Preface to the Second Edition

The differences between the present work and the first edition of *Truth* have three sources. First, some new objections to the minimalistic theory have appeared in the last eight years, and I attempt in what follows to reply to them, particularly to the doubts raised by Anil Gupta, Hartry Field, Mark Richard, Donald Davidson, Bernard Williams, Scott Soames, Michael Devitt, Crispin Wright, and Paul Boghossian. Second, several easily correctable mistakes have come to my attention, along with quite a few points at which my original formulations could be simplified and strengthened. And third, I now have a worked-out version of the view of meaning that minimalism requires: namely, a *use* theory. This point of view and its implications have been incorporated here, although somewhat sketchily: a full account of it can be found in another book, *Meaning*, published in conjunction with this one.

The upshot of these forces for change is that nearly every section of the book has been revised to one degree or another, especially the parts concerning propositional structure (Question 2), the 'liar' paradoxes (Question 10), the grounds for concluding that truth has no underlying nature (Question 14), the intrinsic value of truth (Question 19), negation (Question 25), vagueness (Question 28), utterance truth (Question 34) and the concept of 'correspondence with facts' (Questions 35 and 36). In addition, there is now a Postscript, which restates the essence of deflationism about truth and responds to the above-mentioned critics. Besides these philosophers I would also like to thank Philippe de Rouihlon, Jesus Mosterin, Marcello Pera, Gabriele Usberti, and Albert Visser for their suggestions about how the first edition could be improved.

Preface to the First Edition

Perhaps the only points about truth on which most people could agree are, first, that each proposition specifies its own condition for being true (e.g. the proposition *that snow is white* is true if and only if *snow is white*), and, second, that the underlying nature of truth is a mystery. The general thrust of this book is to turn one of these sentiments against the other. I want to show that truth is entirely captured by the initial triviality, so that in fact nothing could be more mundane and less puzzling than the concept of truth.

This rough idea has been expressed by a fair number of eminent twentieth-century philosophers—including Frege, Wittgenstein, Ramsey, Ayer, Strawson, and Quine—and I certainly claim no originality for advocating it. But in spite of these impressive endorsements, the so-called 'redundancy theory of truth' remains unpopular; and this, I believe, is because a full case for it has never been made. The purpose of the present essay is to fill that gap. I have tried to find the best version of the idea—which I call 'minimalism'—give it a clear formulation, deal with a broad range of familiar objections, and indicate some of its philosophical consequences. I hope that this treatment will help the deflationary view of truth finally gain the acceptance that it deserves. Not only would this be good in itself, but the effect on many surrounding issues would be quite beneficial. For despite its reputation for obscurity the notion of truth is constantly employed in philosophical theory. One is tempted to rely on it in describing, for example, the aims of science, the relations of language to the world, the character of sound reasoning, and the conditions for successful planning. Yet only in light of an adequate account of truth, and an appreciation of what the notion may and may not be used for, can such ideas be fully understood and evaluated.

My plan is as follows. I begin (Chapter 1) by presenting the minimalist conception, and in the following chapters I refine and defend

it in the course of responding to 39 alleged difficulties (which are set out in the Contents). In Chapter 2, I start to deal with these objections, answering questions about what is required of an adequate theory of truth and distinguishing minimalism from other deflationary accounts. In Chapter 3, I argue that the role of truth in laws of cognitive science is nothing more than a display of its minimalistic function and that it should not lead us to expect there to be a theoretical analysis of truth. I then turn to the use of the concept in philosophical theory and suggest that this is often a cause of confusion: generally the issues are independent of truth, and its introduction merely muddies the waters. This is illustrated in treatments of scientific realism (Chapter 4) and of problems in meta-semantics and in the philosophy of logic (Chapter 5). My assumption throughout is that *propositions* are the bearers of truth and, for those readers not comfortable with this idea, Chapter 6 offers some arguments in its favour. Finally, in Chapter 7, I address the feeling that truths are what *correspond to reality*, and I determine the extent to which this intuition may be squared with the minimalist perspective.

In order to explain this conception of *truth* I have had to say something about various other matters such as *reference*, *meaning*, *belief*, *logic*, *vagueness*, *realism*, and the notion of *proposition*, and I have sometimes taken positions in these areas without providing adequate support for them. I hope that the reader will sympathize with the desire to keep this book focused and short, and will agree that the sketchiness of some of these discussions is justified by that end.

The point of view articulated here is a development of some ideas in my 'Three Forms of Realism' (published in *Synthese* in 1982), which was in turn a reaction to various writings of Michael Dummett and Hilary Putnam. Although I disagree with their conclusions about truth, I have benefited from the depth and ingenuity of their thought. It was only against this rich background that my own contrasting position could be elaborated. Another debt is to Hartry Field, with whom I have had several conversations about truth in the last few years. I'm afraid I don't remember if either of us ever convinced the other of anything, but I do remember coming away from our meetings knowing that I had been helped a great deal. Anyone interested in the issues addressed here should read his essay, 'The Deflationary Conception of Truth'. In addition I would like to thank Ned Block, who saw many of my drafts and, as always, supplied lots of reasonable advice; Marcus Giaquinto, who never quite accepts anything I

Contents

The following is a list of the questions and problems regarding minimalism to which replies and solutions will be proposed in the course of this essay

