intuitive sceptical arguments would present us with a problem if we were inclined to welcome their conclusions. And there are forms of scepticism that a lot of us find congenial. Thus, according to one common definition, scepticism is the opposite of dogmatism. If dogmatists are equated with theorists of any stripe, an equation the ancient sceptics were inclined to make, this is perhaps not such a benign outlook. But if we understand 'dogmatism' in our current way, as indicating a blind adherence to a fixed set of beliefs, then some degree of scepticism is an essential component of any rational outlook.

Another form of scepticism, equally benign, is indicated by the contrast between scepticism and credulity. To be a sceptic, in this sense of 'sceptic', is to take a tough-minded attitude towards standards of evidence and, as a result, to find oneself dubious about various commonly accepted claims. So, for example, the sceptically minded person is unlikely to believe, on the basis of reports in the tabloid press, that hundreds of people are abducted every year by visitors from outer space. If this were all it meant to be 'sceptical', it would be a good thing if some people were more sceptical than they are. Scepticism as either open- or tough-mindedness might be problematic at certain times and places: in Europe in the Middle Ages, in what was until recently the Communist bloc, in parts of the Islamic world, or in fundamentalist colleges. But it will not seem threatening to most readers of this book.

Two features of philosophical scepticism make it very different from such everyday 'sceptical' outlooks.

The first is its unusually wide scope. Philosophical scepticism is highly general: the sceptic wonders whether we know anything whatsoever, or anything whatsoever in certain very broad domains. This is why philosophical scepticism is formulated in terms of the *possibility* of knowledge. The philosophical sceptic does not claim merely that we *in fact* know a good deal less than we like to think (which is probably true), but rather that aspiring to knowledge is somehow *inherently* problematic. It isn't just a question of paying closer attention or working longer hours.

The second feature is its strength or depth. The most interesting sceptical arguments imply *radical* scepticism, the thesis that we never have the slightest justification for believing one thing rather than another. The important point about radical scepticism is that it is formulated in terms of justification. Knowledge is impossible because we cannot even get to the point of having justified beliefs, never mind beliefs that are well-enough justified to count as knowledge strictly so-called.

Radical scepticism contrasts with 'high-standards' or 'knowledge-specific' forms of scepticism. Many philosophers have argued, plausibly, that knowledge

demands more than just some degree of justification. Some have even argued that knowledge requires certainty, in some sense of 'certainty'. Accordingly, it is possible to hold that, while we may have justified true belief, our justification falls short of the demands of knowledge. Scepticism that concedes the possibility of justification, but denies the possibility of knowledge, is what I mean by scepticism that is knowledge-specific.

The distinction between radical and knowledge-specific scepticism becomes especially clear in the light of Gettier's problem. Suppose that in the light of this problem we decide that knowledge demands indefeasibly justified true belief. The sceptic's claim that knowledge is impossible could mean that no belief is ever *indefeasibly* justified. This would be to argue for a non-radical form of scepticism, since the possibility of justified belief would still be left open. Although scepticism of this type would not be wholly without interest, it would not be all that threatening, since many important epistemological distinctions would remain unscathed. (This is why defeasibility analyses of knowledge raise the question of the value of 'knowledge strictly so-called'.) However, the sceptic could be arguing that knowledge is impossible because nothing we believe is justified, even to the slightest extent. This is radical scepticism.

In my view, non-radical, knowledge-specific scepticism is a shallow form of scepticism that is not a serious challenge to everyday epistemic practices. The reason is that, even if we saw no way of turning it aside, we could always execute what Crispin Wright has termed 'a Russellian retreat', renouncing talk of knowledge (strictly understood) in favour of talk of justified belief. Or we could decide to use 'knowledge' in a more relaxed way. The burden would then be on the sceptic to show that such a retreat would entail significant loss. By contrast, precisely because it leaves no obviously acceptable line of retreat, radical scepticism poses a much more serious threat to our ordinary epistemological outlook. Radical scepticism threatens to wipe out all significant epistemological distinctions. This result would not be so easy to live with, if it is possible to live with it at all. Generalized radical scepticism implies that, rationally speaking, we can believe anything or nothing. Or as Bertrand Russell once put it, the man who believes he is a poached egg differs from the rest of us only by being in a minority.[1] It is not obvious that anyone could take this conclusion seriously in matters of practical concern.

This is all very different from ordinary (unphilosophical) scepticism. The everyday 'sceptical' view is that most people's epistemic standards are too low or too laxly applied, not that the very idea of justification is fatally flawed. Ordinary scepticism is *demanding* and *selective*. Philosophical scepticism, as radical and general, undermines the very epistemological distinctions on which everyday scepticism depends. It is not simply different from but *precludes* scepticism of the ordinary kind.

Philosophical scepticism's radical and general character harks back to the

simplicity and intuitiveness of sceptical arguments. Such arguments point to radical and general conclusions precisely because they exploit only 'lowest common denominator' features of knowledge. But for the same reason, it is hard to say how or even whether they go astray. The result is that philosophical scepticism is hard to dismiss but impossible (sincerely) to accept. It produces a kind of intellectual split personality. And this is why it is a problem.

Sceptical strategies

There are two basic families of sceptical problems, 'Agrippan' and 'Cartesian', each rooted in a distinctive pattern of argument. Some philosophers will find this claim objectionable, since it seems to ignore a third family: variations on the argument from error.

Certainly, arguments for philosophical scepticism have often made great play with the fact (if it is fact) that, even in the best of circumstances, the possibility that we are in error can never be entirely excluded. Given a demonstrative conception of knowledge, this point is potentially very damaging. Since acknowledging the possibility of error seems incompatible with laying claim to absolute certainty, if knowledge demands such certainty, we are well on the way to concluding that knowledge is impossible. However, while this knowledge-specific form of scepticism provides a strong motive for adopting a more modest, fallibilist conception of knowledge, it implies no concession at all to the *radical* sceptic. To get to radical scepticism, we need to embed acknowledgments of fallibility in a richer argumentative context, which is where my two basic families come in.

I shall begin with the Agrippan argument, not simply on account of its antiquity, but because it embodies the simplest sceptical strategy and, I am inclined to think, the most profound.

Agrippa's trilemma

The ancient sceptics distinguish various 'modes' or 'tropes': argumentative strategies for inducing universal suspension of judgement. The Agrippan argument is found in the Five Modes of Agrippa (about whom nothing is known, except that he is credited with the Five Modes), though the Five Modes simply codify a pattern of argument that occurs widely in Sextus Empiricus' account of sceptical procedures. The Five Modes are Discrepancy, Relativity, Infinity, Assumption, and Circularity.[2]

The Mode of Discrepancy makes the point that people can disagree about virtually anything. That of Relativity suggests that any claim can, perhaps should, be qualified with the rider 'according to you' (you personally, your school of thought, your culture generally, perhaps even your species).

However, we are not yet in sight of scepticism, still less radical scepticism. Sure, people disagree, but some people's views are wrong and can be shown to be wrong. And even where conclusive proof or disproof is not to be expected, some views may be better supported than others.

This is what the sceptic wants us to say. The point of the first two modes is to get us to acknowledge that, if anyone puts forward a claim as more than a personal opinion or a report on how things seem to him, he can reasonably be asked to explain why. Once this apparently innocent acknowledgment has been extracted, the Agrippan argument shows its teeth.

Suppose, then, that I make a claim—any claim. You are entitled to ask me whether what I have said is something that I am just assuming to be true or whether I know it to the case. If I reply that it is something I know, you are further entitled to ask me *how* I know. In response, I will have to cite something in support of my claim: my evidence, my credentials, whatever. But now the question can be renewed: is what I cite in defence of my original claim something that I am just assuming or something that I know? If the former, it will not do the job required of it: you can't base knowledge on a mere assumption. But if the latter, it will in turn need to be backed up, and so on.

Of course, attempts to provide justification come to a halt. But how? The sceptic will say that we just run out of ideas: either we have nothing to say, or we find ourselves going back over old ground. As an implied claim to knowledge, then, every statement I make invites a new challenge; and in the face of these constantly renewed challenges, I can do only one of three things:

1. Keep trying to think of something new to say—i.e. embark on an infinite regress (Mode of Infinity).
2. At some point, refuse to answer—i.e. make a dogmatic assumption (Mode of Assumption).
3. At some point, repeat something I have already said—i.e. reason in a circle (Mode of Circularity).

None of these gives us what we want.

With options (2) and (3), the sceptic seems on strong ground. Surely, statements offered in justification of an initial claim must themselves be justified: you can't base justifications on mere assumptions. So (2) is a dead letter. Equally, reasoning in a circle is a paradigmatically poor sort of reasoning: how can a statement support *itself*? Supposing that it could embodies a kind of pragmatic inconsistency, treating the same statement as needing support (in its role as conclusion) and as already in order (in its role as premise). So (3) is a dead letter too.

We might wonder whether option (1) is so bad. To be sure, new claims can lead to new questions, but what is so terrible about that? This reply misses the sceptic's point, which is that the regress is *vicious*. The sceptic is not saying simply that justifications *can* always be questioned further, should we find

reason to question them. His conclusion is rather that no claim is *ever* justified—to the slightest degree—unless, *per impossible*, we first run through an infinite series of prior justifications. Option (1), then, fares no better than the others. The conclusion seems to be that justification is a complete illusion.

I call the apparently forced choice between assumption, circularity, and infinite regress 'Agrippa's trilemma'. The Agrippan argument is also often thought of as 'the regress problem'. I prefer talking of the trilemma because it reminds us that the problem is to escape the regress without getting impaled on either of the remaining horns.[3] This challenge is not easily met.

Agrippa's trilemma lies at the heart of the ancient 'problem of the criterion': that is, the problem of determining what should be the standard or method ('criterion') for distinguishing between genuine knowledge and mere opinion. The Aristotelian solution to this problem is that knowledge depends on self-evidence: anything we know must either itself be self-evident or deducible by self-evidently correct steps from self-evidently true premises. The sceptical counter-move is to inquire why something's striking us as self-evident should be a guarantee of its correctness? If we ignore this question, the sceptic will say, we are just making a dogmatic assumption. But if we try to find some further validation of intuitive self-evidence, we threaten to open up the regress all over again. Finally, if we say that it is self-evident that certain things are self-evidently correct, we will be arguing in a circle. Alternatively, we could note that the demonstrative conception of knowledge places enormous weight on Reason, understood as a faculty of logical insight. But why should Reason, as opposed to sense-experience, be the arbiter of knowledge? How are we to decide: by Reason, or by experience? Whatever we do, knowledge seems to elude us.

How serious?

I have argued that philosophical scepticism is problematic only if it is intuitive, highly general, and radical. Does Agrippan scepticism meet this standard?

On the question of its naturalness or intuitiveness, the Agrippan argument's chief presupposition—that knowing differs from merely assuming or surmising and that this difference has something to do with an ability to back up or justify whatever can properly be said to be known—seems to be one of those lowest common denominator ideas that just about everyone (except perhaps extreme externalists) is ready to concede. Beyond this, the argument does not appear to be committed to any particular view about what 'back up' or 'justification' consist in. It is not, therefore, in any obvious way theoretically loaded.

The question of the generality of Agrippan scepticism is more complicated. Everyone has experienced arguments that run into an Agrippan dead end. Disputes that engage deep political differences sometimes get into a rut, with

the parties shouting at each other when not going round in circles. Or imagine a debate between a devout believer, who puts his faith in divine revelation, and someone with a thoroughgoing scientific outlook, who thinks that all hypotheses need to be verified (or falsified) by systematic research involving observation and experiment. Such opponents will disagree, not just over particular facts, but over how the facts are to be determined; and how is such a disagreement to be resolved? Perhaps each disputant will dig his heels in, that is, make a dogmatic assumption. Or perhaps the believer will say that God tells him to trust revelation, leading his opponent to accuse him of begging the question. But if the scientifically minded person says that we cannot rely on revelation, because there is no scientific evidence for the existence of God, the believer will return the compliment. Perhaps, then, they should look for some third way of adjudicating their epistemological differences. But what would that be, and what would give us the right to suppose that it was more to be trusted than the criteria already at issue?

While examples like this contribute to the argument's intuitive appeal, they may suggest that the sceptic sees his problem as restricted to admittedly controversial opinions. My having posed the Agrippan problem in terms of claims, challenges, and responses might itself encourage such a reading. This impression may be further reinforced by our recalling that two of the Five Modes—Discrepancy and Relativity—explicitly call attention to differences of opinion. So it is natural to conclude that the point of the trilemma is to show that such differences cannot be resolved. On this account, the Agrippan problem may be recurrent but far from completely general.

The problem is much more general than this. As the sceptic means his argument to be understood, nothing of importance turns on the existence of actual challenges or challengers. If, in a reflective moment, I take any belief I happen to hold and ask myself why I hold it, I am off and running. The problem, therefore, concerns justified belief in general, whether actually challenged or not. The Agrippan argument applies not just to things that anyone would agree are hard to know about but to anything whatsoever, no matter how obvious. Indeed, it is especially problematic in connection with things that are so obvious that we have no idea how to argue for them in a non-question-begging way. Agrippan scepticism is not just highly general: it is absolutely universal.

Does the Agrippan argument promote radical scepticism? Historically, the argument emerged when the Platonic-Aristotelian conception of knowledge was dominant, and its characteristic applications by the ancient sceptics—for example, in posing the problem of the criterion—often take that ideal for granted. However, we must not confuse applications of the Agrippan argument with the argument itself. Although the argument was originally formulated in the context of a demanding conception of knowledge, it is not tied to that conception. The argument has the capacity to promote radical scepticism,

whether or not its inventors saw it in this light. Because of this, we cannot avoid the argument simply by abandoning the classical conception for a more relaxed, fallibilist ideal of knowledge. This distinguishes the Agrippan argument from the pure argument from error. Indeed, abandoning the classical conception of knowledge for more modest epistemological aspirations is precisely what brings the distinction between radical and non-radical forms of scepticism into sharp focus.

External constraint

Given that the Agrippan problem is supposed to extend beyond matters of actual dispute, a natural reaction is that there is something artificial about it. After all, although disputes sometimes bog down or peter out through exhaustion, they can also come to a satisfactory end. When they do, it will be because whoever is on the defensive is able to cite evidence that is acceptable and convincing to all parties to the discussion. The production of such evidence is what we ordinarily call 'justification'. Claims that everyone finds too obvious to be worth discussing are not ordinarily called 'assumptions'. So what is the problem?

While the Agrippan sceptic will not dispute what we may call the *phenomenology* of everyday justification, he will caution against reading too much into it. When we say that justification proceeds from 'acceptable' claims, we really mean claims that are in fact *accepted*. This 'acceptability', which is entirely psychological and person-relative, implies no special epistemological status. The same goes for terms like 'evident' or 'obvious'. So the 'common knowledge' that terminates everyday justificational procedures can be seen, on reflection, to be nothing more than a body of common assumptions (and maybe not so common, if we broaden our imagined audience).

We can put the point like this: the reply to the sceptic just envisaged represents ordinary justification as entirely *dialectical*. But we can all see on reflection that even our ordinary concepts of knowledge and justification demand more. We connect knowledge with justification because we want to distinguish knowledge from *accidentally* true belief. Justification is supposed to establish—or increase the likelihood—that a belief is true. Justification is supposed to be *truth-conducive*. Showing that a claim follows from convictions that some or even all of us happen to share, if those convictions themselves rest on nothing at all, does not meet this requirement.

Agrippa's trilemma calls attention to the need for beliefs that meet two conditions. First, they must be non-inferentially credible: justified without deriving their justification inferentially from further beliefs. Second, given that justification cannot be seen as a purely dialectical exercise, they must be credible in a way that reflects some kind of external constraint. Empirical

knowledge cannot just be a matter of bandying words about: it must have something to do with the objective world.

The demand for external constraint is not a demand merely on knowledge or even justification, narrowly conceived. Rather, external constraint is a precondition for our beliefs having any genuine content or meaning, for our claimings and acceptings and inferrings being more than a game with meaningless counters.[4] If we cannot think of ourselves as justified, in a way that involves control of our beliefs by the world, we cannot even understand ourselves as thinkers.

If this is right, Agrippan scepticism is a very deep problem indeed.

Scepticism as paradox

Agrippan scepticism is absolutely universal. However, someone might suggest that this is its weakness, not its strength. The point of giving an argument is to justify one's conclusion: so isn't it obviously self-defeating to *argue* that nothing can ever be justified?

There are three points to make about this quick way with scepticism. The first is that inconsistency is readily avoided by a little care in the formulation of the sceptical position. The sceptic can claim that the one and only thing he knows is that he knows nothing else. This claim is entirely consistent. (Socrates said that believing it made him the wisest man in Athens.) True, this version of scepticism is not absolutely universal, but it is surely wide-ranging enough to pose a serious threat to our pre-theoretical epistemic attitudes.

There is something unsatisfactory about this response, for the suspicion is bound to remain that, if argument *can* sometimes be effective, why only here? This *caveat* leads to the second point, which is that it is not clear that sceptical arguments have to produce stable conclusions, in which case no restrictions on the scope of sceptical theses may be necessary. Sceptical argumentation does its work if it produces doubt, uncertainty, or suspension of judgement. Viewed from this angle, the situation looks like this: compelling arguments seem to establish that knowledge is impossible, and so incline me to claim to know that this is so. On reflection, I see that it would be inconsistent to claim to know that I am incapable of knowing anything, which in turn inclines me to withdraw the claim. But now, encouraged by the thought that knowledge must be possible after all, I can run through the argument again, once more reaching the conclusion that knowledge is *not* possible, whereupon I reflect that it would be inconsistent . . . and so on. Similar considerations apply to sceptical attacks on justification and to relativist attacks on objective truth. Sceptical arguments do not need to lead to stable—or even credible—conclusions in order to pose problems.[5]

Another way to put this point is to say that a form of scepticism that applies to itself may be not so much self-undermining as paradoxical. The reasoning

that leads to the sceptical conclusion seems to be compelling and yet no stable or credible conclusion emerges from it. In fact, this is how the Pyrrhonists treated all sceptical arguments: as intended to induce suspension of judgement, not assent to a negative, epistemological conclusion. This was because they thought that assenting to anything—including scepticism itself—would be a form of dogmatism, hence not scepticism.

Finally, we should remind ourselves that there is nothing necessarily self-defeating about assuming something to be true for the sake of refuting it. This is just the technique of arguing by *reductio ad absurdum*. We concede something for the sake of argument and go on to show that what we have conceded cannot be so. Many sceptical arguments can be seen in this light.

There is, then, no substitute for a deeper diagnosis of sceptical arguments. But I think that this would be true even if the objection in question were well taken. The problem with the argument is that it is purely dialectical. If we find ourselves drawn to sceptical arguments, it is not enough to be shown that their conclusions are unacceptable. This is something we believed anyway. What we want to know is how they go wrong and why they are, nevertheless, so seductive. Only through such understanding can their spell be broken.

The space of theories

Supposing that we decide not to throw in the towel right away, what are our choices? Assuming that an infinite regress is intolerable, we seem to have two, each defining a fundamental type of epistemological theory. The fact that the Agrippan argument seems to define the space of theories is powerful evidence for the view that the problem of scepticism plays a central role in constituting epistemology as a distinctively philosophical subject.

Perhaps the most natural response to Agrippan scepticism is to identify beliefs that are justifiably held without requiring further back-up. Perhaps some beliefs concern self-evident truths, or are in some other way beyond rational questioning. If there are such beliefs, they will bring requests for justification to a halt without being mere assumptions. Call such beliefs 'basic beliefs'. The quest for them is characteristic of *foundational* theories of knowledge and justification.

Foundationalism is under a cloud these days. But if we abandon the quest for basic beliefs, the only remaining way of avoiding a sceptical conclusion seems to be to argue that circular reasoning isn't as bad as it looks. Philosophers who take this approach point out that we do not have just one or two beliefs, and that our beliefs are not just a grab-bag of unrelated opinions. Rather, our beliefs about the world constitute an extensive and complicated *system*. The thought is thus that the members of such a system can give each other *mutual* support. This is the fundamental idea behind *coherence* theories of knowledge and justification. From this standpoint, our beliefs are not

justified because they rest on a foundation but because they are systematically interconnected. Systematic interconnection is not to be equated with simple circularity.

We shall examine the strengths and weaknesses of both strategies. But first we need to introduce our second family of sceptical problems.

Notes

1. Russell (1961), 646. Russell is discussing Hume's scepticism about induction. See Chapter 17, below.
2. Sextus Empiricus (1933). Clearly, the argument builds on considerations already explored by Aristotle. As we saw in Chapter 3, Aristotle defends a demonstrative conception of knowledge by arguing that, since knowledge requires proof, and since there cannot be an infinite process of prior justification, there must be ultimate premises that are simply self-evidently true. However, though Aristotle brings in regress considerations, it is far from clear that he appreciates the sceptical problem lurking in this argument, or even that he saw himself as raising and responding to sceptical issues at all. The precise function of the Modes in ancient 'Pyrrhonian' scepticism is controversial. See Williams (1988).
3. This point is stressed by Fogelin (1994), ch. 6.
4. McDowell (1994), ch. 1.
5. This point is clearly recognized by Hume. See Hume (1976), 1. 4. 1.

6

Experience and Reality

The external world

Whereas the Agrippan argument goes back to the classical Greek sceptics, our next problem, scepticism about knowledge of the external world, is comparatively new. It receives its first clear formulation in the writings of Descartes, around 350 years ago. Despite the enormous quantity of ink spilled over it, there is still no agreed solution.[1]

It seems undeniable that we discover what is going on in our surroundings by way of our senses. We are all the subjects of a constantly changing stream of sensory experience—visual, auditory, tactile, olfactory—that puts us in touch with objects and events in our environment. However, reflection on this seemingly obvious fact leads to strange conclusions.

Our experience of the world is the end-product of a complex chain of events. Take the case of visual experience: light of various wavelengths is reflected off the surface of an object; it is focused on my retinae, inducing a certain pattern of neural excitation; impulses from the retinal neurons combine to produce a further pattern of firing in my optic nerve and eventually in the occipital region of my cerebral cortex. As a result, I am aware of a red object, say an apple, a few feet in front of me.

It seems to be a consequence of this picture that the same experiential outcome could be produced by intervention at several points in the chain. And there do seem to be odd examples of illusions produced in just this way. In 'phantom limb' cases, amputees feel pain where a limb used to be. Presumably, the nerve impulses reaching the brain are the same as they would have been if they had originated in a more normal way, so it feels just as though the amputated limb were still there. But if our awareness of the world could be produced in several different ways, why are we so sure that it is produced one way rather than another? Descartes presses this question by way of two vivid examples.

The first appeals to the fact that we have dreams. Dreams are conscious experiences that we have while we are asleep. They do not correspond in any reliable way to ongoing events in our surroundings and so do not yield knowledge of them. But there is no waking experience that could not, in principle,

Notes

1. *Meditations on First Philosophy*, in Cottingham *et al.* (1984). On the novelty of Descartes's problem, see Burnyeat (1982) and Williams (1998).
2. On the essential generality of the sceptic's interests, see Barry Stroud, 'Understanding Human Knowledge in General', in Clay and Lehrer (1989).
3. e.g. DeRose (1995).
4. Dretske (1970).
5. Fogelin (1994), 93 f.
6. This account of Cartesian scepticism is given by Bernard Williams (1978).
7. This pragmatist line is associated especially with Dewey (1984).
8. This analysis of the pattern of Cartesian argumentation is due in all essentials to Ayer (1956). For a detailed defence of Ayer's analysis, see Williams (1996*a*), esp. chs. 2, 5.

7

Foundations

The goals of theory

What should we demand of an epistemological theory purporting to solve the Agrippan problem?[1]

The first desideratum is that the theory offer a coherent and plausible account of its basic ideas, showing how they indeed yield a solution. This means blocking the regress without covertly lapsing into assumption or circularity. While such a requirement may seem too obvious to be worth stating, it proves extremely difficult to meet. Closely examined, many theories fail to meet it.

A second requirement is that the theory explain how we can have a workably extensive body of knowledge. The Agrippan sceptic's suggestion that no belief or judgement is ever justified is formally contradicted by the thesis that there is at least one justified belief. But establishing this minimal thesis would not be a satisfactory response to Agrippan scepticism. Descartes famously argued that it is impossible rationally to doubt that one is thinking or that one exists, since doubting is a form of thinking and one has to exist to be deceived (*Cogito ergo sum*). However, if this were all Descartes was able to salvage from the sceptical wreck, he would not have accomplished very much. Not everything we believe need be justifiable, but enough must be.

Finally, it is not enough to show that knowledge is logically or 'theoretically' possible. While a theory of knowledge may involve a degree of idealization or 'rational reconstruction', it should show how knowledge is within the grasp of creatures more or less like ourselves. Knowledge should be humanly possible.

The foundationalist's dilemma

Foundationalism and the coherence theory are generally taken to represent the two main strategic approaches to the Agrippan problem. Typically, they are distinguished by what they imply about the *structure* of epistemic justification.

Foundationalist theories claim to identify basic beliefs (or perhaps some even more primitive form of awareness, such as experiences themselves),

which provide stopping-points for chains of justification. Such beliefs (or other states) are justified without deriving this positive epistemic status from any further beliefs. They are *intrinsically* credible. In this picture, a system of justified belief is like a building: there is a bottom level—a foundation—on which all the upper storeys stand. By contrast, coherence theories make justification a matter of the interconnections between beliefs. An individual belief derives its credibility from its playing a role in a larger system of beliefs. One might think of a system of beliefs as like a space station, held together by its internal structure and wheeling along through space without *resting* on anything. Or one might use the analogy of an ecosystem, in which organisms, by playing their distinctive roles, keep each other going. Where the foundationalist sees an *architecture* of knowledge, the coherence theorist sees something like an *ecology*, with beliefs occupying interdependent niches.[2]

Let us call the foundationalist's structural thesis *structural foundationalism*:

(STF) (i) There are basic beliefs, beliefs that are *in some sense* justifiably held without resting on further evidence. (ii) A belief is justified if and only if it is either itself basic or inferentially connected, in some appropriate way, to other justified beliefs.

Notice that clause (ii) makes use of the idea of justified belief. This may strike some readers as circular. In fact, it isn't. Rather, structural foundationalism treats the class of justified beliefs as specifiable recursively. To specify a class recursively, we being by identifying a basis: anything in the basis is a member of the class in question. We then introduce one or more generating relations: other things can get into the class by standing in an appropriate relation to already accredited members. Thus, starting with the base members, we add whatever is appropriately related to them. Then we add whatever is appropriately related to our expanded collection, and so on. Finally, we stipulate that nothing else is a member. (This is why I have 'if and only if' in clause (ii) of (STF).)

For a simple example, consider the natural numbers (0, 1, 2, 3 . . .). The basis is 0 and the generating relation is that of being the immediate successor of a number. So 0 is a number; so is its immediate successor (call it '1'); so is 1's immediate successor (call it '2'); so is 2's immediate successor, and so on. Nothing else is a number.

Foundationalists treat the class of justified beliefs in exactly this way. There are basic beliefs (specified some way or other) and there are justification-transmitting inferential connections (deductive, inductive, etc.) A belief is justified if and only if it is either itself basic or can be reached by a succession of accredited inferential moves from some initial collection of basic beliefs, just as a number is either 0 or generated from 0 by repeatedly finding immediate successors.

This structural thesis is important. But it does not fully characterize

traditional foundational epistemologies, which embody what I call *substantive foundationalism*:

(SUF) (i) and (ii) as above. (iii) There are certain kinds of beliefs (or other conscious states) that by their very nature—that is, in virtue of their content—are fitted to play the role of terminating points for chains of justification. These beliefs (or other states) are epistemologically basic because *intrinsically credible* or *self-evidencing*.

Where structural foundationalism says only that justification comes to an end with some beliefs or other, substantive foundationalism makes the much stronger claim that there are beliefs of certain broad kinds, identifiable by their distinctive content, with which justification always comes to an end.[3]

In saying that basic beliefs are basic in virtue of their distinctive content, I am not insisting that their content be sufficient to make them justified. The point is rather that only beliefs with a certain kind of content have the potential to be basic. So, for example, many foundationalists have held that one important source of basic knowledge is our 'immediate awareness' of our own thoughts and sensations. Thus my belief that I now have a headache is foundationally justified by my having a headache: not the content of my belief but the fact that it points to. I need not deny this. My point is simply that, for substantive foundationalists, basic beliefs can be sorted into kinds on the basis of their distinctive content: for example, as beliefs about immediate experience. So the difference between structural and substantive foundationalism is like that between 'All roads lead somewhere' and 'There is somewhere (Rome perhaps) to which all roads lead'. Substantive foundationalism implies structural foundationalism, but not vice versa.[4]

What is the appeal of the stronger thesis? One factor is that structural foundationalism, treated as a *purely* structural thesis, is not obviously responsive to Agrippan scepticism. The Agrippan sceptic claims not to challenge the phenomenology of everyday justification, conceding that ordinary justification involves citing evidence that no party to the discussion is *in fact* prepared to challenge. But he sees this 'justification' as merely dialectical, thus vulnerable to the Mode of Assumption. A foundationalist response to the Agrippan sceptic needs basic beliefs that are not simply *unchallenged* but *unchallengeable*. Perhaps not absolutely unchallengeable, but at least not automatically challengeable and thus at least *prima facie* justified.

While correct as far as it goes, this cannot be the whole story. Presenting foundationalism and the coherence theory as the two main theoretical strategies for dealing with the Agrippan problem, while illuminating theoretically, is historically off-key. Foundationalist ideas pre-date Agrippa's trilemma. (As we have seen, they are strongly evident in Aristotle, whose concern with sceptical problems is much less certain.) As for coherence theories, though we may find coherentist elements in Plato's epistemological theorizing, pure coherentism—understood as a clear alternative to foundationalism—is a

relatively recent invention, belonging (at the earliest) to the nineteenth century. The choice between foundationalism and the coherence theory did not spring into being as an immediate consequence of Agrippan scepticism.

That foundationalist ideas pre-date Agrippan scepticism suggests that they have at least some appeal independently of their anti-sceptical potential. This is just as well: wanting there to be terminating judgements does not guarantee that there are any. Foundationalism would never have been plausible had there not been an example of knowledge apparently conforming to the foundationalist ideal. The deep (or at least original) source of foundationalism's appeal lies in the axiomatic method. Axiomatized theories—such as Euclid's presentation of plane geometry—offer a compelling example of foundationalism in action.

The geometrical model, however, is problematic as a model of knowledge in general, and not just because it encourages a demonstrative—hence extremely restrictive—conception of knowledge. Euclid's geometry is an example of a finitely axiomatized theory: it involves a fixed number of proper axioms (together with a few definitions and general principles). This gives the theory a certain independence of epistemological theory. Because its basic assertions can be displayed for the reader's inspection, their self-evidence can be directly experienced, case by case, and so does not need to be explained. However, a general theory of knowledge must allow for an expanding stock of basic beliefs. Accordingly, a generalized foundationalism demands that we identify one or more *kinds* of basic belief or judgement. This is substantive foundationalism, which from now on is what I shall have in mind whenever I speak of 'foundationalism' without qualification.

According to substantive foundationalism, evidence for the truth of belief of a given non-basic kind *must*, in the last analysis, be sought in beliefs belonging to some more basic, 'epistemologically prior' kind. These epistemological kinds cross-cut and are more fundamental than ordinary subject-matter divisions. They correspond to what Descartes called 'the order of reasons'.

Viewed in this light, beliefs and judgements stand in natural relations of epistemic priority and fall into natural epistemological kinds. These relations and kinds are natural in the sense that they are fixed and permanent. They exist independently of our interests and decisions. They define 'our epistemic position', a framework of constraints within which all our knowledge-seeking activities are fated to take place. Substantive foundationalism is thus committed to what I call 'epistemological realism'. This is not realism as a position within epistemology (the view that there is a real world out there which, in favorable circumstances, we can learn a lot about) but rather a kind of extreme realism about the subject-matter of epistemological theory.

The dominant strain of foundationalism today is empiricism, in one form or another. Modern foundationalists have almost invariably taken it that some basic beliefs must be empirical: in fact, perceptual. Most have taken it that

basic perceptual beliefs have to do with the contents of experience—how things appear to us—experience, or experiential beliefs, thus being epistemologically prior to beliefs about the external world.

The second limitation that must be overcome, if the geometrical model is to be extended to knowledge generally, concerns the connection between basic and non-basic knowledge. The individual steps in Euclid's reasoning are meant to be self-evidently deductively valid. Again, this gives geometry a certain independence of epistemology. (Indeed, it was not until the work of Frege, at the end of the nineteenth century, that logical theory progressed to a point where mathematical reasoning could be represented as formally valid.) But a generalized foundationalism, especially a version that has abandoned the demonstrative conception of knowledge, cannot restrict itself in advance to deduction as the only warrant-conferring form of inference.

These points suggest two problems for foundationalism. The first is to explain how it is that some beliefs manage to be genuinely basic, in the foundationalist's strong sense of intrinsically credible. Call this 'the problem of the basis' or 'the problem of security'. The second is to explain the inferential connections between one's chosen basis and a workable superstructure of empirical knowledge. Call this 'the problem of the superstructure' or 'the problem of adequacy'.

There is an obvious tension between the two problems. The more ambitious a claim, the less our assurance of its truth: only beliefs or judgements that are fairly modest in their assertional content are plausible pretenders to the role of intrinsically credible basic beliefs. Accordingly, the quest for security puts the foundationalist under pressure to accept a severely restricted basis for knowledge. But the more restricted the basis, the more challenging the task of recovering a workable superstructure. This is the foundationalist's dilemma: to define a basis for knowledge modest enough to be secure but rich enough to be adequate. This problem offers the critic plenty of room for manoeuvre.

Sources of knowledge

Prior to the nineteenth century, virtually all philosophers concerned with knowledge were foundationalists of one sort or another. But they tended to approach questions about the epistemic status of various kinds of beliefs by investigating their origins, rather than by talking, as I have done, about the structure of justification. So is my account of foundationalism misleading?

A good way of casting light on this issue is to further investigate the distinction between *a priori* and *a posteriori* knowledge. This distinction is generally explained in terms of dependence and independence with respect to observation. What kind of 'dependence' is at issue here?

Traditionally, this question has been answered in terms of how the different

types of knowledge can (or must) be acquired: particularly in terms of which faculties—the senses or reason—need to be exercised in their acquisition. *A priori* knowledge can be established by the use of Reason alone. By contrast, to find out whether there a squirrel on my lawn right now, I *have* to go and look ('use my senses'). Or if I don't go myself, I have to get someone else to look for me. There are many ways to learn particular facts about the world—through testimony, by reading books, and so on—but for such information ever to have got into circulation, at some point somebody had to go and see for himself. This dependence on observation as its *ultimate source* is the hallmark of *a posteriori* knowledge.

Some philosophers argue that the traditional focus on the origins or sources of different kinds of knowledge was always misguided: that the correct way to distinguish between *a priori* and *a posteriori* knowledge is not in terms of how the two types of knowledge *originate* but rather in terms of how they are *justified. A posteriori* knowledge is knowledge that involves beliefs that require support from observational evidence. By contrast, *a priori* knowledge, as in pure mathematics, depends solely on logical intuition and reasoning. Observation is irrelevant.

There is something initially appealing about this view. We acquire beliefs in all sorts of ways, which need not line up in any neat way with important epistemological distinctions. A mathematician looking at the petals of a flower may be stimulated to have some new thoughts about symmetry. Her experience may be powerful enough to convince her of the correctness of her new ideas. And they may even be correct. Still, whether they count as knowledge will depend on her providing proofs. Similarly, a scientific hypothesis might come to us in a dream. But what converts it from speculation to knowledge is confirmation by experimental evidence.

However, there is less disagreement here than first appears. Although whether a belief amounts to knowledge has traditionally been linked to its sources, those sources have themselves been understood in ways that make them relevant to questions of justification. Knowledge is belief from *reliable* sources; and to trace a belief to a reliable source is to justify it. Recognizably reliable sources are thus *authoritative.* This is a perfectly commonsensical view. If we know that our informant is generally truthful, we will be justified in believing what he tells us. If an expert on English furniture tells us that the chair is a Chippendale, we will take his word for it. And so on.

Traditionally, Reason has been conceived on the model of a hyper-reliable informant. In generating mathematical knowledge, Reason operates in two ways. The most basic mathematical truths are grasped by a kind of pure rational insight: they are self-evident because they literally 'stand to reason'. Less obvious truths are deduced from these self-evident beginnings by individually obvious steps. Reason, when operating properly, leads us to a grasp of necessary truths and so cannot misinform us. To trace a belief to Reason is thus to provide it with impeccable credentials and so to justify it. There is no tension between treating theories of knowledge as theories of justification and the traditional tendency to understand different kinds of knowledge in terms of their distinct, authoritative sources.

Rationality

Since there are many versions of foundationalism, foundationalists can diverge in their epistemological views. Nevertheless, there are certain general ideas that go naturally with a foundationalist conception of knowledge and justification.

Let us begin with the problem of rationality. According to an immensely influential view, it is always irrational to hold beliefs that are not adequately justified. Call this a strongly *justificationist* conception of rationality. This conception is deeply embedded in the foundationalist outlook. Indeed, it is hard to say which is the chicken and which the egg. Foundationalists feel the need for a strong response to scepticism because scepticism attacks justification, without which there is no rationality. With a different conception of rationality, we might not feel the need for such a strong response to scepticism.

I must stress that I am thinking here of substantive foundationalism. As we know, Agrippan scepticism is compatible with structural foundationalism. The Agrippan sceptic does not deny that everyday 'justification' generally comes to a halt with beliefs or claims that no party to the discussion feels the need to deny. His point is that this 'common knowledge' is no more than conveniently shared assumptions. Furthermore, there is no reason why any given dispute *must* so terminate. In many cases, we may not share enough assumptions to bring a dispute to a mutually satisfactory conclusion, in which case our dispute will bog down and may stay bogged down indefinitely. To a philosopher attracted to a strongly justificationist conception of rationality, this is intolerable, for it implies that there we have no guarantee that important disagreements can be rationally adjudicated. This explains why purely dialectical justification is distasteful to so many philosophers: it seems hostage to contingent agreement. In the absence of such agreement, disputes will be resolved by non-rational means: rhetorical devices, or worse still, force.

Even if extensive disagreement is acceptable, the Agrippan sceptic's tolerance for structural foundationalism is a standing invitation to relativism. By contrast, substantive foundationalism holds out the promise of a neutral basis for the (in principle) adjudication of all disputes, or all that are genuinely significant. This promise is an important source of foundationalism's appeal.

Epistemic risk

Moving on, what does foundationalism look like as a method of inquiry? Some philosophers, Descartes perhaps, have used the foundationalist picture

of justification as a literal method: a procedure to be followed in building up a correct system of beliefs. So understood, foundationalism amounts to a counsel of extreme caution: we are to add to our stock of beliefs only such beliefs as we can infer, by evidently warrant-conferring procedures, from a secure basis of primitively evident beliefs. This should make us wonder whether a strongly justificationist conception of rationality is really so rational. Of course, we should like to know the truth: however, avoiding errors and maximizing our number of true beliefs are not the same thing. Methodologically speaking, foundationalism is dramatically tilted towards error-avoidance. Perhaps we would do better in the long run if we loosened up and took more epistemological risks. Our mistakes might be more than offset by extra discoveries.