5 Meaning and Logic 68

1

The Minimal Theory

A Sketch of the Minimalist Conception

'What is truth?', we sometimes ask—but the question tends to be rhetorical, conveying the somewhat defeatist idea that a good answer, if indeed there is such a thing, will be so subtle, so profound, and so hard to find, that to look for one would surely be a waste of time. The daunting aura of depth and difficulty which surrounds this concept is perfectly understandable. For on the one hand the notion of truth pervades philosophical theorizing about the basic nature and norms of thought and action—e.g. '*truth* is the aim of science'; '*true* beliefs facilitate successful behaviour'; '*truth* is preserved in valid reasoning'; 'to understand a sentence is to know which circumstances would make it *true*'; 'evaluative assertions can be neither *true* nor false'. So insight into the underlying essence of truth promises, by helping us assess and explain such principles, to shed light on just about the whole of our conceptual scheme. But, on the other hand, this very depth can suggest that in inquiring into the nature of truth we have run up against the limits of analysis; and, indeed, it will be widely agreed that hardly any progress has been made towards achieving the insight we seem to need. The common-sense notion that truth is a kind of 'correspondence with the facts' has never been worked out to anyone's satisfaction. Even its advocates would concede that it remains little more than a vague, guiding intuition. But the traditional alternatives—equations of truth with 'membership in a coherent system of beliefs', or 'what would be verified in ideal conditions', or 'suitability as a basis for action'—have always looked unlikely to work, precisely because they don't accommodate the 'correspondence' intuition, and this air of implausibility is substantiated in straightforward counterexamples. Hence the peculiarly enigmatic

character of truth: a conception of its underlying nature appears to be at once necessary and impossible.

I believe that this impression is wholly wrong and that it grows out of two related misconceptions: first, that truth *has* some hidden structure awaiting our discovery; and secondly, that hinging on this discovery is our ability to explain central philosophical principles such as those just mentioned, and thereby to solve a host of problems in logic, semantics, and epistemology.

The main cause of these misconceptions, I suspect, is linguistic analogy. Just as the predicate 'is magnetic' designates a feature of the world, *magnetism*, whose nature is revealed by quantum physics, and 'is diabetic' describes a group of phenomena, *diabetes*, characterizable in biology, so it seems that 'is true' attributes a complex property, *truth*—an ingredient of reality whose underlying structure will, it is hoped, one day be revealed by philosophical or scientific analysis. The trouble is that this conclusion—which we tend to presuppose in the question, 'What is truth?'—is unjustified and false. An expression might have a meaning that is somewhat disguised by its superficial form—tending, as Wittgenstein warned, to produce mistaken analogies, philosophical confusion, and insoluble pseudo-problems. The word 'exists' provides a notorious example. And we are facing the same sort of thing here. Unlike most other predicates, 'is true' is not used to attribute to certain entities (i.e. statements, beliefs, etc.) an ordinary sort of property—a characteristic whose underlying nature will account for its relations to other ingredients of reality. Therefore, unlike most other predicates, 'is true' should not be expected to participate in some deep theory of that to which it refers—a theory that articulates general conditions for its application. Thus its assimilation to superficially similar expressions is misleading. The role of truth is not what it seems.

In fact, the truth predicate exists solely for the sake of a certain logical need. On occasion we wish to adopt some attitude towards a proposition—for example, believing it, assuming it for the sake of argument, or desiring that it be the case—but find ourselves thwarted by ignorance of what exactly the proposition is. We might know it only as 'what Oscar thinks' or 'Einstein's principle'; perhaps it was expressed, but not clearly or loudly enough, or in a language we don't understand; or—and this is especially common in logical and philosophical contexts—we may wish to cover infinitely many propositions (in the course of generalizing) and simply can't have all of them

in mind. In such situations the concept of truth is invaluable. For it enables the construction of another proposition, intimately related to the one we can't identify, which is perfectly appropriate as the alternative object of our attitude.

Consider, for example,

(1)   What Oscar said is true.

Here we have something of the form

(2) x is F,

whose meaning is such that, given further information about the identity of x—given a further premise of the form

(3) x = the proposition that p

—we are entitled to infer

(4) p.

And it is from precisely this inferential property that propositions involving truth derive their utility. For it makes them, in certain circumstances, the only appropriate objects of our beliefs, suppositions, desires, etc. Suppose, for example, I have great confidence in Oscar's judgement about food; he has just asserted that eels are good but I didn't quite catch his remark. Which belief might I reasonably acquire? Well, obviously not that eels are good. Rather what is needed is a proposition from which that one would follow, given identification of what Oscar said—a proposition equivalent to

(1*)   If what Oscar said is *that eels are good* then eels are good, and if he said *that milk is white* then milk is white, ... and so on;

and the *raison d'être* of the concept of truth is that it supplies us with such a proposition: namely (1).

To take another example, suppose we wish to state the logical law of excluded middle:

(5)   Everything is red or not red, and happy or not happy, and cheap or not cheap, ... and so on.

Our problem is to find a single, finite proposition that has the intuitive logical power of the infinite conjunction of all these instances; and the concept of truth provides a solution.

(6)   Everything is red or not red

is known to be equivalent to

(6*)   The proposition *that everything is red or not red* is true.

And similarly for the other instances. Thus the infinite series of universal disjunctions may be transformed into another infinite series of claims in which the same property, *truth*, is attributed to all the members of a class of structurally similar propositional objects. And in virtue of that form the sum of *these* claims may be captured in an ordinary universally quantified statement:

(5*)   Every proposition of the form: ⟨everything is F or not F⟩ is true.

It is in just this role, and not as the name of some baffling ingredient of nature, that the concept of truth figures so pervasively in philosophical reflection.[1]

What permits the notion of truth to play that role is simply that, for any declarative sentence

---

[1] Notice that one could design an alternative way of putting the things that we actually express by means of the truth predicate. With the introduction of *sentence* variables, *predicate* variables, and *substitutional* quantification, our thoughts could be expressed awkwardly as follows:

(1**)   For any sentence such that Oscar claimed that *it*, then *it*,

or in logical notation

(1***) (p)(Oscar claimed *that p* → *p*);

and

(5**)   Given any predicate, a thing is either *it* or not *it*,

or

(5***) $(F)(x)(Fx \lor -Fx)$.

However, the variables '*it*', '*p*', and '*F*' are not the usual kind which replace *noun* phrases and refer to objects. Rather, '*F*' must be construed as a 'pro-predicate', and '*p*' as a 'pro-sentence'. Moreover, *generalization* with respect to these variables cannot be understood in the usual way as saying that every object has a certain property, but must be construed as asserting the truth of every legitimate substitution instance. Thus (1**) means intuitively that any result of substituting an English declarative sentence for '*p*' in 'Oscar claimed *that p* → *p*' is true.