Contemporary philosophers do not usually interpret foundationalism in procedural or genetic terms. They recognize that we are all acculturated into a complicated system of beliefs and that no one does or could construct such a system from the ground up. Foundationalism is neither a psychological account of how such systems are acquired, nor a set of prescriptive rules governing their construction, but an account of their justificational structure. Foundationalism reveals constraints to which a belief-system must conform to count as justified, hence rational.

From this standpoint, the naive problem of method—how should we conduct inquiry—is ill-stated. Epistemologically speaking, what matters is not where ideas come from but how they can be verified (or falsified). This understanding of foundationalism goes naturally with a sharp distinction between 'contexts of discovery' and 'contexts of justification'. There are no strict rules for thinking up new ideas, only rules of thumb. But there are strict criteria of justification. So although wild ideas can be entertained, they cannot be retained if they fail to find adequate support. Accordingly, the non-genetic understanding does not affect foundationalism's bias towards error-avoidance.

Empiricism and the *a priori*

In principle, substantive foundationalism can take either rationalist or empiricist forms. However, in recent times the dominant form of foundationalism has been empiricist. For modern foundationalists, at least some basic beliefs must be empirical.[5]

This brings us to what I called the problem of unity: are there fundamentally different minds of knowledge, associated with different methods of verification? Empiricist foundationalists think that there are. They want to allow for both demonstrative knowledge, as found in pure mathematics, and empirical knowledge, as found in the natural sciences. This puts them under pressure to recognize at least two sources of basic knowledge: some faculty of

rational or logical intuition, by which we recognize primitive logico-mathematical truths and the validity of elementary inferences; and some forms of primitive experiential knowledge, which give us an expanding database of contingent information.

Since this database is apt to be highly restricted, we are not going to be able to make a workably extensive system of empirical knowledge out of its deductive consequences alone. This pushes foundationalists towards recognizing two forms of warrant-conferring inference: deductive and inductive. As we shall see, foundationalists are required to regard both sorts of inference as *a priori* valid.

How should empiricists understand the *a priori*? To address this question, we need to review (and introduce) some distinctions:

1. A Priori *versus* A Posteriori. *A priori* knowledge is independent of empirical justification or verification. *A posteriori* (or empirical) knowledge depends on experience or observation. Virtually all contemporary epistemologists regard scientific knowledge, which depends on experimental confirmation, as *a posteriori*. (Historically the *a priori/a posteriori* distinction has been closely associated with that between the innate and the learned. The very idea of the *a priori*, as that which can be known prior to experience, encourages this assimilation, which should nevertheless be resisted.)

2. *Necessary versus Contingent.* A necessary truth is one that holds 'in all possible worlds'. Truths of pure mathematics and elementary logical truths have almost (though not quite) universally been held to be necessary. Two and two don't just happen to make four: no other result is so much as possible. (Of course the word 'two' might have had a different meaning, but that is beside the point.) Anything that is not necessary is contingent: contingent truths are those that happen to obtain but might not have. I happen to sitting at my desk right now. But I might have decided to take a walk or have done various things that would have placed me elsewhere. Early modern Rationalists thought that some fundamental principles of the natural sciences are necessary. Today, however, virtually all philosophers agree that science deals exclusively with matters of contingent fact.

3. *Analytic versus Synthetic.* Analytic truths are truths that hold by definition or in virtue of meaning. They are 'conceptual' truths. (Analytic falsehoods are statements that are—perhaps implicitly—oxymoronic.) Anything that is not analytic is synthetic. 'Bachelors are unmarried males' is analytic. 'Bachelors tend to die before married men' is synthetic. Anyone who understands an analytically true statement must recognize that it is true: there is no gap here between understanding and knowledge. On the other hand, analytic truths convey no information about the world. By contrast, knowledge of synthetic truths is clearly 'substantial' in that it is not guaranteed simply by understanding the meaning of one's terms.

These are distinct contrasts: the first is epistemological, the second metaphysical, and the third semantic. How they line up is an interesting question. It has been characteristic of empiricism in its more radical forms to say that they

are coextensive. That is: a truth is *a priori* knowable if and only if it is necessary; and a truth is necessary if and only if it is analytic. On this view, there is no *a priori* knowledge of substantial or factual matters, so that attempts to demonstrate factual truths *a priori*, such as the Ontological Argument for the existence of God, must be deeply misguided. Kant saw this as Hume's challenge to philosophy, understood as the quest for *a priori* knowledge of fundamental truths. How is *a priori* knowledge of synthetic truths so much as *possible*? Of course, perhaps it isn't. Empiricists are sure that it isn't.

The reason behind the empiricist outlook is not far to seek. It is plausible to hold that whatever is necessarily true must be knowable *a priori*. Experience can only tell us how things are, not how they must be. At the same time, empiricists are suspicious of the idea that purely *a priori* procedures can tell us anything about how the world actually is. The solution is to explain necessity in terms of analyticity. To be sure, there are statements that cannot be false: but this is because their truth is guaranteed by our linguistic conventions. There is no *a priori* insight into external reality. Modern empiricism is characterized by a linguistic doctrine of necessity. A consequence of this doctrine is a deflationary attitude towards *a priori* knowledge. Though there is such knowledge, it is not really 'substantive' or 'factual'. It merely reflects our linguistic conventions.

The plausibility of this epistemological standpoint owes a lot to developments in mathematics and science. For Rationalists, geometry was always the paradigm of rational knowledge. Epistemologically and metaphysically, it seemed on a par with arithmetic: *a priori* and necessary. At the same time, by revealing the structure of space, it seemed to tell us important truths about the real world. However, the development of non-Euclidean geometries calls this paradigm into question. From a purely mathematical standpoint, there are all kinds of geometrical spaces. Which mathematical theory best describes real space is an empirical issue, not decidable independently of physical considerations. The development of non-Euclidean geometries contributes powerfully to the appeal of the empiricist idea that pure mathematics is a wholly conceptual form of inquiry.[6]

Because they continue to be widely recognized, it is important to mark these distinctions. However, none is beyond controversy and not everyone accepts the empiricist position.[7] But I have a different question, which is whether some (or even all) of these distinctions have outlived their usefulness. For foundationalists, who are constrained to take the *a priori/a posteriori* distinction seriously, it is important to determine how the lines should be drawn. For others, the significance of this task is much less clear.

Drawing lines

This brings us to the problem of demarcation. Not only does foundationalism suggest a sharp internal boundary between *a priori* and *a posteriori* knowledge, it produces an almost irresistible impulse towards fixing external boundaries. Some things we like to think we know resist justification on the restricted basis to which the foundationalist's quest for security forces him to retreat. Such things will fall outside the province of knowledge.

Some philosophers connect questions of justification with questions of meaning. Thus logical positivists put forward the Principle of Verification, according to which a statement is not meaningful unless it is possible to explain what would count as verifying or falsifying it. Truly interminable disputes are not just pointless: they have no real content. Positivists saw most 'philosophical' disagreements in this light. But philosophical theses are not alone in falling outside the charmed circle of knowledge, hence of genuine meaningfulness. By foundationalist standards, moral claims—indeed value-judgements generally—cannot plausibly be regarded as empirically grounded. Positivists tended to accord them 'emotive' meaning. They make no literal or factual claims, which might be true or false, but express our attitudes of approbation or disapproval. This is sometimes called the 'Boo–Hooray' theory of moral judgement. Value-judgements are not really judgements: they are cheers or jeers.[8]

The ideal of clarity

We have seen that empiricist foundationalism has pronounced demarcational tendencies. But it has certain unifying tendencies too. Because foundationalism purports to be a fully general theory of epistemic justification, it is unfriendly to the idea of a plurality of methods. At best, foundationalists will recognize different ranges of data.

Some philosophers have held that there is a fundamental methodological distinction between the natural and the 'human' sciences (*Natur-* and *Geisteswissenschaften*). In physics or chemistry we seek to explain events by bringing them under laws; and we test putative laws experimentally. In history or anthropology we want something different, a kind of interpretative understanding (*verstehen*). We want to get a feel for past circumstances or exotic ways of life 'from the inside'. This is more a matter of sympathetic imagination than experimentation.

On the whole, the empiricist outlook has not been friendly to such ideas: hence the tendency for empiricist psychology to take behaviouristic forms. But it is doubtful whether classical Rationalism should think much better of them. This is not surprising: it is not easy to assimilate notions like *verstehen* to the foundationalist model of verification. But this is not all. There is a tendency, deeply entrenched in the foundationalist outlook, to place a high value, not only on securely grounded knowledge, but also on absolute conceptual clarity. 'Clear and distinct ideas' was Descartes's epistemological ideal, and 'Define your terms' remains a characteristically foundationalist slogan. These

demands—for epistemic security and conceptual clarity—are two sides of the same coin. The meaning of a basic statement or belief must be absolutely self-contained: its meaning cannot shift, depending on whatever context of further statements or beliefs it comes to be embedded in. This kind of variability would be incompatible with the requirement of intrinsic credibility. Epistemically and semantically, foundationalism is *atomistic*. As for non-basic statements, each of these must have a precisely delimited meaning all of its own. Only thus will it stand in definite relations to some specifiable range of basic evidence. Foundationalists therefore distrust metaphor. Metaphorical description, by its very suggestiveness and open texture, resists association with precisely and individually delimitable verification-conditions. For the foundationalist, this automatically degrades its cognitive value.[9]

Some will conclude that the demarcational consequences of foundationalism are reason enough for seeking another approach to epistemology. However, we should remember that foundationalism promises significant rewards for the costs it exacts. Can it deliver them?

Notes

1. My discussion here is heavily indebted to Fogelin (1994), ch. 6.
2. Ayer (1956) treats the full range of sceptical problems within a broadly foundational perspective. In post-war American philosophy, the most significant defender of foundationalism is Roderick Chisholm. See Chisholm (1982).
3. In Williams (1992), I make the same distinction using the terminology of formal versus substantive foundationalism. This terminology was introduced by Ernest Sosa and I have changed my terms because Sosa's distinction is very different from mine. For Sosa, substantive foundationalism contrasts with the coherence theory while formal foundationalism contrasts with what he calls 'epistemic pessimism'. Formal foundationalism (in epistemology) says that

   > (i) every belief with a certain non-epistemic property F is justified, (ii) if a belief bears relation R to a set of justified beliefs then it is itself justified, and (iii) every belief is justified in virtue of (i) or (ii).

   In effect, formal foundationalism is the view that the class of justified beliefs can be specified recursively, by identifying a basis and a generator: elements you start with and a procedure for generating further elements from them (plus what you add as you go along). The epistemic pessimist is the philosopher who doubts that any such 'interesting' specification of the class of justified belief is possible. My notion of structural foundationalism is weaker, since it does not require that the class of basic beliefs be specifiable in wholly non-epistemic terms, still less that it be specifiable (as Sosa sometimes seems to suggest) in non-normative terms. By contrast, my notion of substantive foundationalism is stronger, involving as it does the commitment to content-based classifications of basic beliefs. I am not sure whether our different ways of dividing up the territory reflect deep disagreements. I adopt mine with an eye to setting up the contextualist epistemology I present in Chapters 13 to 15. For Sosa's views, see his essay 'The Foundations of Foundations' and his justly renowned 'The Raft and the Pyramid', reprinted as chapters 9 and 10 of Sosa (1991).

4. I am indebted to Ernest Sosa for forcing me to clarify this point.
5. Rationalism has not had a very good press at all in recent philosophy. But for an intriguing defence of traditional Rationalism, see BonJour (1998).
6. BonJour (1998) argues forcefully that empiricists, whether 'moderate' or 'radical', have no good account of *a priori* knowledge. He is also willing to take a Rationalist stance towards geometry as providing *a priori* knowledge of the structure of space.
7. For an important discussion of why these distinctions should not be thought to line up, see Kripke (1980), Lecture 1.
8. For the positivist outlook, see the essays in Ayer (1959).
9. For a critical discussion of the foundationalist ideal of absolute clarity and a penetrating critique of foundationalism generally, see Elgin (1996).

8

The Problem of the Basis

Intrinsic credibility

Foundationalism aims to bring demands for justification to an absolute stopping-point with epistemologically basic beliefs. In its traditional form, foundationalism requires basic beliefs to be *intrinsically* credible: justified independently of all inferential connections to further beliefs or factual presuppositions. I say 'in its traditional form' because one might try to combine a foundationalist picture of the structure of justification with a pure reliabilist account of basic knowledge. In this 'externalist' version of foundationalism, basic knowledge would be non-inferential but not independent of external facts (concerning the reliability of various cognitive processes). Since we have already found reason to reject pure reliabilism, traditional foundationalism is what we need to examine here. We shall take up the question of what we can learn from the reliabilist alternative in Chapter 15.

The idea of intrinsic credibility entails an *atomistic* view of justification. Indeed, it entails an atomistic conception of meaning and understanding. If a belief can be justifiably held in the absence of any further beliefs, *a fortiori* it can be held (justifiably or not) in such conditions. The foundationalist's commitment to intrinsic credibility is thus a commitment to *encapsulated* knowledge: knowledge that is independent, justificationally and semantically, of any further knowledge.

Since intrinsic credibility is credibility that attaches to a belief independently of all external factors—even factors concerning what else a person believes—it must somehow depend only on content: what a basic belief means or is about. For the foundationalist, *content determines (epistemic) status*, in the sense that only beliefs with a certain kind of content are even candidates for playing the role of basic beliefs. Foundationalism is committed both to a content-based theory of epistemic justification and to an atomistic account of content. How content, or meaning, might be constituted, given the foundationalist's semantic atomism, is a question we shall discuss in due course.

Traditionally, foundationalists have insisted that basic beliefs be rationally indubitable or (a slightly weaker notion) incorrigible: exempt from rational

correction. One way for a judgement to be indubitable, *prima facie* consistent with the content theory, is to take as its object a proposition expressing a simple necessary truth. It would be absurd to think 'I know exactly what it means to claim that two plus two equals four, I'm just not sure that its true'. Because the falsity of such a proposition is inconceivable, you cannot have a clear grasp of its meaning and be in doubt as to its truth. However, this is not a good model for a generalized foundationalism, which is supposed to apply to *empirical* knowledge of *contingent* truths. The foundationalist needs to explain how indubitability or incorrigibility can attach to propositions that are not necessarily true.

Descartes's famous 'I think' and 'I exist' perhaps fit the bill. If I am in a position to think such thoughts, they must be true. But since I might never have been born or might have been, at this moment, in a state of dreamless sleep, it is not a necessary truth that I exist or that I am thinking. However, such 'pragmatically self-verifying' propositions seem too specialized to serve as models for foundational judgements generally. Here our second success condition—that an epistemological theory ought to salvage a usably extensive body of knowledge—makes itself felt. For the foundations of empirical knowledge generally, we need a form of *perceptual* knowledge that is non-inferential in the foundationalist's demanding sense.

Appearance and reality

Most modern foundationalists have sought basic empirical knowledge at the level of experience: our awareness of how things appear to us or, more generally, in our immediate knowledge of our own mental states, including all occurrent thoughts and feelings. Their thought has been that, although we can always go wrong in judgements about how things objectively are, we cannot go wrong about how they *seem* or *look* to us to be or how we *think* they are. Not even the Evil Deceiver or scientists running brain-in-vat experiments can deceive us here. They may systematically deceive us about how things are, but the deception depends on our awareness of how things appear to be. The whole point of the brain-in-vat experiment is to simulate normal experience, which means that the sceptical scenario concedes that experiential information is not subject to the same kind of doubt as knowledge of the external world. Experiential knowledge is basic by virtue of being the 'highest common factor' in 'veridical' and 'non-veridical' perception.[1]

There is already something disturbing about this line of thought. We saw that Agrippa's trilemma can be seen as revealing the need for some kind of external constraint on our thoughts and beliefs. In a way, empiricist foundationalism respects this demand. Our experiences just come to us: we have no control over how things appear to us. Our beliefs, in so far as they are subject to experimental control, are subject to a kind of external constraint, with the

result that justification is more than dialectical. But is this constraint sufficiently 'external'? Looked at from another angle, it seems not. It seems too subjective: not constraint by objects in the world but constraint by our experiences of objects (objects which, pending a refutation of Cartesian scepticism, may not even exist). But setting this problem aside, let us see how far we can get with the foundationalist's basic ideas.

There seems to be something right in the claim that we cannot be wrong about how things appear to us. 'Looks' or 'appears' talk cannot be iterated. I can say 'This tie looks green to me, but the light is a bit funny in here so perhaps it is really blue'. But I can't say 'This tie looks as if it looks green to me, but perhaps it really looks blue'. What would such a claim even mean?

While this is suggestive, we must go carefully. One reason why 'looks' talk does not iterate might be that it is a *guarded* form of reporting on how things are. I find myself inclined to report that the tie is green but, realizing that the lighting is not ideal, I am reluctant to commit myself. I express this diffidence by saying that the tie looks green, rather than that it definitely is green. Once I have withheld commitment, there is no further guardedness for additional 'looks' talk to express, and so such talk cannot be iterated. To be sure, this is not the last word on perceptual appearing. Still, it is clear that there is much more to the foundationalist idea of basic knowledge than is obviously justified by the logic of ordinary 'looks' talk.[2]

Against acquaintance

What, according to foundationalism, does empirical knowledge ultimately rest on? Does it rest on beliefs or judgements about experience? Or does the foundation of empirical knowledge consist in our experiences or sensations themselves? Foundationalists have been tempted to say 'in a way, both'. Experiential judgements are the most basic form of judgement; but their epistemic status as absolutely non-inferential depends on a more fundamental kind of awareness, below the judgemental level. Thus, I know that I have a headache simply by having one; or I know that something looks red simply by virtue of being in the appropriate state, a state in which redness is present in my visual experience. In this spirit, Bertrand Russell wrote of 'knowledge by acquaintance'.[3] For reasons we shall come to, this is more than a casual temptation: foundationalists are under strong theoretical pressure to take this line.

I remarked in Chapter 1 that our epistemological tradition has tended to take propositional knowledge as its primary focus and to treat such knowledge as theoretically fundamental. Now I have to enter a caveat. Philosophers like Russell, who see empirical knowledge as resting on knowledge by acquaintance, think that even the most rock-bottom propositional knowledge rests on something more basic still.

What are we acquainted with? Typically, foundationalists take experience to involve the 'immediate' presence to consciousness of a special kind of mental particular: a sensation or sense impression. Such things are immediately present or 'given' in the sense that we know about them simply by having them. To stress this feature, the particulars we 'sense', the objects of acquaintance, have often been called 'sense-data', though this terminology is no longer as popular as once it was. The point of saying that the sensing of sense-data involves 'immediate' presence to consciousness is that this form of awareness is not mediated by any kind of representation, including the sort of conceptual representation involved in linguistically articulated thought. The reason for insisting on such absolute immediacy is that representation seems always to involve the possibility of misrepresentation. But our awareness of our own thoughts and sensations is supposed to be error-proof; for it is by being error-proof that it underwrites the non-inferential judgements that occur at the next level up.

This approach to the foundations of knowledge is fraught with difficulties. One problem stands out immediately: how could merely having experiences or sensing sense-data justify anything? However basic knowledge is understood, it must be capable of standing in logical relations to whatever judgements rest on it. For example, it must be capable of being consistent or inconsistent with them. But this means that even basic knowledge must involve propositional content and so cannot consist in a mere relation to a particular. Sensing a sense-datum is no more knowing anything than is standing next to a lamp-post. For this reason, some philosophers reject the idea that experience—if experiencing is sharply distinguished from judging or believing—can justify anything.[4]

We may baulk at this: surely experience makes an essential contribution to our knowing about the world around us. No doubt it does, in some sense. But if we are to treat experience as the foundation of knowledge, experience must itself be understood to involve propositional content: the sort of content expressed by complete sentences. We cannot understand experience in terms of the mere presence of a particular, even a mental particular, and hold on to the idea that experience constitutes a form of knowledge, capable of playing a justificatory role. This does not mean that, in experiencing the world, we are always muttering sentences under our breath. It does mean, however, that the content of experience, if it is to have any epistemological significance, must be propositional in form. This amounts to rejecting any sharp distinction between experiencing and judging or believing.