The advantage of the truth predicate is that it allows us to say what we want without having to employ any new linguistic apparatus of this sort. It enables us to achieve the effect of generalizing substitutionally over sentences and predicates, but by means of ordinary variables (i.e. pronouns), which range over *objects*. See Chapter 2, the answer to Question 6, for discussion of this point.

(4)   *p*

we are provided with an equivalent sentence

(4*)   The proposition *that p* is true,

where the original sentence has been converted into a noun phrase, 'The proposition *that p*', occupying a position open to object variables, and where the truth predicate serves merely to restore the structure of a sentence: it acts simply as a *de-nominalizer*. In other words, in order for the truth predicate to fulfil its function we must acknowledge that

(MT)   The proposition *that quarks really exist* is true if and only if quarks really exist, the proposition *that lying is bad* is true if and only if lying is bad, . . . and so on,

but nothing more about truth need be assumed. The entire conceptual and theoretical role of truth may be explained on this basis. This confirms our suspicion that the traditional attempt to discern the *essence* of truth—to analyse that special quality which all truths supposedly have in common—is just a pseudo-problem based on syntactic overgeneralization. Unlike most other properties, *being true* is unsusceptible to conceptual or scientific analysis. No wonder that its 'underlying nature' has so stubbornly resisted philosophical elaboration; for there is simply no such thing.

This sort of deflationary picture is attractively demystifying.[2] Nevertheless, it has not been widely accepted, for it faces a formidable array of theoretical and intuitive objections. My aim in this book is to work out a form of the approach that is able to deal with all the alleged difficulties. Some of them expose genuine deficiencies in certain versions of the doctrine and reveal the need for a better formulation of the deflationary position. But most of the complaints have simply been given more weight than they deserve. Indeed, I tend to think that the approach has been underrated more because of the sheer number of objections to it than because of their quality. Put in

---

[2] More or less deflationary views about truth are endorsed and defended (in various forms and to various degrees) by Frege (1891, 1918), Ramsey (1927), Ayer (1935), Wittgenstein (1922, 1953), Strawson (1950), and Quine (1970). In recent years the idea has been developed by Grover, Camp, and Belnap (1975), Leeds (1978), the present author (1982*a*), A. Fine (1984), Soames (1984, 1997), Field (1986, 1994), M. J. Williams (1986), Loar (1987), Baldwin (1989), and Brandom (1988, 1994).

more positive terms, my plan is to provide a highly deflationary account of our concept of truth—but one that can nevertheless explain the role of the notion in scientific methodology and in science itself, and enable us to find answers to such questions as: In what does our grasp of truth consist? Why is it practically useful to believe the truth? Can there be, in addition, any purely intrinsic value to such beliefs? Does science aim and progress towards the truth? How does our conception of truth bear on the nature of various types of fact and on our capacity to discover them? Is truth an explanatorily vital concept in semantics or in any of the empirical sciences? I shall start by giving what I believe is the best statement of the deflationary point of view. Because it contains no more than what is expressed by uncontroversial instances of the equivalence schema,

   (E)   It is true *that p* if and only if *p*,

I shall call my theory of truth '*the minimal theory*', and I shall refer to the surrounding remarks on behalf of its adequacy as '*the minimalist conception*'. With a good formulation in hand, I want to show that the standard criticisms of deflationary approaches are either irrelevant or surmountable, to display the virtues of the theory in comparison with alternatives, and, by answering the above questions, to draw out the implications of minimalism for issues in semantics, psychology, and the philosophy of science. For the sake of simplicity and conformity with natural language I begin by developing the account of truth for *propositions*. However, I shall go on to argue that the minimalist conception applies equally well to the 'truth' of utterances, mental attitudes, and other types of entity.

   It might be thought that minimalism is too obvious and too weak to have any significant philosophical implications. Let me try, in at least a preliminary manner, to quell this misgiving. The real proof, of course, will be in the execution of the project. We should start by distinguishing (very roughly) two types of 'philosophical implication' that might be expected. First, there are general principles involving truth: for example, the fact that verification indicates truth, and that true beliefs are conducive to successful action. And, secondly, there are solutions to philosophical problems such as the paradoxes of vagueness and the issue of scientific realism. According to the minimalist conception, the equivalence schema, despite its obviousness and weakness, is not too weak to have significant philosophical implications—at least within the first category. On the contrary, our

thesis is that it is possible to explain all the facts involving truth on the basis of the minimal theory. This may indeed appear to be a rather tall order. But remember that most of the interesting facts to be explained concern relations between truth and certain other matters; and in such cases it is perfectly proper to make use of theories about these other matters, and not to expect all the explanatory work to be done by the theory of truth in isolation. When this methodological point is borne in mind it becomes more plausible to suppose that the explanatory duties of a theory of truth can be carried out by the minimal theory.

As for the second class of 'philosophical implication'—namely, solutions to problems—one would expect these to flow, not from the minimal theory as such (i.e. instances of the equivalence schema), but rather from the minimalist *conception* (i.e. the thesis that our theory of truth should contain nothing more than instances of the equivalence schema). Philosophical questions are typically based on confusion rather than simple ignorance. Therefore an account that makes plain the character of truth will permit a clearer view of any problems that are thought to involve truth. The account itself may well never entail, or even suggest, any solutions. But, in so far as it elucidates one of the sources of confusion, it will help us to untangle the conceptual knots that are generating the problems, and thereby facilitate their solution. In the limiting case, a conception of truth can achieve this result by enabling us to see that, contrary to what has been generally presupposed, the notion of truth is not even involved in the problem. The recognition that truth plays no role can be vital to achieving the clarity needed for a solution. Thus, to put the matter somewhat paradoxically, the *relevance* of a theory of truth may lie in its import regarding the *irrelevance* of truth. We shall see, I think, that this is very often the situation. Consider, for example, the debate surrounding scientific realism. It is commonly assumed that truth is an essential constituent of the problem; one sees reference to 'realist conceptions of truth' and to 'anti-realist conceptions of truth'; and questions about the meaning of theoretical assertions, our right to believe them, and what it would be for them to be true, are all lumped together as components of a single broad problem. This intertwining of philosophically puzzling notions is why the realism issue has proved so slippery and tough. What I am claiming on behalf of the minimalist conception of truth is not that it, by itself, will engender realism or anti-realism; but rather that it will make it easier for us to

see that the central aspects of the realism debate have nothing to do with truth. By providing this clarification of the main problems, minimalism will take us a long way toward being able to solve them.