Foundationalists have tended to vacillate on this point. Many foundationalists have contrasted the givenness of experiences or sense-data with any kind of judgement. Russell's 'acquaintance' is supposed to be a direct relation to a sense-datum, conceived as a sensory particular: for example, a red patch in one's visual field. This account of experience gains an air of intelligibility from its analogy with ordinary talk about acquaintance with people and places:

knowing the Prime Minister, knowing Manhattan, and so on. But the analogy is superficial. Russellian knowledge by acquaintance is supposed to be prior to—thus independent of—all propositional knowledge, which Russell refers to as 'knowledge by description'. Knowing people and places is nothing like that. It is impossible to 'know Manhattan' without knowing anything whatsoever *about* Manhattan.

Acquaintance with sense-data is supposed to be a form of non-propositional knowledge. This is what I claim we cannot understand: how something can be non-propositional and yet knowledge. Knowledge provides evidence, grounds for further inferences. But only that which can be true or false—thus prepositionally contentful—can confirm or refute. Under pressure, Russell tends to write of our being acquainted with sense-data *as* red, square, or whatever. But how is this different from being aware *that* something in one's visual field is (or appears) red or square? It looks like an attempt to have things both ways.

There are, of course, ways in which the richness and complexity of perceptual experience exceeds our capacity for precise verbal expression. Some philosophers regard this as a decisive reason for crediting experience with non-conceptual content (perceptual content, say), just as Russell urges. But we should not make too much of our verbal limitations. True, we can discriminate many more shades of colour that most of us remember the names for. But even here, we are aware that the precise shade of blue in the curtains is different from that in the carpet. The main point stands: only where there is propositionally articulable content can there be relations of justification.[5]

Someone might object that this argument assumes an excessively 'intellectualist' conception of justification, according to which justification is entirely dependent on logical relations among propositions (or the propositional contents of beliefs and other mental states). But there are other ways in which experiences might be relevant to justification: for example, they may serve (causally) as reliable guides to our environment, so that the policy of relying on them (except where there is reason to think that appearances are deceptive) is justified by its consequences (and not by any intrinsic feature of the experiences themselves).

I have a lot of sympathy with this line of thought. But it is no defence of traditional foundationalism, which itself takes an 'intellectualist' conception of justification for granted. Challenging that conception will lead us eventually to epistemological views profoundly at variance with traditional foundationalist ideas.[6]

The ambiguity of 'experience'

The idea of non-propositional knowledge represents a reaction to a deep tension in empiricist foundationalism, a tension clearly identified by Wilfrid

Sellars. According to Sellars, the classical empiricist conception of 'experience' conflates two ideas:

> (1) The idea that there are certain inner episodes—e.g. sensations of red or of C# which can occur to human beings (and brutes) without any prior process of learning or concept formation; and without which it would in some sense be impossible to see, for example, that the facing surface of a physical object is red and triangular; or hear that a certain physical sound is C#.
>
> (2) The idea that there are certain inner episodes which are the non-inferential knowings that certain items are, for example, red or C#, and that these episodes are the necessary conditions of empirical knowledge as providing the evidence for all other empirical propositions.[7]

The first idea concerns the sort of 'awareness' or 'experience' that we enjoy simply in virtue of being conscious or 'sentient' beings. The second idea concerns the sort of awareness that is involved in knowledge properly so-called: the sort of knowledge that involves the ability to make propositionally articulable claims. Robert Brandom calls the two types of awareness 'sentience' and 'sapience'.[8] So in Brandom's terms, the traditional empiricist conception of experience as the foundation of knowledge involves a mistaken attempt to reduce sapience to mere sentience.

Talk of sensing sense-data is derived from familiar talk about sensations: headaches, tickles, itches, and so on, the kind of awareness or consciousness we share with animals and prelinguistic infants. The capacity for experience in this sense, the sense captured by the first idea, is indeed primitive, unacquired. However, the sort of knowing involved in the second idea is not plausibly regarded as primitive in this way. Knowing that x is F—for example, that there is something red and triangular in my visual field—requires familiarity with systems of classification, systems that bring 'particulars' under 'universals'. It therefore depends on training in the use of such systems, training that results in our knowing where the boundaries of various descriptive categories lie. It depends on learning a language.

If we want to think of conceptual abilities as independent of the social process of language learning, we must think that such abilities—at least those relating to basic observable facts—are innate. This is something that empiricists have been reluctant to claim. However, this reluctance threatens them with self-contradiction. Sensing a sense-datum, call it 's', is supposed to be sufficient for one's non-inferentially knowing that s is (say) red. But while the ability to sense sense-data is supposed to be unacquired, the ability to know facts like 's is red', which presupposes conceptual mastery, is not.

To say that merely having sensations is not sufficient for knowledge is not to say that it is irrelevant. Experiences can certainly be involved in *causal* relations to beliefs and judgements, as when my having a certain visual experience

causes me to exclaim 'How blue the sky is today'. But causal relations are not what the foundationalist needs. A foundationalist can allow that causal explanation is sometimes epistemically relevant, in that it can excuse me from charges of epistemic irresponsibility. If I hold certain strange beliefs as the result of a brutal regime of indoctrination or brainwashing, I will not be held epistemically accountable in the usual way. But no foundationalist can allow that a purely causal explanation of my holding a belief is sufficient to justify it, on pain of undermining the case for manifestly credible basic beliefs. If causal relations were sufficient for justification, there would be no need to detour through experience. We could say that we have basic, non-inferential knowledge when we stand in appropriate causal relations to objects or events in our external surroundings. This is what contemporary externalists do say. Foundationalists appeal to experience because they want experience to involve a kind of knowledge. However, what they give with one hand they take away with the other by treating that knowledge as 'non-propositional'.

This is not surprising. Propositional content involves conceptual or descriptive content, and description is inseparable from the possibility of misdescription. It seems, then, that propositional content is inseparable from the possibility of error. If this is so, no judgement, however modest, is absolutely indubitable. So if basic experiential knowledge has to be indubitable, there is no such knowledge.

This is another reason why foundationalists have been attracted to the idea of non-propositional knowledge. In thinking of relations like acquaintance as absolutely 'direct' or 'unmediated', philosophers like Russell mean to exclude even the 'mediation' involved in bringing an object under a concept: that is, describing it or thinking of it in some definite way. But this very tactic embodies a tacit recognition that descriptions can apply or not. It is true that by eschewing all description, we avoid the possibility of getting things wrong, but only because we also give up the possibility of getting them right. If you refuse to play the game, you can't lose. But you can't win either.

The content problem

We have been led to what we may call 'the content problem'. The problem is to show how basic beliefs can be indubitable without losing all content, thus ceasing to be beliefs. In my view, it is insoluble.

Basic judgements must be synthetic—answering to contingent facts—yet indubitable. Many attempts to explain how this is possible employ the idea that basic judgements involve a strongly demonstrative element. Thus the content of a typical basic experiential belief would be something like '*This* is (or appears) red *now*'. In thinking 'This', we engage in a kind of 'mental pointing', a deliberate focusing of attention on some aspect of our current experience.[9] Such judgements, while not analytic, are *like* analytic judgements

in that anyone who has mastered the rules for the use of demonstratives and elementary descriptive terms, and who uses those terms in a fully mindful way, cannot judge mistakenly. The only sort of mistake that can be made is a slip of the tongue, or its mental equivalent, thus a 'verbal' not a 'factual' error. Considered misjudgement is impossible.[10]

Behind this account of basic knowledge lies a certain picture of meaning and understanding. In this picture, two sorts of rules or 'definitions' determine the conceptual abilities that go with understanding a language. Some words get their meaning by *discursive* definitions: definitions that link words with other words. Such definitions state analytic truths, such as 'A bachelor is an unmarried male'. But not all words can get their meanings this way. Some meanings must be established by *ostensive* definitions, which set up rules or conventions linking words with extra-linguistic reality. Ostensive definitions apply first and foremost to objects and qualities that we can grasp directly in experience: we learn what things are properly called 'red' by being presented with examples. Of course, the redness in question must be 'phenomenal' or 'experiential': the highest common factor redness that is present to consciousness when something looks red, whether or not it really is red. With this restriction in place, it is clear that one cannot understand a term like 'red', the meaning of which is fixed by direct correlation with something consciously present, and make mistakes when mindfully applying it in basic demonstrative judgements.

We have gone right back to the doctrine of acquaintance. The model of meaning and understanding just presented postulates a kind of grasping of red things as red that is prior to any mastery of linguistic or conceptual systems of classification. The possibility of empirical knowledge depends on there being what Sellars calls 'self-authenticating, non-verbal episodes'. It is questionable whether we even understand this idea.

Still, assume that we do, to some extent. The fact remains that basic demonstrative judgements are not purely demonstrative, but contain a descriptive term, and this threatens to leave the content problem untouched. Foundationalists emphasize the demonstrative element in basic judgements because they want to tie basic knowledge to an immediate knowledge of present facts. They want to do this because they admit that, if a judgement has implications for what must obtain in other situations, it cannot be immune from error and cannot really be *intrinsically* credible. But description in and of itself carries such implications. In thinking of something as 'red', I am assimilating it to the class of objects that I would have thought of as red on other occasions. Foundationalists meet this objection by postulating a 'non-comparative' or 'purely phenomenological' use of terms like 'red'.[11] However, if this use is absolutely non-comparative—devoid of all implications for further applications of the concept—how does 'This is red' differ from 'This is this'? Once more, we have lost all empirical content.

Here we come to the heart of the content problem. The foundationalist's basic judgements have to be 'stand alone' judgements in a very strong sense: judgements that we could entertain, and which would amount to knowledge, without our knowing or even believing anything else. This is what genuinely *first* principles have to be like, and it is this conception of 'first principles' for which Sellars coined the pejorative description 'the Myth of the Given'. The key point is that the foundationalist's epistemic atomism commits him to a kind of semantic atomism: an atomism about meaning and understanding. The difficulty is to see how any judgement could be wrenched out of all inferential connections to further judgements and retain any content at all.

The reliability problem

The quest for certainty forces foundationalists to make their candidates for basic judgements ever more modest in content. But so long as content persists, the possibility of error remains. Eventually, the temptation arises to eliminate propositional content altogether. The price of giving in to this temptation is epistemic irrelevance.

Perhaps the demand for strict indubitability is too strong. Maybe foundational beliefs can be epistemically privileged without being strictly indubitable. One suggestion is to distinguish more precisely between indubitability and incorrigibility. Incorrigible beliefs, we might say, could in principle be false, though we could never be in a position to correct them. This is still a striking epistemological asymmetry: incorrigible beliefs serve as the touchstone for accepting or rejecting other beliefs while remaining themselves immune to correction.

This won't work. If we can go wrong in our basic judgements, how can our errors systematically fail to create problems in our inferential superstructure? If incorrigibility is supposed to be different from absolute indubitability, the difference ought to make a difference. But if errors at the basic level have consequences higher up, why shouldn't they be retrospectively detectable? It is not clear that there is a workable notion of incorrigibility, distinct from strict indubitability.

If we abandon incorrigibility, we begin to drift away from foundationalism towards a coherentist model of justification. In retrospective self-correction, we are trading off our basic against our non-basic beliefs, trying to find an optimum balance. This kind of mutual adjustment is central to the coherence theory's account of empirical justification. This suggests that it is difficult to give up the requirement of strict indubitability for basic beliefs while remaining faithful to foundationalism.

It may not be impossible. According to 'modest foundationalists', the intrinsic credibility attaching to basic beliefs only amounts to *prima facie* justification, which may be inferentially overridden in the context of a developed belief-system. A basic belief is thus always justified unless there is some special reason to suspect that it might be false. This is all we need to block the regress of justification; and it spares us the need to defend an impossible ideal of empirical certainty.

Modest foundationalism offers an attractive prospect, but will it really work? The *prima facie* credibility of basic beliefs is still supposed to be intrinsic. So it must somehow belong to the nature of basic beliefs, however picked out, to be right most of the time. If they are no more likely to be correct than any other beliefs, we lose the basic/non-basic distinction altogether. The question is, what sort of truth is it—necessary or contingent—that beliefs with the defining feature of basic beliefs are more likely to be right than wrong? This is the reliability problem.[12]

Since we are in search of foundations for empirical knowledge, some of our basic beliefs must be empirical. But it is hard to see how it could be a necessary truth that particular empirical beliefs are highly truth-reliable. On the other hand, if our reliability in certain matters is an empirical fact, the sceptic will want to know why we should accept it. As a general claim, he will say, it will need inductive support from particular observations. If these observations differ from the sort of observations whose credibility we are trying to explain, we will need to know why they should be trusted, reopening the regress of justification. If they are the same, we will simply be reasoning in a circle: assuming our reliability in certain matters to underwrite that very reliability. And if we say that it is just obvious that we are reliable reporters on our thoughts and sensations, the sceptic will point out that we are simply making an assumption. Agrippa's trilemma reasserts itself.

One reply is that the reliability problem assumes that the foundationalist is committed to a person's being able to justify the reliability of his basic beliefs. This might be resisted. Maybe it is enough that a person *be* reliable in certain matters. Maybe it is unreasonable to insist on his always being able to account for his reliability. In other words, a foundationalist might deny the K-K thesis, opting for an externalist approach to basic knowledge. This is the externalist version of foundationalism I mentioned at the beginning of this chapter.

From all that has been said, it should be clear that externalism can hold no appeal for a philosopher sensitive to the concerns of traditional foundationalism. It is not just that this kind of reliabilism obviates the need to give 'experience' a distinguished role in empirical knowledge. Rather, to explain non-inferential justification in terms of *de facto* reliability is to abandon altogether the foundationalist quest for a fixed class of intrinsically credible basic judgements. The boundaries of non-inferential knowledge would be fixed by what, as a matter of empirical fact, we find ourselves capable of being trained to respond to. An epistemology like this might contain an element of formal foundationalism, but it would not be substantively foundationalist. (I shall explain all this further in Chapter 15.)

I laid down three explanatory goals for a theory of justification: to give a coherent account of its basic ideas, explaining how they defeat the fatal trilemma; to show how the theory allows for a workably extensive system of justified belief; and to show how knowledge is humanly possible. There is a strong case for thinking that foundationalism falls at the first hurdle.

Notes

1. This formulation is due to McDowell (1994).
2. Sellars (1997), 32 ff.
3. Russell (1969), 113.
4. Davidson (1986).
5. For an important recent discussion of the issues broached here, see McDowell (1994).
6. This objection is urged by Sosa (1980) in the context of responding to what he calls 'the Coherentist Critique of Foundationalism'. I agree with Sosa that coherentists have no business complaining about the intellectualist conception of justification: they are themselves fully committed to it.
7. Sellars (1997), 21–2. The following discussion is deeply indebted to Sellars. For more details on the Sellarsian attack on the appeal to the Given, see Williams (1999*a*).
8. Brandom (1994).
9. Pollock (1974), 74. A classic early statement of this view is by Schlick (1934).
10. In his essay 'Basic Propositions', in Ayer (1954), Ayer endorses the impossibility of mistake at the basic level. However, he changed his mind: see Ayer (1956), 61 ff.
11. See Chisholm (1982), 141 ff.
12. My argument here is much influenced by BonJour (1985), 30 ff.

9

Reduction and Inference

Out of the frying pan

Supposing that a foundationalist could solve the problem of security; could he solve the problem of adequacy? Could he retrieve a workable edifice of empirical knowledge?

The foundationalist picture of a hierarchy of broad classes of beliefs, ordered by general relations of epistemological priority, leads straight to a variety of underdetermination problems. In other words, foundationalism escapes Agrippan scepticism only at the cost of exposing us to various forms of Cartesian scepticism. The problem of adequacy is the problem of how to deal with scepticism in its Cartesian form.

Let us focus on the problem of the external world. The Cartesian sceptic argues that, although knowledge of the world rests ultimately on experiential knowledge, there is no justifying inference—deductive or inductive—connecting what we know on the basis of experience alone with our beliefs about objective reality. This leaves the foundationalist three options. He can:

1. argue that knowledge of the external world is not necessarily inferential but can be 'direct' or immediate. This is the strategy of direct realism;
2. argue that further reflection reveals logical or conceptual links between experiential knowledge and beliefs about the external world. This is the approach of reductionism and the criterial theory;
3. argue that the sceptic underestimates the resources of inductive inference. This is the inductive approach.

We shall examine all three strategies.

Direct realism

In the form in which I am about to discuss it, direct realism does not take issue with the foundationalist picture of knowledge: it simply claims that the foundations can be set at a higher level than the sceptic allows.

It is obvious right away that, with respect to some Cartesian problems, this strategy is not even minimally plausible. Barring telepathy and precognition,

the direct realist approach has no purchase on the problem of other minds or the problem of induction. Conceivably, we might claim that memory puts us in direct touch with the past, though this too has a whiff of the supernatural about it. And in any case, even granting personal memories the status of direct knowledge, we would still be left with an impossibly restricted basis for historical knowledge generally.

The direct realist's most promising theatre of operations is the problem of the external world. Obviously, he cannot plausibly claim absolute certainty for his basic beliefs: we do make mistakes. Still, we also think of ourselves as often 'just knowing' this or that about objects in our surroundings. I know that there is a computer screen in front of me because I can see that there is. It isn't a matter of inference. But in what sense isn't it a matter of inference?

Here we must bear in mind that, *as a version of substantive foundationalism*, direct realism is committed to the existence of non-inferential knowledge in a very strong sense of 'non-inferential'. The direct realist is not just saying that, in everyday situations, our opinions about what is going on around us are generally made off the cuff and accepted without special scrutiny. Rather, direct realism credits certain observational beliefs with intrinsic credibility. Such beliefs constitute foundational knowledge, knowledge that we could have on its own, in isolation from collateral knowledge of the world and our relations to it.

It is doubtful that we ever think of perceptual knowledge as absolutely non-inferential in this way. The problem is not just that we do not credit ourselves with unqualified reliability: our reasons for regarding ourselves as fallible matter crucially here. The extent to which we regard ourselves as trustworthy observers is determined by empirical considerations: it has nothing to do with intrinsic credibility, even if the credibility is only *prima facie*. We realize that we are more or less reliable, depending on a wide range of environmental and psychological conditions. We are good at taking in facts about familiar objects, relatively close to hand, in good light, when we are wide awake and not distracted . . . and so on. If the world were different, or we were different, our reliability as perceivers would be different. We might be better at spotting some things, worse at spotting others. Ordinary cases of 'just knowing' or 'just seeing' are embedded in an extensive body of collateral knowledge about the world and our place in it, not encapsulated in the way that foundationalism demands. (We shall see in Chapter 15 that this embedding is nevertheless compatible with the existence of observational knowledge that is genuinely non-inferential.)

Reductionism

The next possibility to explore is that there are logical links of some kind between privileged and problematic beliefs. According to reductionism,

statements from the problematic class can be translated into logically complex statements from the privileged class. Thus, according to *phenomenalism*, talk about the physical world is really just shorthand for logically complex claims about sensory experiences: claims about the experiences we do or, in various conditions, would have.[1] External objects are, in this sense, 'logical constructions' out of sense-data, actual and possible. Precursors of modern phenomenalism include Berkeley, who identifies objects with collections of 'ideas', and J. S. Mill, who defines external objects as 'permanent possibilities of sensation'.

Like direct realism, reductionism is probably at its best in connection with the external world. With respect to the problem of other minds, the analogue of phenomenalism is logical behaviourism, the thesis that talk about someone else's thoughts and sensations is just a roundabout way of talking about behaviour and behavioural dispositions. However, it is difficult to believe that when I describe someone as being in great pain, *all* I mean is that he is disposed to groan or grimace and to quieten down when given an analgesic. Such behaviour is a symptom of pain, not the pain itself. Similarly with respect to the problem of the past. Surely, talk about the past is more than shorthand for talk about documentary and other evidence existing *in the present*.