The Space of Alternative Theories

It will help us to focus on what is at stake in accepting the minimalist conception of truth if I contrast it with some of the well-known alternatives.

Correspondence

First, there is the venerable notion that truth is the property of *corresponding with reality*. In its most sophisticated formulations this has been taken to mean that the truth of a statement depends on how its constituents are arranged with respect to one another and which entities they stand for. One strategy along these lines (Wittgenstein, 1922) is to suppose that a statement as a whole *depicts* a fact whose constituents are referents of the statement's constituents, and that the statement is true if and only if such a fact exists. Another strategy (Austin, 1950; Tarski, 1958; Davidson, 1969) is to define truth in terms of reference and predicate-satisfaction without importing the notions of fact and structure. Either way, these correspondence theories further divide according to what is said about reference. For example, one might suppose, with Wittgenstein (1922), that it is simply indescribable; or, with Field (1972) and Devitt (1984), that reference is a naturalistic (causal) relation; or, with Quine (1970) and Leeds (1978), that it is merely a device for semantic ascent. From our minimalist point of view, the last of these ideas is along the right lines—reference and truth being parallel notions—although, as we shall see, it is a mistake to explain truth in terms of reference.

Coherence

The second most popular view of truth is known as the coherence theory. A system of beliefs is said to be coherent when its elements are consistent with one another and when it displays a certain overall

simplicity. In that case, according to the coherence theory, the whole system and each of its elements are true. Thus truth is the property of *belonging to a harmonious system of beliefs*. This line was urged by the idealists Bradley (1914) and Blanshard (1939), embraced by Hempel (1935) as the only alternative to what he regarded as the obnoxious metaphysics of *correspondence*, and resurrected for similar reasons by Dummett (1978) and Putnam (1981) (as the 'verificationist' or 'constructivist' theory) in their identification of truth with idealized justification. What has struck most philosophers as wrong with this point of view is its refusal to acknowledge what would appear to be a central feature of our conception of truth, namely the possibility of there being some discrepancy between what really *is* true and what we will (or should, given all possible evidence) *believe* to be true. For it seems quite conceivable that there are facts we are not capable of discovering (for example, that there are infinitely many stars, or that every even number is the sum of two primes); conversely, our evidence might on occasion point incontrovertibly toward some conclusion that happens to be false; so truth and verification are not the same.

Pragmatism

In the third place we have the so-called pragmatic theory of truth devised by James (1909) and Dewey (1938), and recently elaborated by Rorty (1982) and Papineau (1987). Here truth is *utility*; true assumptions are those that work best—those which provoke actions with desirable results. However, just as in the case of the coherence theory, one must guard against taking a strong correlation for an identity. Although there is indeed an association between the truth of a belief and its tendency to facilitate successful activity, the tightness of the association should not be exaggerated. After all, actions based on true belief can none the less work out badly. Moreover, the link, to the extent that it holds, is surely something to be *explained*, not merely stipulated in the very definition of truth.

Unanalysable quality

Fourthly—perhaps the least attractive conclusion—there is the one-time thesis of Moore (1899, 1910/11) and Russell (1904) that truth is

an indefinable, inexplicable quality that some propositions simply have and others simply lack—a fundamental property of which no account can be provided.[3] This gives a sense of impenetrable mysteriousness to the notion of truth and can be the resort only of those who feel that the decent alternatives have been exhausted.

These traditional approaches do not typically impugn the correctness of the equivalence schema,

(E) $\langle p \rangle$ is true iff p,[4]

but question its completeness. They deny that it tells us about the essential nature of truth, and so they inflate it with additional content in ways that, I will argue, are, at best, unnecessary and, at worst, mistaken. To explain this point a little further it is useful to imagine various dimensions on which alternative accounts of truth may be characterized—each dimension varying with respect to some form of theoretical commitment.

1. An account may or may not be compositional—it may or may not define the truth of an utterance or a proposition in terms of the semantic properties of its parts. For example, a theory inflated in this way might involve the principle,

(T/R)   'a is F' is true iff there exists an object x such that 'a' refers to x and 'F' is satisfied by x.

The minimalist policy is not to *deny* such principles relating truth, reference, and satisfaction, but to argue that our theory of truth should not contain them as *axioms*. Instead, they should be *derived* from a conjunction of the theory of truth and quite distinct minimalist theories of reference and satisfaction.

2. One may or may not insist on a conceptual analysis of truth, a specification (in philosophically unproblematic terms) of the content of every statement employing the concept. Minimalism offers no such definition, and denies the need for one.

---

[3] For an examination of this view as it appears in the early writings of Moore and Russell see Cartwright (1987).

[4] I shall write '$\langle p \rangle$' for 'the proposition that p', and 'iff' for 'if and only if'.

3. An account may or may not suppose that there is some substantive, reductive theory of truth—some non-definitional analysis specifying the underlying property in which truth consists. In the context of a compositional account, the parallel issue is whether *reference* and *satisfaction* are complex relational properties—according to some philosophers, reducible to certain causal notions. The minimalist denies that truth, reference, or satisfaction are complex or naturalistic properties.

4. One may or may not attempt to formulate a non-trivial, finite theory of truth itself—a succinct body of statements about truth that can be tacked on to our other theories (in physics, mathematics, etc.) to enable the deduction of everything we believe about truth. According to minimalism, there is no such thing. We can say what is in the basic theory of truth—an infinity of biconditionals of the form '⟨p⟩ is true iff p'—but we can't formulate it explicitly because there are too many axioms.

5. One may or may not propose an account which inextricably links truth with other matters: for example, assertion, verification, reference, meaning, success, or logical entailment. Minimalism involves the contention that truth has a certain purity—that our understanding of it is fairly independent of other ideas.

Thus my account will take the less theoretically loaded view with respect to each of these dimensions of commitment. The theory of truth it proposes involves nothing more than the equivalence schema; it is non-compositional; it denies that truth and reference are complex or naturalistic properties; and it does not insist on an eliminative account of truth attributions. In this way minimalism aims for a maximally deflationary theory of truth, which, though complete, has no extraneous content—a theory about truth, the whole of truth, and nothing but truth.

   I should stress that the minimalist critique of the correspondence, coherence, constructivist, pragmatist, and primitivist accounts of truth is *not* that they are false. On the contrary, it seems quite likely that carefully qualified, true versions of each of them could be concocted. The main objection is rather that none can meet the *explanatory* demands on an adequate theory of truth. Specifically, none provides a good account of why it is that instances of the equivalence

schema are true. Minimalism involves a reversal of that explanatory direction. We shall find that on the basis of the equivalence biconditionals it is easy to see why, and in what form, the traditional principles hold. Indeed, every fact about truth can be naturally derived from those biconditionals. Therefore it is they that should constitute our basic theory of truth.

Summary of Alleged Difficulties

Objections to deflationary approaches have concerned six related topics.

The Proper Formulation. It has been no easy matter to provide even a prima facie plausible version of such a theory of truth—something that meets the normal methodological standards of fidelity to obvious fact, simplicity, explanatory power, etc., and that is not falsified by the 'liar' paradoxes.

The Explanatory Role of the Concept of Truth. The concept of truth is apparently employed in certain forms of scientific explanation (e.g. to help account for the contribution of language use to the achievement of practical goals), and it has been argued on this basis that deflationism must be missing something—namely, the naturalistic character that provides truth with its causal properties.

Methodology and Scientific Realism. A natural (realist) view of science is that it aims for, and gradually progresses towards, the truth—a goal that exists independently of our capacity to reach it, and that we value partly for its own sake, independently of any practical benefits that might accrue. This position would seem to require a substantial notion of truth—a conception of just the sort that the deflationary point of view eschews. In other words, any deflationary account of truth would seem to entail an anti-realist perspective on science.

Meaning and Logic. A further body of objections concerns the role of truth in semantics, and the ability of any deflationary approach to explain this role. For example, it is usual to analyse understanding in terms of knowledge of truth conditions, to use the concepts of truth

and reference to show how the meanings of sentences depend on the meanings of their parts, to suppose that truth must be a central concept in the appraisal of alternative rules of inference, and to treat various semantic phenomena (e.g. vagueness, empty names, expressive utterances) by exploiting the idea that a proposition might be neither true nor false. It is commonly assumed that deflationary theories of truth are precluded by these demands.

Propositions and Utterances. Propositions can appear to be such obscure and bizarre entities that it may seem undesirable to base an account of truth on the schematic principle

(E)   The proposition ⟨*p*⟩ is true iff *p*

which presupposes them. At the same time, the natural deflationary account of truth for *utterances*, the disquotational schema

(D)   Any utterance of the sentence '*p*' is true iff *p*,

has difficulty with indexical expressions (try 'I am hungry'), foreign languages ('Schnee ist weiß'), and indeed with all sentence-tokens whose truth or falsity depends on the context in which they are produced.

The 'Correspondence' Intuition. The idea that a representation is made true by its correspondence to reality has great intuitive appeal, yet there appears to be no room for any such conception within the deflationary picture.

Each of these topics receives separate treatment in the following chapters. However, they need not be taken in their order of appearance. Readers uncertain about what exactly the minimal theory is should certainly not miss Chapter 2. But otherwise one can proceed directly to Chapter 3 where some influential arguments against minimalism are rebutted and a case in favour of it is made, or Chapters 4 and 5 where its implications are examined. Anyone wary of *propositions* should not delay long before looking at Chapter 6 where I hope their concerns will be assuaged. And those philosophers who are fond of the correspondence theory of truth should perhaps not wait until the end before reading Chapter 7 and seeing that most of their intuitions may be accommodated.

   I have organized the above-mentioned areas of criticism into 39
specific questions and objections. In what follows I shall articulate
these problems in more detail and, in each case, sketch what I think is
an adequate response. What will emerge, I hope, is a view of truth
that is clear, plausible, and fairly comprehensive.

2

The Proper Formulation

The conception of truth to be defended in this essay is similar in spirit to other deflationary accounts that have appeared during the past hundred years or so, maintaining, in one way or another, that truth is not a normal property and that traditional investigations into its underlying nature have been misconceived. None of these accounts, however, has won over very many adherents, and the vast majority of philosophers either still subscribe to some form of correspondence, coherence, pragmatist, or primitivist picture, or else think that no decent theory has yet been made available. One cause of dissatisfaction with deflationary proposals in the literature is that they are not described fully or precisely enough to be properly evaluated. For instance, it isn't always said whether the theory concerns the nature of *truth itself*, or merely the meaning of the word 'true'. Secondly, and exacerbating the evaluation problem, there is a tendency to omit explicit statement of what a satisfactory account is supposed to do. The adequacy conditions for a theory of truth are left unclear. A third common defect of deflationary views is their commitment to certain blatantly implausible theses: for example, that *being true* is not a property at all, or that *every* instance of ' "*p*" is true iff *p*' is correct. And, in the fourth place, objections are often left standing that could in fact be rebutted: for example, that the theory fails to say *what truth is*, and that it cannot be reconciled with the *desirability* of truth. The purpose of this chapter is to reach an exact characterization of the minimalist conception and, whilst doing so, to show how to deal with some of the problems that have notoriously afflicted previous deflationary proposals.