These implausibilities notwithstanding, reductionist theories give voice to a powerful idea: the verificationist conception of meaning. On this view, meaning and understanding cannot be detached from verification in the way that the sceptic supposes. The sceptic wants to argue that the world could be wildly different from the way we normally take it to be, even though our experience, the ultimate basis of all empirical knowledge, remained unchanged. Accordingly, we can understand all sorts of stories about how the world might be but can never know how it really is. The phenomenalist replies that we only *think* we understand this suggestion. In fact, we have no clear conception of a 'reality' somehow 'behind' our experience. The world-as-we-experience-it—the world we can find out about—is the only world we can even conceive. The distinction between the objective and the merely apparent features of the world must be one that we draw within our experience. It cannot be a contrast between experience and an ineffable 'reality'.

O. K. Bouwsma once offered a cute variant on Descartes's tale of the Evil Deceiver. In his first attempt to deceive someone, call him 'John', the Deceiver turns the world (victim included) into paper. Naturally, John sees through the ruse: he makes crinkling noises when he moves, is highly inflammable, easily folded up, and so on. So the Deceiver embarks on a series of improvements, eventually getting to the point where the 'deception' is absolutely undetectable. But now where is the deception? What is John deceived *about*? In the world as he experiences it, everything remains the same; and what other world do we care about? What other world is there? What other world can we even imagine?[2]

It is easy to feel sympathy for philosophers who resist the sceptic's attempt to create an unbridgeable gulf between 'experience' and 'the world'. Nevertheless, reductionism expresses this resistance in the form of a bold theoretical programme that is impossible to carry through. It is not just that no one ever *has* reduced an external-object statement to a conjunction of experiential statements: no one *can*.

Laws of experience

The phenomenalist wants to treat external objects as logical constructions out of experiences. However, one of our most salient beliefs about external objects is that they exist independently of being perceived. The bookcases in my study do not cease to exist when I turn my attention to the computer screen. Indeed, there are countless things that no one will ever see.

The phenomenalist deals with this problem by claiming that translating an external-object statement into experiential (or 'sense-data') statements will involve liberal use of *conditionals*. I believe that my bookcases exist when I am not looking at them: what I mean by this is that *if I were* to turn around I *would* get the appropriate as-of-bookcases experiences.

The conditionals in question here are *subjunctive*. Unlike indicative conditionals, which say what follows if such and such *is* the case, subjunctive conditionals say what *would* be true if such and such a condition *were* to hold. In many cases, the relevant conditionals will be *counterfactual*: their antecedents will be false. My bookcases have continued to exist while I have been typing these words. During this time, I have *not* turned to look at them. But if I *had*, I would have enjoyed perceptual experiences of the relevant kind.

Subjunctive conditionals are connected with *natural laws*. When I say that the match would have lit had I struck it, I am relying on there being certain physical and chemical laws having to do with friction, heat, and combustion. Similarly, what the phenomenalist needs for his conditionals to work is that there be laws of experience. Since the phenomenalist wants to reduce *all* external-object talk to experience-talk, he needs *purely* experiential laws: law-like regularities logically analogous to the laws of physics but concerning only relations between experiences, without reference to physical events.

No one has the faintest idea as to whether there are any such regularities. It is true that we think of experience as being regular and predictable; but what we mean by this is that we can predict the sort of experiences we will have *in particular physical and physiological conditions*. If I swivel round in my chair with my eyes open, if nothing falls on my head knocking me unconscious, and so on, then I know how things will appear to me: my bookcases will come into view. But these experiential regularities are object-dependent, whereas the phenomenalist needs object-independent regularities concerning experiences alone. The trouble is, such regularities are vulnerable to disruption by every blink of the eye, turn of the head, lapse of attention, change in the environment, and so on indefinitely.[3]

Although this may not be a knock-down proof that no such regularities exist, it is certain that they would be so complicated that no one knows, or even could know, that there are any. We lack the ability to track experience at the level of detail phenomenalism requires. If verifying an external-object statement were dependent on verifying its phenomenalist translation, knowledge would be humanly impossible, even if the phenomenalist programme were feasible in principle.

A second point is that the phenomenalist has not really followed through on his promise to display a logical link between experiential evidence and external-object statements. Given a phenomenalist translation, the simplest particular claim about an external object implies various *general* claims about experience. At most, phenomenalism would reduce the problem of the external world to the problem of induction. This would not be a negligible achievement. But it would not be a complete answer to scepticism either.

Semantic mass

Is the phenomenalist programme of sentence-by-sentence translation of external-object claims into experiential equivalents feasible even in principle? It seems not. The problem is that single external-object statements do not have *any* experiential consequences. By itself, the existence of my bookshelves implies nothing about my experience, *unless further facts concerning the external world are taken for granted.*

In committing myself to the unperceived existence of my bookshelves, I commit myself to supposing that, if I were to swivel round in my chair, I would enjoy such and such experiences. Of course, this conditional won't do, since this talk about swivelling round involves unreduced reference to events and objects in the external world. The problem for the phenomenalist is that such reference is ineliminable. If it is night and there is a sudden blackout, I won't get the usual experiences; if as I swivel the roof falls in and knocks me unconscious, I won't enjoy any experiences at all, except perhaps briefly seeing stars. If someone has slipped a hallucinogen into my coffee, my bookshelves may appear to melt: *and so on indefinitely*. No external-object statement, *taken alone*, has any particular experiential implications.

This problem is not unique to phenomenalism. Consider logical behaviourism. According to a logical behaviourist, talk about mental states is roundabout talk about behaviour. Not only actual behaviour: just as external objects can exist unperceived, mental states can exist without being acted upon. You believe that it is going to rain this morning. This belief *might* show itself by your remembering to take an umbrella when you go out. But of course, you might stay home, so that you get no chance to express your belief in action.

Like the phenomenalist, then, the logical behaviourist has recourse to subjunctive conditionals. He analyses mental states in terms of behavioural dispositions: what you *would* do if appropriate circumstances *were* to present themselves.

This leaves him with the same problem: an ascription of a particular mental state does not appear to have any behavioural consequences, not even dispositional consequences. You believe that it will rain this morning: does this mean that if you were to go out you would take your umbrella? Yes: if you care about not getting wet, if you intend to walk rather than drive, if you are not in such a hurry that you forget, and so on . . . There is no prospect of appealing to such dispositions to effect a reduction of the mental to the behavioural, since the relevant behavioural dispositions associated with a given mental state are themselves conditional on further mental states.

Once more, foundationalism is frustrated by its epistemic and semantic atomism. For the foundationalist, knowledge rests on encapsulated items of basic perceptual knowledge. Accordingly, if the link with non-basic knowledge is reductive, the reduction must be carried out on a sentence-by-sentence basis, and this is what cannot be done. While many beliefs indeed depend for their justification on confirmation by experience, they have no experiential consequences they can call their own. They have definite experiential consequences only as members of groups of beliefs with critical semantic mass, and even then only subject to vague *ceteris paribus* qualifications. Sentence-by-sentence translation into basic experiential statements is a non-starter.

The criterial theory

The 'criterial' view is less ambitious than strict reductionism. It holds that it belongs to the 'logic' of physical-object statements to be responsive to experiential criteria. In other words, it belongs to the meaning of a physical-object statement that certain experiences count for or against it. Similarly, it belongs to the logic of statements about the inner life of another person that certain outward expressions constitute good evidence for the ascription of particular states. Without such public criteria, we would not even be able to think of other people—and perhaps not even ourselves—as having an inner life.[4]

The criterial theory respects the fundamental idea behind reductionism: that understanding a claim is importantly connected with grasping what counts for or against it. But it avoids the impossible task of showing how full-scale translations of, say, physical-object into experiential statements might proceed. And it has the added advantage of not committing us to thinking that, in speaking of someone's inner life, we are *only* talking in a roundabout way about his outward behaviour. There must be some truth in this. Nevertheless, *as a defence of foundationalism*, the criterial view is a failure.

We are invited to suppose that it is analytically or conceptually true that certain experiences give me good reason to think that there are two bookcases in my study. At the very least, this must mean that having such experiences is a reliable indicator of the presence of bookcases. But just as no single statement about objects in the external world strictly entails that I would have such and such experiences if such and such conditions were to obtain, no single statement entails anything about what experiences are reliable indicators of its truth. At most, we can say that such relations of reliable indication hold 'other things being equal' or 'in normal conditions'. But since these qualifiers point to an indefinite range of further empirical presuppositions, this shows that no analytic evidential connections hold between any single statement about an external object and any definite body of experiential knowledge. The holistic character of the relation between worldly and experiential commitments is as much a problem for the criterial theory as it is for strict reductionism.

Are there analytic truths?

Reductionism and the criterial theory lean heavily on the notion of analytic or conceptual truth. I have been arguing that the connections postulated by these approaches between experiential evidence and statements about the world are not plausibly viewed as analytic. But an even more fundamental problem is that the analytic–synthetic distinction is dubious in itself.

Suspicion of the analytic–synthetic distinction is much associated with Quine.[5] Quine's point is not just that the distinction is vague, like the red–orange distinction, but that it should be rejected altogether. In Quine's eyes, the analytic–synthetic distinction reflects false theoretical views about meaning, just as the witch/non-witch distinction reflects false views about supernatural powers. The fact that adepts of such views think that they can recognize cases is neither here nor there. Nobody is really a witch; and no statement is really analytic.

One reason for being suspicious of the analytic–synthetic distinction is that it depends on an obscure notion of absolute synonymy. It is supposedly a hallmark of analytic truths that their denials are self-contradictory. However, the self-contradiction is not explicitly logical. Compare the tautological sentence 'All male siblings are male siblings' with the allegedly analytic sentence 'All brothers are male siblings'. Denying the first results in the formally self-contradictory claim 'Some male siblings are not male siblings'. But denying the second yields only 'Some brothers are not male siblings'. To be sure, we can turn this latter sentence into an explicit contradiction by substituting 'male siblings' for 'brothers'. The justification for the substitution is that 'brother' and 'male sibling' have 'the same meaning'. But do they? *Alle menschen wirden bruder*: all men will be brothers, though presumably not male siblings. The reply will be that Schiller is not using 'brother' in the 'literal'

sense. But what sense is that? The sense that makes 'Brothers are male siblings' analytic? We are not getting anywhere.

The notion of synonymy needed to explain the analytic–synthetic distinction is not the everyday notion of 'having the same meaning'. That notion is highly contextual and interest-relative. No one denies that *for certain purposes* (legal proceedings involving a disputed will, say) 'brother' can be *treated* as meaning 'male sibling'. But 'brother' stands in complex relations to other words: there are blood brothers, lay brothers, brothers in arms, and brothers under the skin. Prescinding from all purposes and interests, there is no way of deciding when words 'mean the same thing'. Here we see another instance of foundationalism's stake in encapsulated 'meanings'.

Analytic truths are often explained as 'true by definition'. However, the notion of definition is as context-sensitive and interest-relative as that of sameness of meaning. For expository purposes, a scientific theory can be sorted out into definitions and empirical postulates; but if the theory runs into trouble, modifications can be made anywhere. Moreover, this kind of sorting into definitional and empirical elements can be done for theories that are comprehensively false (like witch-theory), so being 'analytic-in-a-theory' does not mean being analytically true. Being designated a 'definition' confers no privileged epistemological status, no immunity from revision. Yet explaining epistemic privilege is the main reason for accepting the analytic–synthetic distinction in the first place.

It is no accident that the examples used to introduce the analytic–synthetic distinction often involve quasi-legal terms like 'bachelor' or 'brother', for which it is at least *prima facie* plausible that there is a single criterion of applicability. Even here, as we just saw, it is doubtful that matters are so simple. But for other terms, particularly natural kind terms like 'gold', the idea of definitionally fixed meanings is not even *prima facie* plausible. Natural kind terms involve 'law-cluster' concepts: we identify gold by a cluster of physical and chemical characteristics, none of which is absolutely essential. Maybe we will rethink the periodic table so radically that we no longer count gold as a metal.

I do not deny that the analytic–synthetic distinction has a certain *prima facie* plausibility. It does seem that certain proposition are 'meaning-constituting' in the sense that, if anyone denied them, it would be doubtful that he understood them. If you understand what it means to say that $2 + 2 = 4$, you see that it is true. To say of such a proposition, 'I know what it means, it just doesn't seem very plausible to me', would be (at best) a bad joke. Our acceptance of this and other elementary propositions fixes what we mean by '2', '4', and '+'. In this way, it is claimed, truth-by-virtue-of-meaning explains *a priori* knowledge.

Quine denies that there is any principled way of sorting things we accept into those that are 'meaning-constituting' and those that are empirical. To be

sure, if someone denies something crashingly obvious, there will always be a question of whether he understands what he is saying. But this provides no reason for supposing that some propositions are true by virtue of meaning alone. Furthermore, even if we agreed on which propositions are (in a common-sense way) meaning-constituting, we would still not have an explanation of *a priori* knowledge because, as already remarked, there is no automatic connection between a proposition's being meaning-constituting—'analytic-in-a-theory'—and its being true. After all, we can understand theories that we know to be false. So while, if we like, we can say that part of what we mean by 'phlogiston' is that it is a substance released by combustion, our commitment to such meaning-constituting claims about phlogiston remains merely conditional. If there is such a substance as phlogiston, it is released during combustion; but since there is no such thing, it is false that phlogiston is released during combustion. We can understand 'phlogiston' without supposing that any claims belonging to the phlogiston theory of combustion are true.[6]

Quine is right about scientific concepts like phlogiston. Postulates governing their use can be meaning-constituting and nevertheless not true. *A fortiori*, they are not true by virtue of meaning alone. But what about logic and mathematics? Quine would say that, since logic and mathematics are both used in science, they too are in principle open to revision. All that matters is getting a theory that is optimal with respect to empirical adequacy and theoretical simplicity. In this sense, even logic is responsive to empirical considerations. It has been suggested that, in quantum mechanics, a revised logic might be part of the best overall theory. Of course, this suggestion may prove unfruitful. But then again, it may not. Not even logic is unrevisable in principle.

Suppose that we did revise logic in the interests of improving quantum theory: would we have to give up using classical logic in non-scientific contexts, such as courts of law? Surely not. Noting this, Paul Horwich suggests that we might end up with two logics: an *a priori* logic involving our familiar concepts of 'not', 'or', and so on, and an *a posteriori* logic involving technical variants of familiar logical notions.[7] However, there is another way to look at the situation: that theories do not have to be unqualifiedly true to be practically preferable in many (or most) contexts. Geocentric astronomy is good enough for navigation, Euclidean geometry for surveying, and Newtonian physics for aiming rockets at the moon. Everyday indispensability does not distinguish logic from more obviously empirical theories.

The doctrine of analytic truth and the doctrine of acquaintance are part of the same theoretical package. Foundationalist empiricism recognizes two source of meaning: some terms derive their meanings ostensively, by correlation with whatever is 'directly presented' in experience. Given this stratum of primitive, encapsulated meanings, other terms can be introduced. This may be

by explicit term-by-term definitions; or it may be by rules for the use of whole sentences, rules that thus implicitly define the terms that these sentences contain. As 'purely conventional', such definitions and rules are true in virtue of meaning alone, hence knowable *a priori*. We have just seen that the second component in this package—analyticity—is in no better shape than the first.

Induction and explanation

We come finally to the inductive approach. Anyone who adopts this approach will concede to the sceptic that simple induction is no help with the problems in hand. But there are more sophisticated forms of inductive inference, notably what is sometimes called 'inference to the best explanation'. On this view, inductive inference is not a matter of reading off hypotheses from observational data. Rather, given a range of data to explain, we formulate various hypotheses, selecting the best in the light of various epistemic desiderata: for example, empirical adequacy and theoretical elegance.[8]

So conceived, inductive justification is far from algorithmic. If competing theories perform well with respect to different criteria, it can be difficult to pick a winner. Sometimes we have to choose between empirical and theoretical virtues: a clumsy and complicated theory may fit the data better than a simple and elegant one. In such a case, we might even go with the less empirically adequate theory, trusting that its empirical performance can be improved over time. But a choice can be justified without being uniquely justifiable.

That inductive inference is not always inference to a uniquely best explanation does not mean that there are no cases in which clear winners emerge; and it is clear that our ordinary ways of understanding the world are far better than the sceptic's trumped-up alternatives. At best, sceptical hypotheses cover the same experiential data. They have no predictive superiority and are vastly inferior theoretically. Even if logically coherent, Descartes's Deceiver hypothesis is thin and underdescribed, compared with our usual account of how experience arises. (Why does he do it, and how?) It is a worse explanation and may be justifiably dismissed.

There is something in this: sceptical hypotheses *are* thin and underdescribed. Moreover, this kind of inductive approach avoids commitment to sentence-by-sentence evidential connections. In adopting the 'hypothesis' of an external world, we are elaborating a whole theory with respect to the course of experience. There is no reason to expect individual statements in the theory to have specific experiential consequences.

These advantages notwithstanding, the inductive approach, as a defence of foundationalism, shares deep theoretical commitments with its rivals. To see this, we need only ask what the 'hypothesis' of the external world is supposed to explain. The answer, which goes back to Hume, is the *coherence* of our

experience. Our experience is not a phantasmagoria, but displays predictable patterns. This is what makes it justifiable to trace experience to our causal interactions with an external world of persisting objects. But this takes us back to square one. The coherence of experience is just its conformity to purely experiential laws. In the relevant sense of 'coherent', no one has the slightest idea whether experience is coherent or not.

Someone might reply that this objection misunderstands inference to the best explanation. In particular, the objection takes for granted an impossibly rigid distinction between observation and theory. In science, observations are not any random facts that investigators happen to have registered. Observations are selected on the basis of theoretical commitments, which means that other facts may be rejected as irrelevant. It used to be thought necessary to explain the exact number of the planets and not just their motions, but after Newton this idea was abandoned. More than this, observations often cannot be made at all unless a certain theoretical orientation tells us how and where to look. For centuries, astronomers puzzled over various irregularities in the behaviour of the planets or 'wandering stars'. But irregularities only count as such given background expectations as to what 'regular' behaviour would consist in. In these and other ways, observation is 'theory laden'. Maybe the same is true with respect to the coherence of experience: we need to think in terms of an objective world in order to appreciate the ways in which experience is stable and regular.

In considering this sophisticated defence of the inductive approach, we must remember the current context of our discussion, which is foundationalism. The whole point of foundationalism is to insist on the absolutely basic character of experiential knowledge. From a foundationalist standpoint, it must at least be possible to identify the facts that the external world 'hypothesis' is supposed to explain independently of adopting the hypothesis. Once we treat our belief in an external world and the recognition of certain regularities in experience as part of a single explanatory package—each component gaining credibility from the other—we have abandoned foundationalism in favour of some form of coherence theory.

Can the coherentist approach succeed where foundationalism fails? That is our next question.

Notes

1. Ayer was one of phenomenalism's best-known and most faithful advocates, although he spent much of his career modifying and weakening his phenomenalist commitments. See 'Phenomenalism', in Ayer (1954), and Ayer (1956), 118–32.
2. 'Descartes' Evil Genius', in Bouwsma (1969).
3. My critical line is inspired by Sellars's essay 'Phenomenalism', in Sellars (1963).
4. See e.g. Pollock (1967). The 'descriptivism' defended in Ayer (1956) strikes me as a related view. The appeal to criteria is closely associated with certain followers of

Wittgenstein, though there is much controversy over what Wittgenstein had in mind. For an argument to the effect that Wittgenstein did not embrace what I here call a criterial epistemology, see John McDowell, 'Criteria, Defeasibility and Knowledge', in McDowell (1998).