1. Of what kinds are the entities to which truth may be attributed?

The list of candidates includes: (a) *utterances*—individual sounds and marks located in particular regions of space and time (e.g. Oscar's saying the words 'I am hungry' at midday on 1 January 1988); (b) *sentences*—types of expression in a language: syntactic forms that are exemplified by particular utterances (e.g. the English sentence 'I am hungry'); (c) *statements*, *beliefs*, *suppositions*, etc.— individual, localized actions or states of mind (e.g. Oscar's state at midday of believing that he is hungry); (d) *propositions*—the things that are believed, stated, supposed, etc.; the contents of such states (e.g. *that Oscar was hungry at midday on 1 January 1988*). I shall follow ordinary language in supposing that truth is a property of propositions.[1] Thus, if we agree with Oscar, we attribute truth to *what he said*, to the proposition he asserted. Evidently the *sentence-type of English* that he used is not true; for that very sentence-type is used on other occasions to make false statements. Nor would one normally characterize the noises he made, or his belief state, as true. These entities are more naturally described as 'expressing a truth' and 'being *of* a true proposition'. No doubt we do attribute truth to statements, beliefs, suppositions, and so on; but surely what we have in mind is that the propositional objects of these linguistic and mental acts are true, and not the acts themselves.

Most of the time I will conform to this way of speaking. To some extent this decision is non-trivial; for it involves a commitment to the existence of a breed of things called 'propositions'. However this commitment, though controversial and in need of some defence (to be supplied in Chapter 6), is much less substantial than it might seem at first. For it presupposes very little about the *nature* of propositions. Granted, minimalism entails that the notion of proposition not depend on the notion of truth. For the minimalist's direction of conceptual priority is the other way round: in so far as our concept

---

[1] In light of the locution 'It is true that *p*', it might be thought that a theory of the truth *predicate* would have to be supplemented with a separate theory of the truth *operator*; but this is not so. We can construe 'It is true *that p*', on a par with 'It is true *what Oscar said*', as an application of the truth predicate to the thing to which the initial 'It' refers, which is supplied by the subsequent noun phrase, '*that p*'.

of truth is constituted by our acceptance of instances of 'The proposition *that p* is true iff *p*', we must already be capable of grasping propositions. But this requirement leaves open many possibilities. As far as the minimal theory of truth is concerned, propositions could be composed of abstract Fregean senses, or of concrete objects and properties; they could be identical to a certain class of sentences in some specific language, or to the meanings of sentences, or to some new and irreducible type of entity that is correlated with the meanings of certain sentences. I am not saying that there is nothing to choose amongst these answers. The point is rather that the minimal theory does not require any particular one of them. So that someone who wishes to avoid commitment to 'propositions' of any specific sort need not on that score object to the conception of truth that will be elaborated here.

Moreover, the view that truth is not strictly speaking attributable to utterances, or to linguistic or mental acts, is not substantial, and nothing of importance in what follows will depend on it. If someone holds that an utterance may be 'true', in a certain sense, then he can simply regard my claims about the property of *expressing truth* as claims about 'truth' in his sense. Similarly for those who think that a truth predicate may be applied to acts of asserting, states of believing, etc.

2. What are the fundamental principles of the minimal theory of truth?

The axioms of the theory are propositions like

    (1) ⟨⟨Snow is white⟩ is true iff snow is white⟩

and

    (2) ⟨⟨Lying is wrong⟩ is true iff lying is wrong⟩;

that is to say, all the propositions whose structure is

    (E*) ⟨⟨*p*⟩ is true iff *p*⟩.[2]

---

[2] This claim will be modified slightly in the answer to Question 10 in order to accommodate the 'liar' paradoxes.

In order to arrive at this 'propositional structure' we can begin with any one of the axioms and note that the sentence expressing it may be divided into two complex constituents. First there is a part that is itself a sentence and which appears twice. In the case of (1), this is

(3)	'snow is white'.

And second there is the remainder of the axiom-formulation—namely, the schema

(E)	'$\langle p \rangle$ is true iff p'.

Now we assume that if a complex expression results from the application of a schema to a sequence of terms, then the meaning of the expression is the result of applying the meaning of the schema to the sequence of the terms' meanings. In particular, since, as we have just seen, the sentence

(1*)	'$\langle$Snow is white$\rangle$ is true iff snow is white'

is the result of applying (E) to (3), then the proposition expressed by (1*) is the result of applying what is expressed by schema (E) to what is expressed by (3). That is to say, the axiom (1) is the result of applying the propositional structure (E*) to the proposition

(3*)	$\langle$snow is white$\rangle$.

Similarly, if (E*) is applied to the proposition

(5)	$\langle$Lying is wrong$\rangle$,

it yields

(2)	$\langle\langle$Lying is wrong$\rangle$ is true iff lying is wrong$\rangle$.

Indeed, when applied to any proposition, y, this structure (or function) yields a corresponding axiom of the minimal theory, MT.[3]

---

[3] I am employing the convention that surrounding any expression, e, with angled brackets, '$\langle$' and '$\rangle$', produces an expression referring to *the propositional constituent expressed by e*.

It might be argued that the two sentence-tokens in each MT axiom-formulation do not have the same content as one another, since the first occurs in an opaque context (after 'The proposition that . . .') and the second does not, and, consequently,

In other words, the axioms of MT are given by the principle

(5)   For any object x: x is an axiom of the minimal theory if and only if, for some y, when the function E* is applied to y, its value is x.

Or in logical notation

that these axioms cannot really be regarded as the various results of applying a single one-place function (the propositional structure) to the various propositions. If this is correct (which I doubt) then we must proceed differently.

One alternative is to characterize the axioms of the minimal theory as anything that is expressed by instances of the sentence schema

(E)   '$\langle p \rangle$ is true iff p'.