5. 'Two Dogmas of Empiricism', in Quine (1961). The analytic–synthetic distinction is the first dogma, reductionism the second. An important reply to Quine is Grice and Strawson (1956). Harman (1969) is still the best exposition of Quine's views about meaning. For recent defences of ideas closely related to the classic logical empiricist approach to analyticity and *a priori* knowledge, see Peacocke (1993) and Boghossian (1996). Peacocke and Boghossian are critically discussed in Horwich (1998*a*), ch. 6.
6. See Horwich (1998*a*) for the best recent discussion of this and related points.
7. Ibid. 146 n.
8. On inference to the best explanation, see Harman (1973), ch. 10. These ideas about inductive inference are applied to Cartesian scepticism in Vogel (1990).

10

Coherence

Radical holism

Whereas foundationalist theories of justification are *atomistic*, coherence theories are *holistic*. For the coherence theorist, there is not question of a belief's being justified all by itself, as the foundationalist's basic beliefs are supposed to be. To be justified, a belief must fit into a justified system; and the system is more or less justified depending on how well it 'hangs together' *considered as a whole*.[1] This reference to whole systems is crucial. Coherence theories do not claim merely that the epistemic significance of a given belief must always be assessed in the context of some further commitments. Rather, they insist that the justification of individual beliefs depends on certain properties of total belief-systems. Coherence theories are *radically* holistic.[2]

The difference between foundationalism and the coherence theory is sometimes explained in terms of two competing models of justifying inference. The foundationalist conceives justifying inference on a *linear* model, in which justification proceeds from given 'premises' to 'conclusions' by justification-transmitting rules. The coherence theorist's holistic model of justification is decidedly *non-linear*. While the beliefs that comprise a given system will be logically interconnected in various ways, these connections are not in themselves relations of justification. Rather, the density of such interconnections contributes to the coherence of the system. Justification, which depends on coherence, is primarily a property of the whole system. Individual beliefs are justified derivatively in virtue of belonging to a coherent total view.

The coherence theory's radical holism can be obscured by a tendency for coherence theorists to think of 'coherence' in two ways. Sometimes a belief is said to be justified if it fits in—*coheres with*—a suitable background system. Since, in this sense, a foundationalist insists on a justified belief's cohering with the relevant set of basic beliefs, this *relational* coherence does not distinguish the coherence theory from foundationalism. The coherence theory's distinctiveness arises from tracing the epistemic status of the background system to the way the entire system fits together. This *systematic* coherence implies radical holism.

Because they are radically holistic, coherence theories reject even formal

foundationalism. This seems to create tension with the phenomenology of justification. Ordinarily, disputes are resolved by finding evidence that is acceptable to all parties, not by assessing anyone's total system of beliefs.

The coherence theorist has a reply to this: ordinary justification, which involves bringing *specific* evidence to bear on a *specific* disputed claim, is *local* rather than *global*. However, specific evidence is only the tip of the iceberg. When such evidence is found credible and relevant, this is always in virtue of a shared system of background beliefs. Usually, as its name suggests, the background system operates tacitly and, in matters of local justification, is simply assumed to be in order. However, when a dispute proves intractable, we may need to make some aspects of the background system explicit, so as to locate the source of the disagreement. More important still, if we are to cope with scepticism, we cannot rest content with a mere assumption that our beliefs at large are justified. The coherence theorist's task is thus to explain what it is for our submerged system to be 'in order', to explain the global justification that local justification takes for granted.

Objections

A recurrent criticism of holistic approaches to justification is that they allow any belief, no matter how absurd, to be 'justified' by surrounding it with suitable supporting beliefs. These can be as absurd as the original belief, for all that matters is that the beliefs in the system support *each other*. Accordingly, there is no distinction between a 'reasonable' belief-system and the delusions of a logically adept paranoid. As one critic famously puts it, a coherence theorist 'must consider arbitrary fairy stories to be as true as a historical report, or as statements in a textbook of chemistry, provided the story is constructed in such a way that no contradiction ever arises'. But since we can 'arrive at any number of consistent systems of statements which are incompatible with one another', the coherence theory 'fails altogether to give an unambiguous criterion of truth' and is therefore 'logically impossible'.[3]

Not everyone will see this line of thought as an objection. It can just as easily be taken as an argument for relativism. Once we abandon foundationalism, the argument will go, we must admit that there is no *external* standard by which to evaluate different belief-systems. If there are different ways of viewing the world, each of which hangs together internally, there are no grounds for saying that one is more justified, or even more true, than another. But while a vaguely coherentist epistemology sometimes lies behind relativism, whether the coherence theory as such is committed to relativism is another question. Most of its philosophical defenders deny that it is.

The charge of 'logical impossibility' depends on framing the objection in terms of truth. However, contemporary coherence theorists respond that they are advancing a theory of justification, not truth. Truth is not a matter of degree: either the cat is on the mat or it isn't. But my belief as to the animal's whereabouts may be more or less justified. So while incompatible views can't all be true, there can be circumstances in which they are more or less equally justified. No theory of justification is committed to giving a criterion that is 'unambiguous' in the sense of always picking a winner.

To be sure, some theorists hold that truth, as well as justification, should be explained in coherentist terms. But they do not identify truth with whatever degree of coherence is present in a person's current system of beliefs. Rather, they explain truth in terms of *ideal* coherence or coherence at the limit of inquiry. It is not so obvious that this fails to yield an 'unambiguous' criterion.

While well taken, these points do not dispose of the difficulty. Since the aim of a theory of justification is to articulate constraints on what we can justifiably or rationally believe, no such theory can allow a person to be justified in believing anything whatsoever. Even if there is not always a unique best system, some systems must be better than others. Failing this, there will be no defence against sceptical hypotheses involving Evil Deceivers or brains-in-vats. The worry behind the argument from fairy stories is that the coherence theory cannot satisfy even this minimal requirement.

On closer inspection, the argument resolves itself into at least two objections. The first is that, since coherence reduces in the end to logical consistency, and since there is no limit to the number of consistent systems of statements, there is no limit to the number of 'justified' systems. Call this the 'many systems' objection. The second is suggested by the contrast between science and fairy stories. In what sense are the latter 'arbitrary' while the former are not? The answer, surely, is that fairy stories are *simply* made up. By contrast, scientific views, even if they originate in wild flights of fancy, stand or fall with *empirical evidence*: facts about the world, ascertained by observation. The coherence theory, because it makes justification supervene exclusively on belief–belief relations, fails to allow for empirical constraint: constraint *by the world*. Call this the 'isolation objection'. Exploring these objections will give us a clearer view of how the coherence theory is supposed to work.

Coherence and explanation

With respect to the many systems objection, an obvious first move is to reject its identification of coherence with mere logical consistency. Thus coherence theorists often claim that coherence is more than 'absence of conflict' between beliefs: it also involves 'positive connections'. These connections may include logical entailments but also, perhaps, relations of inductive confirmation, weaker than strict entailment but stronger than consistency. To indicate this kind of epistemological connectedness, coherence theorists often speak of our system of beliefs being held together by 'inferential connections' between its

elements. But many theorists also lay great emphasis on the existence of *explanatory* relations: our beliefs should not just be compatible, they should hang together theoretically. However, on a standard view of explanation, confirmation is the inverse of explanation. Newton's laws, together with certain initial conditions, predict and explain the movements of the planets. In turn, the details of those movements confirm Newton's laws. Identifying coherence with explanatory coherence therefore implies epistemological connectedness as well.

We can relate these ideas to the distinction between local and global justification. The logical, epistemological, and explanatory relations between individual beliefs are precisely what we exploit in 'linear' local justification, where we back up a given belief by citing other beliefs. However, these belief–belief relations are also significant holistically in that, the denser they are, the more a given system is a genuine system and not just a ragbag of unrelated opinions. With richer internal connections, a belief-system is more coherent, thus from a global standpoint more justified.

Coherence theorists typically emphasize a second way in which coherence goes beyond mere consistency: scope or 'comprehensiveness'. A system is more coherent the more it takes in, the greater the range of facts it records, explains, and allows us to anticipate. This is an important corrective to the emphasis on systematicity. Obviously, one way to keep one's outlook tightly integrated is to narrow one's view, ignoring or denying inconvenient facts. But this would not be a rational procedure and would not produce an optimally coherent system of beliefs. The goal of inquiry is to understand as much as possible. This means taking in as much as possible and integrating it as best we can, using the fewest theoretical primitives, minimizing reliance on *ad hoc* hypotheses, establishing connections between our views on various topics, and so on. According to coherence theorists, we are always tacitly modifying our total view, 'making it more complete, less *ad hoc*, more plausible'.[4]

Many coherence theorists also insist on a principle of conservatism: faced with a problem in our belief-system—for example, a previously overlooked contradiction—we look for the least change that will remove the difficulty. However, conservatism is better seen as a consequence of the comprehensiveness requirement than as an independent aspect of coherence. Keeping our belief-system as comprehensive as possible means engaging in damage-limitation when problems arise. For example, we avoid taking on board beliefs that would force us to jettison large chunks of an otherwise functional system. Not that we *cannot* take such beliefs on board: epistemological conservatism is not blind traditionalism. We can accept some loss of comprehensiveness for a large gain in theoretical integration, provided that the loss is not so great that the system becomes less usable. The coherence theory is 'conservative' only in that the need for a useful system gives comprehensiveness a certain priority over integration.

If the coherence theorist is granted these points, a lot of steam goes out of the many systems objection. In its 'fairy story' version, the objection does not even deal with workably comprehensive belief-systems, concentrating rather on individual stories or theories. Clearly, a system that included rich historical and scientific beliefs would be much more comprehensive than one that didn't, so such lore cannot simply be disregarded. But with such a system in place, no 'arbitrary fairy story' would be locally credible. Even to make such stories consistent with a workable system of scientific beliefs we would need a large number of *ad hoc* postulates explaining why fairy-doings are exempt from natural laws. With these hypotheses in place, the fairy-believer's system might perhaps be somewhat more comprehensive than that of his scientific critic. But it would also be much more *ad hoc*, thus on balance less coherent.

These ideas are readily extended to paranoid delusions and sceptical hypotheses. Take Descartes's Evil Deceiver story: it offers no positive advantages over our normal view. There is nothing that it explains that our standard view leaves out. On the contrary, to the extent that we could believe it at all, we could still only get through life by relying on the Deceiver's *simulacra* of ordinary facts. Like the paranoid's delusions, this sceptical hypothesis can accommodate—after the fact—anything that our ordinary view turns up. But it does not allow us to anticipate anything that our standard view does not. At the same time, while offering no gains, the hypothesis entails clear losses. We don't know *why* the Deceiver is up to his tricks. Of course, we could speculate, but nothing in the story suggests elaborating it one way rather than another. A Deceiver-based world-view would be more complicated, more *ad hoc*, but no more functional. Sceptical hypotheses, even if logically possible, may justifiably be dismissed. In this way, the coherence theorist sheds light on the feeling that sceptical hypotheses, if not conclusively eliminable, are nevertheless idle.

This coherentist argument avoids the objections that sink the foundationalist's appeal to inference to the best explanation. That appeal requires holding that experience is 'coherent' in the sense of subject to purely experiential regularities. For the true coherentist, by contrast, our beliefs about the world and about experience are part of a single, integrated system. The true coherentist's appeal to explanatory coherence does not presuppose that there are facts to be explained that we can identify independently of all more 'theoretical' views. Rather, recognizing relevant facts and elaborating theories to explain them are two aspects of a single process: making our total view as coherent as we can.

Coherence and observation

Let us now turn to the isolation objection: that the coherence theory, by reducing justification to belief–belief relations, 'cuts justification off from the world'. Because our discussion so far has only enriched our understanding of

the relations the coherence theorist has in mind, it does nothing to alleviate this worry. But what worry exactly? It is all too easy to become haunted by a vague picture of our belief-system's floating clear of reality, like a helium balloon cut loose from its moorings. We need a sharper formulation of the objection than this.

A better way to frame the objection focuses on the thought that genuine justification must involve some element of external constraint. This is not just a point about knowledge narrowly conceived, but about meaning. If justification were purely dialectical, our 'beliefs' would be no more than counters in a self-contained game. They would not be about the world at all.[5]

Granted that empirical knowledge must be responsive to what goes on in the world, how can this be? Events in the world can *cause* us to have various beliefs, but causation is not justification. Justification requires logical rather than causal relations. This is the point of the slogan that only a belief can justify a belief. How, then are we to distinguish empirical knowing from the weaving of fantasies? Indeed, how are we to distinguish genuine believing from making moves in a self-contained game? Appeal to comprehensiveness and systematicity don't help, for they require only that our fantasies be more elaborate than the average fairy story: that the game be complicated. What we want is for our belief-system to respond to input from *outside*.

The coherence theorist does not deny the importance of external input. We have, he will say, all kinds of *cognitively spontaneous beliefs*: for example, those beliefs that we think of as 'perceptual'. (These beliefs need not be thought of as *sotto voce* verbalizations: they may be carried by experiences, so long as experiences are thought of as involving propositional content.) But these spontaneous beliefs are subject to assessment in the light of a complicated array of general beliefs about our abilities as observers. Each of us knows that his reliability as an observer is apt to vary. We do better in good light, stone-cold sober, than at dusk having downed one aperitif too many. For the coherence theorist, we need both spontaneous beliefs and epistemic beliefs to regulate them. Call this the 'rationalized input requirement'. This is a *material* constraint on coherence, supplementing the *formal* constraints discussed in the previous section.[6]

It is important to see that, for the coherence theorist, input *must* be rationalized. There would be no need to appeal to coherence if we allowed that external constraint could be the result of cognitively spontaneous beliefs that are merely *in fact* reliably caused by the circumstances they represent. To allow this would be to desert the coherence theory for an externalist version of foundationalism. For a coherence theorist, our causal relations to the world are epistemically relevant only to the extent that they are represented in our belief-system or total view. Epistemic or 'reliability' beliefs are an essential component in the coherentist's theoretical package.

The need for epistemic beliefs further explains why the coherence theory

does not encourage blinkered adherence to entrenched views. We might think that, given the desirability of comprehensiveness, it is always rational to reject observations that undermine successful theories. Indeed, it *can* sometimes be rational to do this. For example, if one group of investigators gets experimental results that run counter to a successful theory, but which resist replication, we may be entitled to conclude that their experiment went wrong, even if we are not sure how. But the coherence theory does not license general insouciance about observational evidence. Dismissing otherwise irreproachable observations, because it amounts to an implicit rejection of vitally important epistemic beliefs, can do *more* damage to our belief-system than jettisoning long-standing theories. We cannot make a policy of ignoring consistently replicable results *solely* on the ground that they threaten some favourite views.

The presence of reliability beliefs in belief-systems gives particular observations much more weight than we might have supposed. This is important for the coherence theorist because it undermines one of the foundationalist's strongest and most appealing intuitions: that there is an asymmetry between theory and observation. Because observation has more power to undermine theory than theoretical commitments have to discount observation, we may feel that there must be something to the foundationalist's idea of epistemic priority. Recognizing entrenched epistemic beliefs allows the coherence theorist to simulate the observation–theory asymmetry without committing himself to the idea that perceptual beliefs enjoy an absolute epistemic privilege.

Where foundationalism postulates a fixed observational vocabulary, the coherence theory allows our theoretical views to modify our ideas about what is observable. This is a plus. Armed with enough theory, we can see electrons in a cloud chamber. Special training enables some people to see things others miss: it takes skill to use a microscope. Conversely, we can become sceptical of past observation-claims. Where the witch doctor sees spirits, we may find the effects of hallucinogenic plants. Where some see 'recovered memories', closer investigation reveals irresponsible therapists working on distressed and suggestible patients. Reliability beliefs are no more absolutely privileged than the perceptual beliefs they regulate. However systematically significant, they are open to revision.

The isolation objection is not immediately fatal to the coherence theory.

The coherentist outlook

At first sight, the coherence theory offers a very different picture of justification. Where foundationalism is atomistic, the coherence theory is holistic. Foundationalism is committed to encapsulated knowledge; the coherence theory precludes anyone's having only a single justified belief. Foundationalism

takes seriously the idea of building up a system of justified beliefs one element at a time, starting with a first item of knowledge. For the coherence theory, a belief-system must be acquired more or less whole.

The coherence theory's holism is semantic as well as justificational. The foundationalist picture sees meaning on a 'definitional' model. Primitive terms get their meaning by being defined 'ostensively'; others can then be defined discursively. The meanings of individual terms, or words, is primary; that of sentences is derivative. By contrast, the natural way for a coherence theorist to understand meaning is functionally or inferentially. A belief derives its meaning or content from its inferential connections to other beliefs—connections that reflect its inferential role in a larger system of beliefs—rather as a connecting rod is the engine-component it is in virtue of how it works together with other components to constitute a functioning engine. Sentences, which express the meanings of complete judgements or beliefs, are primary; word-meanings are abstractions from patterns of sentence-use.

Like foundationalism, the coherence theory is committed to a strongly justificationist conception of rationality. At the same time, the coherentist model of justification—hence of rationality—is less constraining than its foundationalist rival. For the coherence theory, justification always involves a trade-off between systematicity and comprehensiveness. In theoretical matters, given the importance of epistemic beliefs, particularly those governing observation, this will often take the form of weighing theoretical neatness against empirical adequacy. Secondly, because the coherence theory accepts that beliefs are always, if only implicitly, evaluated in large groups, there is never a unique solution to a problem. If, in testing a particular theory, we find our expectations disappointed, there will be different ways of making adjustments to our beliefs so as to iron out our difficulties. Maybe our theory is false in its fundamentals; maybe there are problems with ancillary theories, concerning our experimental set-up; maybe the apparatus malfunctioned; or maybe we just misread the dial. Sometimes further tests offer guidance, though sometimes they don't. Most importantly, in no case is there a meta-criterion telling us exactly how much weight to give to the different factors in coherence or which beliefs to revise in order to accommodate recalcitrant observations. The coherence theory is compatible with a range of epistemological temperaments, all equally rational.

Does this mean that the coherentist outlook is relativistic after all? Not obviously. Relativists see individuals or 'cultures' as trapped in self-enclosed systems. Sophisticated forms of the coherence theory challenge this picture. Coherentism, because of its insistence on rationalized input, offers an essentially *dynamic* view of belief. Belief-systems change under pressure to accommodate input that those who hold them do not fully control. Also, the theoretical environment changes with the emergence of new ideas; and no one can anticipate all the ways in which original thoughts may provoke revisions

in how we see the world. Accordingly, belief-systems are not stable enough to constitute the self-enclosed systems or 'perspectives' of the dogmatic relativist. At any given time, there will be issues that nobody knows how to resolve. But we cannot conclude from this that they are intrinsically irresolvable. This is a plausible defence, though not the last word.

The coherence theory is unfriendly to many of foundationalism's favourite *demarcational* ideas. A coherence theorist cannot recognize any fundamental distinction between the 'demonstrative' and 'inductive' sciences. Everything, even mathematics and logic, owes its acceptability to its playing an important role in the overall economy of our belief-system. In consequence, there is no deep, principled distinction between *a priori* and *a posteriori* knowledge. In principle, anything we accept can be modified, if this would make our overall belief-system function better.

The distinction between the necessary and the contingent is also problematic. As we shall see, coherence theorists are under pressure to give a coherentist account of truth, as well as of justification. Roughly, truth becomes idealized justification. This view does not encourage a 'two sorts of truth' position. Older coherence theorists sometimes put this by saying that all truths are necessary, but their point is best taken as a way of resisting the necessary/contingent distinction.

This epistemic conception of truth has often been connected with the idea of degrees of truth. For foundationalists—or for advocates of the correspondence theory of truth generally—a belief or proposition is either true or it isn't. While justification can come in degrees—a belief can be more or less justified—truth is all or nothing. But if truth is understood epistemically, it loses this all-or-nothing character. Some readers may find this an implausible consequence of holism: surely, either the cat is on the mat or it is somewhere else. How can simple factual claims be more or less true? However, we should remember here that holism affects our understanding of meaning as well as our conception of knowledge. On the functionalist or inferentialist understanding of meaning that the coherence theorist adopts, enlarging and revising the inferential connections between our beliefs changes their meanings. As our beliefs evolve, our concepts evolve too. Our beliefs become truer as the concepts they involve become more adequate. We don't mean by 'tree' quite what our animist forebears meant by the word they used to talk about oaks, limes, and beeches. So even their simple factual claims, being inferentially freighted in ways we regard as misguided, were not wholly true. This is not a self-evidently foolish view

From what has just been said, it should be clear that holism is hostile to the analytic–synthetic distinction. We may choose on occasion to regard some things we accept as 'true by definition' or 'true in virtue of meaning', thereby distinguishing them from empirical or factual commitments. But this designation confers no epistemological privilege. If a definition combines with other

views to generate problematic consequences, it is as open to revision as anything else.