However, the theory cannot be restricted to instantiations of (E) by *English* sentences; for presumably there are propositions that are not expressible in current English, and the question of *their* truth must also be covered. So further 'equivalence axioms' are needed, one for each unformulatable proposition.

Although we cannot now *articulate* these extra axioms (any more than we can articulate the propositions they are about), we can nevertheless *identify* them. One way of doing this is by reference to *foreign* languages. We can suppose that the theory of truth includes whatever is expressed by instances of *translations* of the equivalence schema: e.g. instantiations of

(E-f)   '$\langle p \rangle$ est vrai ssi p'

by French sentences, instantiations of

(E-g)   '$\langle p \rangle$ ist wahr gdw p'

by German sentences, and so on, for all languages. If it were assumed that every proposition is expressed in some language, then this would do. But we want to allow for the existence of propositions that are not yet expressible at all. To accommodate these we might suppose that every proposition, though perhaps not expressed by any *actual* sentence, is at least expressed by a sentence in some *possible* language. And we can then regard the theory of truth as whatever would be expressed by instances of translations of the equivalence schema into possible languages.

However, once the need to refer to possible languages has been acknowledged, we can see that there was no reason to have brought in *actual foreign languages*. For we can make do with our own language supplemented with possible extensions of it. In other words, we can characterize the 'equivalence axioms' for unformulatable propositions by considering what would result if we *could* formulate them and *could* instantiate those formulations in our equivalence schema. Thus we may specify the axioms of the theory of truth as what are expressed when the schema

(E)   '$\langle p \rangle$ is true iff p'

is instantiated by sentences in any possible extension of English.

Alternatively, instead of identifying the axioms indirectly in terms of how they would be *expressed*, we can solve the problem by directly specifying the *propositional* structure which all and only the axioms have in common. This is the strategy adopted in the text.

(5*) $(x)(x$ is an axiom of MT $\leftrightarrow (\exists y)(x = E^*(y)))$.[4]

The minimal theory has several striking features—features that might at first be regarded as grounds for dissatisfaction with it. In the first place it does not say explicitly *what truth is*: it contains no principle of the form '$(x)(x$ is true iff . . . x . . .)', or 'What makes a proposition true is its having characteristic P'. And so one might suspect that certain general facts about truth could not be explained by the theory. Secondly, it does not mention phenomena such as reference, logical validity, assertion, and the aim of inquiry—notions whose relation to truth one might have thought any decent theory should describe. And, thirdly, although we have been able to characterize the axioms of MT (as the propositions of a certain form) we cannot explicitly formulate the theory—for two independent reasons. In the first place the number of axioms that we have the terminology to formulate is too great; there are infinitely many, and though each one of them can be expressed, it is not possible to write down the whole collection. In the second place there are many propositions we cannot express in current terminology. And for those the corresponding equivalence axioms are themselves inexpressible—although, as we have seen, it is none the less possible to indicate what they are.

In the following few sections we shall examine our justification for concluding that MT is nevertheless the best theory of truth, and we shall see why the peculiar features of the theory should not be held against it.

3. It seems unlikely that instances of the equivalence schema could possibly suffice to explain all the great variety of facts about truth.

The primary test of this (and any other) theory is its capacity to accommodate the phenomena in its domain. That is to say, if our

---

[4] Patrick Grim pointed out to me that the minimal theory cannot be regarded as *the set* of propositions of the form $\langle\langle p \rangle$ is true iff $p \rangle$; for there is no such set. The argument for this conclusion is that if there were such a set, then there would be distinct propositions regarding *each* of its subsets, and then there would have to be distinct axioms of the theory corresponding to those propositions. Therefore there

theory is a good one, it will be able to account for all the facts about truth. Let me give three examples of the sort of explanation that minimalism can provide.

(I) From 'What Smith said was true' and 'What Smith said was that snow is white', it follows that 'Snow is white'. Given the minimal theory (MT) this fact can be explained as follows:

1. What Smith said is true.
2. What Smith said = ⟨Snow is white⟩.
∴ 3. ⟨Snow is white⟩ is true.   [from 1,2]
4. ⟨Snow is white⟩ is true iff snow is white.   [MT]
∴ 5. Snow is white.   [from 3,4][5]

(II) If one proposition implies another, and the first one is true, then so is the second. Here is a minimalist explanation:

1. Logic provides us with facts like

     [dogs bark & (dogs bark → pigs fly)] → pigs fly,

that is, with every fact of the form

     $[p \ \& \ (p \rightarrow q)] \rightarrow q.$

would be a 1–1 function correlating the subsets of MT with some of its members. But Cantor's diagonal argument shows that there can be no such function. Therefore, MT is not a set. In light of this result, when we say things like '⟨A⟩ follows from the minimal theory', we must take that to mean, not that the relation of *following from* holds between ⟨A⟩ and a certain entity, the minimal theory; but rather that it holds between ⟨A⟩ and *some part* of the minimal theory—i.e. between ⟨A⟩ and some set of propositions of the form ⟨⟨p⟩ is true iff p⟩.

    [5] In order to explain why '*Possibly*, snow is white' follows from 'What Smith said is *possibly* true' and 'What Smith said is that snow is white', we must assume, not merely statement 4, but rather

Necessarily, ⟨snow is white⟩ is true iff snow is white.

    Thus it might seem that the axioms of the theory of truth should be strengthened and taken to consist of *modal* propositions of the form

⟨*Necessarily*, ⟨p⟩ is true iff p⟩.

    An alternative strategy, however—and one that I prefer—is to keep the theory of truth un-modal and simple, and instead *derive* the necessity of its axioms from a separate theory of necessity, specifying, in general, what makes a proposition not merely true but necessarily true. It might be supposed, for example, that the *necessary* truths are distinguished by being *explanatorily fundamental*. In that case, given our argument to the effect that MT is explanatorily basic, it would follow that its axioms are necessary. Thus we might obtain the necessity of instances of the equivalence schema without having to build it into the theory of truth itself.