So much for the foundationalist's main internal boundaries. What about external boundaries? Foundationalism, in its dominant empiricist form, tends to downgrade 'value-judgements' of all types as not really expressions of knowledge. For the coherence theory, matters are not so clear cut. Consider an influential account of how we achieve a considered moral outlook: we attempt to formulate general principles in the light of reactions to particular cases. But once in place, our tentative principles may modify our intuitive judgements. The quest for a considered moral outlook is a search for 'reflective equilibrium' between general principles and particular judgements, with both categories open to revision. This looks very like a particular application of how the coherence theorist understands reasoning generally. It is not at all clear that a coherence theorist does or should recognize deep distinctions between empirical or 'scientific' and moral reasoning.

However, while the coherence theory is hostile to tying demarcational concerns to overly rigid conceptions of empirical reasoning, it is not so obviously inimical to demarcational projects as such. It is compatible with 'explanationist' approaches to demarcation.

A belief-system is more coherent the fewer independent sub-systems it contains. Coherence is also increased by epistemological self-understanding: our understanding of why we accept the things we do. Indeed, given the crucial role of epistemic beliefs, some measure of such understanding is an indispensable component of a coherent view. Now in some cases, the things that my beliefs are about play an indispensable role in the best explanation of why I believe it. If I look out of my window and notice my dog on the lawn, the fact that my dog is on the lawn plays an indispensable role in the best explanation of why I believe that he is. Moral judgements may be different in this respect. I regard a certain type of behaviour as utterly heinous. Is this best explained by the fact that such behaviour is utterly heinous? Or do other factors—my upbringing, my temperament, local customs—sufficiently explain my view, without bringing in moral facts at all? If they do, we may be able dispense with irreducible moral facts. This will increase the coherence of our overall view by reducing the number of independent sub-systems and increasing our epistemological self-understanding. Whatever the merits of this idea, it is not excluded by a coherentist outlook in epistemology.[7]

Notes

1. A coherentist approach to knowledge (and truth) is evident in British Idealism. See e.g. Bradley (1914) and Blanshard (1939). There are clear coherentist elements in Quine (1961), though whether he is a fully paid-up coherence theorist is a difficult question. The same goes for Davidson: see his 'A Coherence Theory of Truth and Knowledge' (1986), in Lepore (1986), but also his 'Afterthoughts' (1987), in Malachowski (1990). Harman (1973) and (1986) develops a number of coherentist ideas in very interesting ways. A recent sophisticated defence of the coherence theory can be found in BonJour (1985). Lehrer is another prominent contemporary coherence theorist, or is widely regarded as defending a coherence theory: see his 'Knowledge Reconsidered', in Clay and Lehrer (1989). However, Lehrer's views have pronounced affinities with what I call 'contextualism'. (See Ch. 14 below.) Bender (1988) is a useful collection of critical essays on the views of BonJour and Lehrer.
2. This claim is mildly tendentious, since some philosophers use 'coherentist' to describe any epistemological outlook that is not (substantively) foundationalist. I think that my usage is both historically correct and, because it allows for a clear distinction between coherentism and contextualism, theoretically well motivated. My picture of the coherence theory owes a great deal to BonJour who, it seems to me, is admirably clear about the theory's fundamental aims and commitments.
3. Schlick, in Ayer (1959), 212–16.
4. Harman (1973), 159.
5. McDowell (1994), Lecture 1.
6. This response is essentially BonJour's. See BonJour (1985), ch. 6. BonJour, in turn, is drawing on ideas from Sellars (1997). For a more extended treatment of the coherence theory of justification, see Williams (1996*a*), ch. 7.
7. Harman (1977), ch. 1.

11

The Myth of the System

Do we have a total view?

The coherence theory looks like an attractive alternative to foundationalism. It builds on ideas that are direct responses to foundationalism's foundamental problems; and it offers a way of taking justification and rationality seriously without pressing them into a rigid mould. On closer examination, however, these advantages disappear. The coherence theory is radically holistic: it ties justification to the comprehensiveness and systematicity of our whole system of beliefs or total view. How clearly we understand this talk of total views is questionable in the extreme.

There is something fishy about the coherentist image of a 'web of belief'. On this picture, a proposition is either in the system or out of it. This seems unrealistic. We have different styles of acceptance. For example, we might accept a certain theory as a promising working hypothesis, though we would not say that we believe it to be true. Like much traditional epistemology, the coherence theory places too much emphasis on belief as an all-purpose attitude of acceptance. Belief is a rather special normative attitude: a relatively unrestricted form of commitment. Not everything we accept (in some circumstances, or for some purposes) is something we unqualifiedly believe. It is not clear that we have a single acceptance-system, active in all contexts of inquiry.

A coherence theorist may deny that he treats belief in a simple all or nothing way. An important aspect of the coherence theory is that beliefs have different degrees of entrenchment. Some beliefs, because of their centrality and wide-ranging explanatory or inferential significance, cannot be rejected without severe damage to the system at large. Since such damage is prohibited by the comprehensiveness requirement, beliefs like this enjoy a high degree of immunity to revision. By contrast, beliefs that are only loosely integrated into the system can easily be rejected or modified. Because beliefs vary widely in their functional importance, the fact that a belief is either in the system or out of it does not imply a uniform acceptance-status.

This is true but does not really meet the objection. Differences in styles of acceptance cannot be assimilated to differences in degrees of entrenchment.

Take the case of accepting something as a working hypothesis or for the sake of argument. In its appropriate context, a proposition accepted in this way may be treated as definitely correct in the sense that it is (temporarily) exempted from scrutiny. In context, it may be deeply entrenched without being, in any unqualified way, believed. Indeed, we may accept something as a working hypothesis even when our long-term purpose is to falsify it. Karl Popper thinks that, in science, we are always subjecting our hypotheses to potentially falsifying tests. For this reason, he holds that 'belief' is a poor term for our attitude towards even the most successful scientific theories. Anyway, if there really are different styles of acceptance, and if an item's acceptance status varies from context to context, there may be no yes-or-no answer to the question of whether a given item is in or out of the system.

To be sure, some things we know play a role in larger, functionally useful systems of commitments: there is a distinction between useful knowledge and mere information. But this distinction does nothing to encourage us to think in terms of a single super-system or total view, rather than a congeries of relatively independent sub-systems.

Coherence theorists often speak of a belief-system's being held together by 'inferential' or 'explanatory' connections. But the idea of such relations tying our beliefs together into a single system is evidently fictional. I could lose an enormous amount of my scattered historical knowledge without much impact on what I know of physics or mathematics. Our 'belief system', if we should talk this way at all, shows a considerable degree of *modularity*: we don't have a single system, but an indefinite number of sub-systems that are at best loosely connected. These sub-systems are always to some degree porous. Developments in the physical sciences—new methods for dating ancient artefacts or for mapping the traces of vanished settlements—may lead to serious revisions in our historical views. But developments like this can be understood without invoking the idea of a total view.

A coherence theorist will see modularity as an indication that our present system of belief is not as coherent as it ought to be. The drive for greater coherence is what keeps inquiry moving. However, while it is true that knowledge is sometimes advanced by bringing together fields of investigation that were hitherto separate—the rapprochement between the biological and physical sciences has given us molecular biology—this point gives little support to the ideal of total integration. Even in the sciences, progress can take the form of proliferation as much as that of integration or reduction; and in any case, there is more to life than science. Why should political theory ever have much to do with quantum physics, or pet care with parliamentary history? No reason. The coherence theory builds in a metaphysical bias towards monism: the idea that everything we know should somehow form one massive 'complete theory of everything'. It ignores the fact that the progress of knowledge also creates new forms of inquiry, capable of standing on their own two feet.

These objections are just preliminary skirmishes, intended to suggest that coherentist metaphors are less than compelling. There are more theoretical problems to be dealt with.

Logic and inference

According to coherence theorists, belief-systems are held together by 'inferential' and 'explanatory' connections. It is far from clear that the existence of such connections is compatible with radical holism.

Recall that the coherence theory is supposed to embody a non-linear model of justifying inference. On this model, although the beliefs in a given system will be interconnected in various ways, these connections are not directly justificational: that is, they are not relations in virtue of which one particular belief or set of beliefs justifies some further particular belief. Interconnectedness among its components, together with its comprehensiveness, determine a system's degree of coherence. Justification, as coherence, is therefore primarily a property of whole systems. Individual beliefs inherit their justification from belonging to a coherent total view. On this model, inference is not the passage from beliefs to other beliefs along lines of inferential connectedness: it is the holistic selection of a particular total view over less coherent alternatives. All justifying inference, even so-called local inference, is implicitly holistic, that is, global.

As Gilbert Harman has argued, if there were such a thing as genuinely linear inference, it would follow so-called 'rules of inference', such as Modus Ponens:

If P then Q.
P.
∴ Q.

Because we often do seem to make inferences in accordance with this rule, it is tempting to read the rule as saying that given 'P' and 'If P then Q' one may (or perhaps must) infer 'Q': a paradigm of linear inference. But according to Harman, it is a mistake to read this or any other logical inference-rule this way. We must remember that an equally good 'rule of inference' is Modus Tollens:

If P then Q.
Not Q.
∴ Not P.

This rule offers a sharp reminder that inference is not always either a matter of drawing the deductive consequences of beliefs one already has or of accepting beliefs that one's current beliefs inductively support. If we find that our current beliefs imply absurd conclusions, we should revise our current beliefs. Modus Ponens may simply provide an occasion to apply Modus Tollens:

nothing in logic tells us which rule to apply in particular circumstances. These 'rules of inference', then, are really rules of entailment. They tell us that a certain conclusion follows logically from the given premises. They do not tell us what to infer from what, if 'inference' is understood as the process of adding to or subtracting from one's system of beliefs.[1]

The fact that premises may need to be re-examined in the light of what they imply appears to support the coherence theorist's view that, even when inference seems to proceed linearly, we are really evaluating the comparative acceptability of whole systems of beliefs. When we accept the logical consequences of some particular beliefs, we are not really 'inferring' conclusions from given premises. Rather, we are judging that continuing to accept the beliefs in question, together with the newly recognized consequences, gives a more coherent end-result than rejecting one or more of those beliefs so as to avoid their implications. Linear inference requires premises that never need to be re-examined, no matter where they lead us: foundationalism in its strongest form.

For a philosopher who accepts this line of thought, it makes no sense to think of belief-systems as coherence theorists often want to think of them: as held together by inferential connections. For there to be inferential connections linking beliefs to other beliefs, our belief system would have to be structured so that some beliefs function as fixed points, while others are more or less up for grabs. This is exactly what coherence theorists, as radical holists, deny.

Explanations and interests

When coherence theorists speak of a total view's being held together by explanatory coherence, they have the following picture in mind. Our belief-system or total view contains a myriad of beliefs about particular matters of fact and a smaller, though still extensive, array of general or theoretical beliefs that explain or make sense of them. The greater the range of particular facts we can explain, and the smaller the number of primitive theoretical commitments we employ in explaining them, the greater the explanatory coherence of our system at large. However, as we just saw, at the global level our beliefs are held together only by logical relations, particularly implication or entailment (and logical relations of confirmation or probabilification, if there are such). Thus the coherence theorist needs explanation to be intelligible in terms of such relations.

There is an account of explanation to hand. According to the 'deductive-nomological' ('D-N') model, which used to be standard, an event or state of affairs is explained by showing that the statement that it occurred or obtains is deductively entailed by one or more natural laws together with statements concerning the surrounding conditions. Schematically:

(E) $L_1, L_2, \ldots L_n$ (Law-statements)
$C_1, C_2 \ldots C_k$ (Statements of conditions)
——— (Deductively entail)
E (Statement that the event to be explained occurred.)

For example, given Galileo's law of falling bodies and information about how long a stone has been in falling towards the Earth, we can deduce—thus explain—the distance the stone has travelled. On this account, explanation is close to being a logical notion. To characterize explanation, we need only the logical concept of entailment and the (supposedly) near-logical concept of a natural law. But although it promises coherence theorists just what they need, the D-N model is problematic. Few philosophers, if any, defend it today.[2]

One objection is that the conditions laid down by the D-N model are not even necessary features of explanation. For example, we can explain particular events by other particular events without invoking laws. The garage window broke because it was hit by a cricket ball. No doubt, physical laws are involved in cases like this; but I don't have to know what they are in order to explain how the garage window got broken. Furthermore, even if I did know, I still wouldn't be able to give a full-dress D-N explanation since I would never be in a position to determine all the relevant conditions: the precise strength of the glass, the mass, velocity, and angle of impact of the ball, and so on.

Someone might reply that my explanation is still D-N, but vague and approximate. If a ball hits with *sufficient* force a *sufficiently* brittle object, that object will shatter; in this case the force and so on were sufficient, therefore . . . But this statement of laws and conditions adds nothing to what we already know: that in this case, the ball broke the window. At most, it indicates a commitment to there being laws, though we may not know what they are.

This all seems correct to me. Still, it may not be such a fundamental criticism, since it allows that the D-N model may be a correct account of *ideally complete* explanation. A more interesting line of attack is that the D-N model does not state sufficient conditions for explanation: that is, that we can have a D-N 'explanation' and still not have an explanation (or have only a *very bad* explanation—the line is not a sharp one).

Here is a simple example, adapted from Hilary Putnam.[3] You notice me working in my office at six p.m. and wonder why. The following explanation is proposed:

> No object could have exited Williams's office in Δt seconds (a very small time interval) without exceeding the speed of light, which is physically impossible.
> Williams was in his office at Δt seconds to 6 p.m.
> Therefore:
> Williams was in his office at 6 p.m.

Putnam would say that this a very bad explanation and perhaps not an explanation at all. The 'explanation' cites things you can be presumed to know already; and if, knowing these things, you still feel the need for an explanation, they must not address what you want explained. Most likely, you wanted to know what I was up to: you wanted an intentional rather than a physical explanation. This gets us somewhere, for it acknowledges that there are different kinds of explanation; but not all the way, since it allows that the explanation given above is a good one, just not the sort that was wanted.

The deep problem with the D-N model of explanation is that to 'explain an event' is always to explain it *as described in a particular way*. We never just 'explain an event': we explain something *about* it, focusing on some features and ignoring others. In our original example of the falling stone, highlighting the stone's speed of descent tacitly projects us into a context where the focus of interest is the physics of motion: the question is why the stone is *falling*. If the question had been why a *stone* is falling rapidly towards the Earth, the explanation might have been that the castle's defenders had to resort to dropping stones on the attackers' heads, having run out of boiling pitch. What counts as a good explanation depends essentially on interests and background knowledge. Explanation is context-sensitive in ways that 'logical' accounts, like the D-N model, fail to capture.[4]

Someone might reply that the fact that we typically focus our explanatory demands on certain aspects of a given situation is only of pragmatic significance and has nothing to do with the *logic* of explanation. Some things—like my failure to exit the office at warp speed—do not 'call for explanation' because we already know what the explanation is. The explanation is *contained in* our background knowledge. It is not, however, *relative to* a context of interests and presuppositions.

This response underestimates the contextual sensitivity of explanation. Consider again 'explaining' my presence in my office by the physical impossibility of my exiting within the stated time constraints without exceeding the speed of light. Putnam's intuition—that this is a very bad explanation, if it is an explanation at all—needs correction. The correct response to this example is that we have *no idea* whether it presents an explanation, still less whether the explanation is good or bad. The reason that the example strikes us as a bad explanation is that we effortlessly project ourselves into the sort of context in which questions about my staying in my office would typically come up; and in such a context, the appeal to physical impossibility would be completely beside the point. But explanations like this are not always beside the point. Suppose that I were charged with a crime: if I could establish that, just before the crime was committed, I was so far away that getting to the scene of the crime would have been physically impossible, I would have a watertight alibi. This would be a very good explanation for why I was not the criminal. Abstracted from a context of questions, interests, and background knowledge, a D-N 'explanation' is *just a deduction*. But to move to the level of our total

view, where all beliefs are up for grabs, is automatically to abstract from all particular contexts of explanation, which depend essentially on there being all sorts of things that are not in question. At the level of a total view, explanation disappears.

Foundationalism in disguise?

While the coherence theory claims to articulate a distinctive, 'non-linear' conception of justifying inference, it may not be the radical alternative to foundationalism that it purports to be.

Let us return to the isolation objection: the thought that the coherence theory cannot allow for genuine external constraint. We met this objection by allowing for rationalized input: 'cognitively spontaneous' beliefs regulated by various epistemic beliefs. However, this apparently innocent talk of 'cognitive' spontaneity masks a crucial ambiguity. We might give it a purely psychological reading: cognitively spontaneous beliefs are conceptually structured experiences, or inclinations to judge, that arise without any conscious process of inference. But this isn't what we want: not every spontaneous thought or daydream is something that amounts to 'input' constraining our belief-system. So perhaps we should say that sufficiently vivid or powerful experiences or inclinations to judge are what matter. This is the ambiguity I just spoke of. Our input from the world is supposed to be more than psychologically vivid: it is supposed to be epistemically, hence normatively, significant, that is, something that we *ought* to take account of and which we can be criticized for ignoring. But then the question is: what is the source of this normative significance? If it is intrinsic to experiences or perceptual beliefs, we have abandoned the coherence theory proper for modest foundationalism: the view that basic beliefs have intrinsic *prima facie* credibility that can be strengthened, and occasionally undermined, by how their incorporation into belief-system at large affects overall coherence. But if they have no intrinsic credibility—if their epistemic significance is wholly derivative from our epistemic beliefs—all constraint is ultimately dialectical: a matter of how beliefs fit together, with no constraint that is seriously external to our way of looking at things. To avoid collapsing into foundationalism, the coherence theory must stay true to its radical holism, thus giving up on the idea of external constraint.

Even taking this line may only put off the evil day. According to radical holism, no beliefs are epistemically privileged. There are no fixed points: in any inferential situation, everything we believe is (theoretically) up for grabs. But if everything is up for grabs, even our criteria of coherence, there is simply no saying—at the time, anyway—why we make one inference rather than another. Coherence theorists are hostile to the idea that any knowledge is *a priori*: even basic logical principles have to pay their way by contributing to

the smooth functioning of a total view. Nothing is immune from revision. But at the very least, this has to be an oversimplification. Belief-systems derive their degree of coherence from the character of the logical relations between their component beliefs. Furthermore, a basic feeling for logical relations is the prerequisite for following any justificatory or evaluative procedure, including the application of criteria of global coherence. How, then, can everything be up for grabs?

This question is particularly pressing with respect to the criteria of coherence themselves. Why not subject them to the same questions that get directed against logical rules? Supposedly, we cannot see inference as regulated by rules like Modus Ponens because the question always arises as to why we accept the consequences of prior beliefs, in accordance with the rule, rather than revise those beliefs because their consequences are unacceptable. But the same goes for whatever principles regulate the selection of total views. We can always ask why we apply these particular criteria of coherence, or trade off competing criteria in the way we do, rather than modify our epistemic principles in order to hold on to a particular combination of beliefs. We had better not answer this question by appealing to a further set of meta-criteria: to do so will open a new regress.

Coherence theorists avoid these problems by implicitly assigning the criteria of coherence a special status. When they deny that any beliefs are epistemically privileged, they really mean any 'first-order' beliefs, beliefs about the world. The criteria of coherence—which embody 'second-order', epistemic beliefs about what makes the beliefs in a system more likely to be true—function as the fixed points by reference to which first-order acceptance is regulated. These epistemic beliefs thus enjoy a *foundational status*. In so far as this status is assigned *a priori*, the coherence theory represents a rationalistic—'top down' as opposed to 'bottom up'—variant of foundationalism.

This complicates the coherence theorist's attitude to 'linear inference'. The criteria of coherence prescribe revisions to total views: for example, they tell us that, if one modification to our belief-system increases its comprehensiveness, without loss of integration, while another reduces both, the system resulting from the first modification is more coherent and is to be preferred. This example of reasoning involves a straightforward application of Modus Ponens. The coherence theory is therefore not unqualifiedly hostile to linear inference: it excludes it only at first order. While there is nothing strictly contradictory in this view, it is oddly unmotivated. Once any element of epistemic privilege or linear inference is admitted, the question arises as to why its role should be so restricted. Why not allow it a larger role, given that doing so would let us stay closer to the phenomenology of everyday justification?

Knowing what we think

The epistemic beliefs implicit in acceptance of the criteria of coherence are not the only locus of epistemic privilege in the coherence theory. To apply these criteria we need to know what all our beliefs are and how they hang together.

This seems unrealistic. No one has the faintest idea how *many* beliefs he has, or even how to go about counting them. This isn't just because we have so many beliefs that we wouldn't know where to begin, though this is perfectly true. Rather, we lack clear criteria for individuating beliefs—that is, saying when beliefs are the same and different—without which there is no possibility of counting. Asking how many beliefs I have is like asking how many drops of water there are in a bucket: who's to say? I believe that my dog is in the garden right now; do I *also* believe that he is not in the house, not in the basement, not in Siberia? Or are these beliefs somehow included in the original belief? This is an odd question. With respect to beliefs, we do not normally have any use for fine-grained criteria of individuation. Unless we become captured by the coherentist image of 'our beliefs' as a finely articulated, complexly interrelated network—a 'web of belief'—I doubt we shall ever have any use for them.

Even if we had a clearer conception than we do of how our beliefs constitute a genuine totality, this totality would be completely unsurveyable. No one has a clue what his actual belief-system looks like *as a whole*, still less how it stacks up against the incalculable number of alternative systems that would result from the endlessly many ways in which his existing system might be modified. The coherence theory is impossibly idealized. It makes human reasoners like Leibniz's God, who conceives (in the minutest detail) every world there could *possibly* be, choosing to create the best.

But waiving this difficulty, what is the epistemological status of our supposed self-knowledge? As the author of the most comprehensive recent version of the coherence theory recognizes, the coherence theory must presuppose two things: that everyone has a primitive sense of what his beliefs are and how they hang together; and that he is entitled to presume that this sense is more or less accurate. Coherentist justification proceeds under a *Doxastic Presumption*.[5]

In an everyday sense, we all do know what we believe: that is, if asked for our opinion on a definite subject, we can give it if we have one to give. But this is a world away from what the coherence theorist has is mind. For his purposes, 'knowing what I believe' means having a sense of the extent and structure of my total belief-system. I have no such sense; and even if I did, I have no idea how I should estimate its accuracy. The Doxastic Presumption is *completely* unwarranted. Why not *Doxastic Assumption*?

Of course, the word 'presumption' is carefully chosen. If our supposed knowledge of our own belief-system is allowed to be questionable, it will need to be backed up by some justifying inference. This can hardly be coherentist: such an inference would require us to estimate the coherence of our beliefs about our beliefs. We would then have to presume these beliefs to be more or less accurate, or else produce a meta-meta-argument about the coherence of our beliefs about our beliefs about our beliefs . . . and so on without end. Invoking a 'presumption' is meant to prevent any such regress getting started. In effect, it confers *foundational status* on the relevant kind of self-knowledge.

Epistemologically basic beliefs concerning the content and structure of our total view are regulated by equally basic epistemic principles. The coherence theory is foundationalism in disguise.

Notes

1. Harman (1986), chs. 1–2, App. A.
2. For discussion of the D-N model by one of its leading advocates, see Hempel (1965).
3. Putnam (1978), 42.
4. The case for the interest-relativity of explanation is developed by Garfinkel (1981).
5. BonJour (1985), 81–2.

12

Realism and Truth

Losing the world

Our objections to reductionism, although decisive, may not get at what is really unsettling about reductionist strategies. The real problem with (say) a phenomenalist analysis of external-object talk is that, even if it could be carried through, it would not salvage our pre-theoretical aspirations towards knowledge of the external world. At best it would provide us with a pale simulacrum of such knowledge. For phenomenalists, the world is not something to which experience gives us access. Rather, talk about the world is just a roundabout form of talk about experience. Phenomenalism cannot satisfy our aspiration to knowledge of an objective world.

There is a logical aspect to this worry. According to the phenomenalist, to say that there are two bookcases in my office, even though the room is currently unoccupied, is to say that *if* you or I were to be in my office we *would* be having such and such experiences. However, simple statements about the external world do not feel the least bit 'iffy'. The statement that there are two bookcases in my office is straightforwardly categorical: true if there are two bookcases there, false otherwise.[1]

Phenomenalism is a form of idealism, the doctrine that thought and reality are really one and the same. The problem with idealism is that it flies in the face of common-sense *realism*. From a realist standpoint, the facts of the world do not depend on what anyone does or would experience. Human beings have existed for only a brief moment in the universe's history. Much has happened that we will never know about. Much would have happened even if human beings had never evolved. It is not easy for an idealist to respond to this thought in a satisfactory way. This is what I was hinting at when I suggested that empiricist foundationalism does not really offer enough in the way of external constraint on our beliefs.

Phenomenalists, or idealists generally, will reply that realism is mistaken or even incoherent. They will claim that realist prejudices are the source of sceptical anxieties. There is something to this. Sceptical arguments, particularly Cartesian arguments, do seem reflect a commitment to realism. The sceptic presents experience as a guide to an independently existing reality:

the only guide we have and, unfortunately, one that is completely inadequate. Phenomenalism calls in question the sceptic's conception of reality. Phenomenalism is anti-sceptical in virtue of being anti-realist. This is its strength, but also its weakness. The sceptic claims that we know nothing about the world: our knowledge extends no farther than experience. The phenomenalist tells us not to worry. Of course we know all sorts of things about the external world, because the external world is only a logical construction out of experience. This solution to scepticism is not easy to distinguish from scepticism itself.[2]

One lesson to learn from all this is the importance of what I call the apparent 'naturalness' of sceptical problems. To concede that sceptical arguments are natural or intuitive, is to give the sceptic a tremendous dialectical advantage. If realism leads to scepticism, and if realism is the common-sense or 'default' position, any anti-sceptical strategy that challenges realism will be *revisionary*: it will not really salvage what we thought of, pre-theoretically, as knowledge of the world. On the contrary, it will implicitly concede that our pre-theoretical epistemic aspirations are incurably paradoxical, which is why they must be changed. Any such change will inevitably involve large concessions to scepticism.

Truth and correspondence

Realism is generally taken to have important implications for our understanding of truth. According to realists, whether a proposition about the world is true depends on the way the world is, whether the proposition 'corresponds to the facts'. Realists therefore insist on distinguishing sharply between a proposition's being true and our being justified in taking it to be true. Sceptics, too, insist on this distinction, though they treat it as the entering wedge for a much more radical assault on the possibility of knowledge.

The correspondence theory of truth has attracted severe criticism. One focus of attack is the notion of correspondence itself. Talk of correspondence naturally calls to mind pictures or mirror-images. We do indeed speak of pictures as being 'true to life'. But the correspondence theory is supposed to explain the truth of beliefs or judgements or propositions, the content of which is expressed by complete sentences. Sentences are not pictures, at least in any straightforward way.

We might think that explaining correspondence is not quite so difficult for early modern philosophers, like Descartes, who discuss knowledge in terms of ideas. Ideas include sense-experiences as well as conceptually articulated thoughts; and we can, perhaps, form a vague conception of what might be meant by supposing our ideas to correspond to the world. We think of our experience as something like a private picture-show, and we ask ourselves whether the events on the screen (our experiences) accurately reflect what is going on outside the cinema (in the external world).

On reflection, this doesn't help. Knowledge requires judgements—thoughts or beliefs that can be true or false. Such thoughts demand propositional content: the sort of content that is expressed by complete sentences. Whatever the merits of thinking of experience as picturing the world, a picture doesn't say anything unless we read something into it. Even for Descartes, knowledge arises when we make judgements about the extent to which our experiences mirror the world. Whether or not we think of ideas as mental pictures, and whether or not we take the linguistic turn, the problem for the correspondence theory is to explain what makes judgements true.

Another problem for the correspondence theory is to explain what thoughts (or propositions or sentences) correspond *to*. The usual answer is *facts*. The trouble with the appeal to facts is that, in general, we have no way of indicating what fact a sentence, when true, corresponds to other than asserting the sentence. Thus the sentence 'Snow is white' is true iff it corresponds to the fact that snow is white. Saying that a thought or sentence corresponds to the facts looks more like another way of saying that it is true than a genuine explanation of what its being true amounts to.

There are sophisticated contemporary versions of the correspondence theory that avoid both mirror-imagery and naive appeals to facts. But examining them would take us too far afield. Moreover, I do not think that examining them would yield epistemologically significant results, at least so far as scepticism is concerned. The reason is that the 'realist' presuppositions of sceptical arguments are very minimal. Sceptical arguments do not depend on metaphysically charged conceptions of truth.

Deflating truth

For traditional theorists of truth, the nature of truth may be difficult to explain, but that truth has a nature is something they take for granted. Just as a scientist might want to explain what it is that makes some substances acidic, so a philosopher will want to explain what makes true sentences (propositions, beliefs, etc.) true. However this preconception has been challenged by proponents of so-called 'deflationary' accounts of truth.[3]

An early version of such an approach to truth is the redundancy theory, according to which 'It is true that Caesar was murdered' *means no more than* that Caesar was murdered. Any difference is entirely 'stylistic': for example, we may use 'It is true that . . .' to speak more emphatically. Because it stresses the use of 'true' in performing such special speech-acts, this approach is sometimes also called the 'performative' theory.

More recent views, such as Quine's disquotational theory, do not claim that 'p' and 'It is true that p' are synonymous. Rather, what matters about 'true' is given by certain logical equivalences. Thus:

> 'Snow is white' is true if and only if all snow is white; 'Grass is green' is true if and only if all grass is green . . . and so on.

Appending 'is true' to a quoted sentence is just like cancelling the quotation marks ('disquotation'). Notice that while, on the redundancy theory, to think of truth as *any* sort of property is, in Ramsey's words, just 'linguistic muddle', on the disquotational theory we *can* see truth as a property, a complete theory of which is given by the appropriate equivalences. Of course, we cannot write this 'theory' down, since it will have infinitely many axioms. But since these axioms share a common structure, we can indicate more or less what they are.[4] These axioms capture all there is to the idea of 'correspondence'.

For deflationists, the function of truth-talk is expressive. As Quine puts it, 'true' offers a way of replacing talk about the world with logically equivalent talk about words. Moving to the level of talk about words ('semantic ascent') gives us new things to generalize over: that is, linguistic objects, sentences. This move to the 'meta' level allows us to express agreement and disagreement with sentences that we cannot specify. One type of case is where we do not know exactly what these sentences are, as in 'What the President said is true'. Another is where there are too many of them. We want to assert all instances of logical laws, like the Law of the Excluded Middle. We can accomplish this with the aid of semantic ascent: every sentence of the form 'p or not p' is true.

This may seem trivial; and deflationary theorists are often suspected of belittling the concept of truth. In fact, however, the generalizing power of truth talk adds enormously to our language's expressive powers. It allows us to express general cognitive goals, such as seeking truths and avoiding falsehoods. It allows us to acknowledge our fallibility by saying that even our best-supported theories may be false: we do not have to list them. The capacity for generalization of this kind is indispensable to reflection on inferential and methodological commitments. And the ability to engage in such reflection is integral to human rationality.[5]

Truth and verification

The deflationary approach to truth has a good claim to capture all there is to the idea of truth as 'correspondence' or 'fitting the facts'. It is 'realist' only in what it does *not* say. It does not make any conceptual connection between truth and epistemological notions like justification or evidence. In this sense, the deflationary perspective retains the realist idea that truth is a *radically non-epistemic notion.*

Contrast this realist conception of truth with that implicit in reductionist theories like phenomenalism. Such theories, together with 'criterial' variants, are an articulation of a more general view of meaning: *verificationism.* According to verificationism, the meaning of a statement is constituted by its

verification conditions: the facts or procedures that go towards confirming or disconfirming it. A proposition's being true is therefore logically connected with the fulfillment of those conditions. Verificationism goes with an *epistemic* conception of truth, in the sense that truth is defined (at least for non-basic propositions) in terms of epistemological concepts like verification or justification.

We have seen that verificationism is a non-starter. Individual statements do not have verification conditions they can call their own. They connect with experience only as part of some set of sentences with critical semantic mass. However, the idea that truth is some kind of epistemic notion is not tied to strict verificationism.

Let us step back and ask why we should understand truth in a realist rather than in some epistemic way. The answer is: because we want to distinguish between how things really are and how we (however justifiably) take them to be. After all, we are not omniscient. Even resting on evidence that is to all intents and purposes conclusive is not the same as being true, for we can never entirely exclude the possibility of our coming to see that evidence in a new light. However, there is a reply to this. The fact that we are fallible only shows that we should not identify truth with justification by current standards. Rather, truth consists in *ideal* verification, verification in the limit, verification at the end of inquiry, or something like that.[6]

This reply detaches the idea of understanding truth in an epistemic way from reductionist forms of verificationism, such as phenomenalism. A proponent of the 'ideal justification' conception of truth need not imagine that individual propositions have their own precisely demarcated verification conditions. But since he postulates a conceptual link between truth and verifiability, he does not treat truth as a radically non-epistemic notion.

Where verificationism is associated with foundationalism, the ideal justification theory appeals to coherence theorists. Foundationalists accept certain basic facts as directly known or 'given'. These given facts constitute the verification conditions for all non-basic statements or beliefs. To the coherence theorist, the doctrine of the given looks like an unpurged element of non-epistemic truth. The coherence theorist wants to unify his account of justification and truth for all types of belief: truth is final coherence—ideal justification—for all beliefs.[7] A coherence theory of justification combined with an ideal justification theory of truth treats justification and truth in a consistently epistemic way.

This line of thought has a certain charm. But how far do we understand the notion of 'ideal' justification? Coherence theorists explain it in terms of incorporation into an ideally coherent system of belief: one that is maximally comprehensive and integrated. I doubt that any clear ideas attach to this kind of talk. The same goes for references to the 'limit' or 'end' of inquiry. We have no clear conception of what it would be for inquiry to have a limit or end, thus

no clear idea of what it would be for a system of belief to be maximally coherent. So far as I can see, the most sense we can make of 'ideal' justification is to understand it as involving procedures that take into account all relevant sources of error. This is as good as to say: the sort of justification we have when we are ideally placed for finding out the truth. If this is right, there is indeed a connection between truth and ideal justification but, from the standpoint of the epistemic theorist of truth, it goes the wrong way round. To the extent that we understand talk of ideal justification, it is because we have a prior understanding of truth.

Even setting this objection aside, it is doubtful whether the epistemic conception of truth really accomplishes its anti-sceptical goal. After all, we have no idea how close we are to making our belief-system ideally coherent. For all we know, our outlook at the end of inquiry might be quite different from our outlook today, so it is not clear that justification and truth have been connected by anything more than verbal sleight-of-hand: both may consist in 'coherence', but we are given no reason to suppose that coherence in the one case has much resemblance to coherence in the other.

Traditionally, coherence theorists have tended to think of inquiry in teleological terms, as heading towards a goal. This conception is implicit in the idea of the 'end' or 'limit' of inquiry. If we can make an assumption like this, perhaps we can do something to answer the sceptic. But of course, we can always answer the sceptic by making assumptions, including assumptions much simpler and more in tune with common sense than this one.

These are all serious concerns. But, as with our original objections to phenomenalism, they may not get at the deepest source of disquiet. The coherence theory of truth, in its own way, involves a retreat from common-sense realism. The association between the coherence theory of truth and Idealism—the metaphysical thesis that Thought and Reality are somehow one—is no accident. It arises from a problem we noted at the end of the previous chapter: that the coherence theory has difficulty allowing for any notion of genuinely external constraint. It insists that everything that is relevant to the epistemic status of a particular belief be swept up into our total belief-system.

The move from a coherence theory of justification to a coherence theory of truth depends on raising the spectre of radical and general scepticism. But this spectre, once raised, can be raised again. We can ask what makes even an ideally coherent system of belief *true of objective reality*. The traditional Idealist response is that 'reality' and 'the ideally coherent system of thought' are two ways of referring to the same thing. But what else is there is to say, once we have travelled this far down the coherentist road?

The more all-encompassing Idealism that grows out of the coherence theory prompts the same anxieties as the subjective idealism involved in phenomenalism, which reduces the world to a construction out of

sense-experiences. Both approaches try to quiet sceptical doubts by retreating from common-sense realism. But common-sense realism is not easily cast aside. It is hard to shake the thought that most of what has happened in the universe happened before sentient beings came along and would have happened even if no such beings had ever existed. How can, then, the facts of the world be constituted by what anyone does or would have reason to think?

The advantage of adopting a deflationary view of the truth-predicate is that it allows us to sidestep problems like this. Deflationism captures everything worth capturing in the realist conception of truth. However, a deflationary view of truth is not an epistemological cure-all.

Scepticism and generality

Great philosophical battles have been fought over the Nature of Truth. The combatants have generally thought that the fate of human knowledge was at stake: that if we misunderstand truth, we will have no defence against radical scepticism. But I think that questions about the nature of truth have much less to do with epistemology than is commonly supposed.

As far as the Agrippan problem is concerned, it seems clear that no elaborate ideas about truth are presupposed. All the sceptic needs on this score is the minimally realist concession, granted even by deflationism, that our believing something to be true and its actually being true are not the same thing. Given this much, he can ask us to back up our belief, and we are off and running. But the concession in question involves no more than the generalizing use of truth-talk that deflationism highlights.

Cartesian scepticism may seem to be another matter, for this form of scepticism does seem to concern whether our experiences or beliefs correctly represent external reality. However, here too the heart of the problem lies with the unusual generality of the sceptic's questions. What the sceptic wants to know is how anything at all that we believe about the world amounts to knowledge: what justification we have to think that any of our beliefs about the world are true. Here, again, he is employing truth-talk in its generalizing capacity, to call in question the indefinitely many things we believe about the world around us, but cannot simply list.

To come to terms with scepticism, we must investigate the apparently innocent generality of the sceptic's questions. We shall make a start on this in the next chapter, though the argument will not be completed until Chapter 16.

Notes

1. Isaiah Berlin, 'Empirical Propositions and Hypothetical Statements', in Berlin (1979).
2. For a subtle elaboration of this theme, see Stroud (1984*a*), esp. chs. 4, 5.
3. For deflationary approaches to truth, see 'Facts and Propositions', in Ramsey (1990); Quine (1990), ch. 5; and esp. Horwich (1998*b*). Blackburn and Simmons (1999) is an excellent anthology of papers about truth, many having to do with the pros and cons of deflationism.
4. Only more or less because, without some restrictions on admissible substitutions, a schema like '"p" is true if and only if p' will generate semantic paradoxes, such as the famous paradox of the liar. Much philosophical work on truth is devoted to explaining how the paradoxes can be avoided. For more on this (somewhat technical) topic, see Soames (1999).
5. This is emphasized by Brandom (1994), ch. 5.
6. This idea, which goes back to C. S. Pierce, has recently been defended by Putnam. See e.g. Putnam (1981), ch. 3.
7. Blanshard (1939), ch. 25. For more detailed criticism of the coherence theory of truth, and its relevance to scepticism, see Williams (1996*a*), chs. 6, 7.