2. Therefore, given MT, we can go on to explain every fact of the form

   $[\langle p \rangle$ is true $\& (p \rightarrow q)] \rightarrow \langle q \rangle$ is true.

3. But from the nature of implication, we have all instances of

   $(\langle p \rangle$ implies $\langle q \rangle) \rightarrow (p \rightarrow q)$

4. Therefore we can explain each fact of the form

   $[\langle p \rangle$ is true $\& \langle p \rangle$ implies $\langle q \rangle] \rightarrow \langle q \rangle$ is true.

5. And therefore, given MT, we get each fact of the form

   $\langle [\langle p \rangle$ is true $\& \langle p \rangle$ implies $\langle q \rangle] \rightarrow [\langle q \rangle$ is true]$\rangle$ is true.

6. But it is a peculiar property of propositions that any general claim about them—any characterization of *all* propositions—is made true by the infinite set of particular facts associating that characteristic with each individual proposition.[6]

7. Therefore, in light of 5 and 6, we can explain the general fact:

   Every proposition of the form, $\langle [\langle p \rangle$ is true $\& \langle p \rangle$ implies $\langle q \rangle] \rightarrow [\langle q \rangle$ is true]$\rangle$, is true.

(III) We would be inclined to endorse the following thesis: 'If all Bill wants is to have a beer, and he thinks that merely by nodding he will get one, then, if his belief is true, he will get what he wants.' This fact would be explained as follows:

We begin with the suppositions

1. Bill wants $\langle$Bill has a beer$\rangle$;
2. Bill believes $\langle$Bill nods $\rightarrow$ Bill has a beer$\rangle$.

In addition, we can make the normal assumption (an instance of the 'practical syllogism') about the relation between Bill's belief, desire, and action:

3. [Bill wants $\langle$Bill has a beer$\rangle$ &
   Bill believes $\langle$Bill nods $\rightarrow$ Bill has a beer$\rangle$]
   $\rightarrow$ Bill nods;   [premise]

∴ 4. Bill nods.   [from 1,2,3]

---

[6] See sect. 5 of the Postscript for further discussion of this premise.

Now let us assume, for the sake of argument,

    5. Bill's belief is true.

That is to say,

    6. ⟨Bill nods → Bill has a beer⟩ is true.   [from 2,5]

Also we have, from the theory of truth,

    7. ⟨Bill nods → Bill has a beer⟩ is true iff
       Bill nods → Bill has a beer;   [MT]

∴ 8. Bill nods → Bill has a beer;   [from 6,7]

∴ 9. Bill has a beer.   [from 4,8]

But again from the theory of truth,

    10. ⟨Bill has a beer⟩ is true iff Bill has a beer;   [MT]

∴ 11. ⟨Bill has a beer⟩ is true;   [from 9,10]

∴ 12. Bill gets what he wants.   [from 1,11]

And, as we shall see in Chapter 3 (in the answer to Question 11), this sort of explanation may be universalized to show in general how true beliefs engender successful action.

According to the minimalist thesis, all of the facts whose expression involves the truth predicate may be explained in such a way: namely, by assuming no more about truth than instances of the equivalence schema.[7] Further explanations of this sort, dealing with a range of philosophically interesting facts about truth, will be given as we proceed. These explanations will confirm the minimalist thesis that no account of the *nature* of truth, no principle of the form '$(x)(x$ is true iff . . . x . . .)', is called for.

> **4.** The minimal theory must be incomplete, for it says nothing about the relationships between truth and affiliated phenomena such as verification, practical success, reference, logical validity, and assertion.

A theory of any phenomenon, X, is a collection of principles (i.e. axioms and/or rules); and the theory is *good* to the extent that it

---

[7] This is a slight exaggeration. As Anil Gupta (1993*b*) has pointed out, the equivalence axioms cannot explain why, for example, Julius Caesar is not true. It would seem that a complete theory of truth will require, in addition to MT, the axiom '$(x)(x$ is true → x is a proposition)'.

captures all the facts about that phenomenon in the simplest possible way. It won't do merely to produce some set of important facts about X and call that the theory. Nor would it suffice even if *every* fact about X were explicitly listed. Rather, the understanding that we want requires some account of explanatory relationships. We have to locate the most basic facts regarding X, from which all the others may be explained. Of course we don't expect our theory of X to do the explanatory work all by itself. It does not follow solely from the theory of electrons that electrons are smaller than elephants; we need a theory of elephants too. Our goal, then, is to find a simple theory of X, which, together with our theories of other matters, will engender all the facts.

Sometimes it will turn out, for certain phenomena, X and Y, that we cannot separate two distinct theories, one for X and one for Y: the simplest adequate body of principles we can find concerns both X and Y. Consider, for example, a geometric theory about points, lines, angles, etc. This cannot be split up into a theory of points, a theory of lines, and so on. Sometimes we are forced to acknowledge that certain theoretical phenomena are, in this way, inextricably entangled with one another. And this is a significant fact about such phenomena. But when this is *not* so, where distinct theories of X and Y *can* be given, then they *should* be given. Otherwise, a misleading illusion of interdependence is conveyed and the cause of simplicity and explanatory insight is poorly served.

For this reason it seems to me not merely legitimate but important to separate, if we can, what we say about truth from our theories of reference, logic, meaning, verification, and so on. No doubt there are interesting relationships amongst these matters. But in so far as we want to *understand* truth and the other phenomena, then our task is to *explain* the relationships between them and not merely to recognize that they exist. We must discover the simplest principles from which they can all be deduced: and simplicity is promoted by the existence of separate theories of each phenomenon. Therefore it is quite proper to explain the properties of truth by conjoining the minimal theory with assumptions from elsewhere. (Note, for example, the use of extraneous premises in the explanations in the previous section—drawn, in those cases, from psychology, logic, and the theory of propositions.) The virtue of minimalism, I claim, is that *it provides a theory of truth that is a theory of nothing else, but*

which is sufficient, in combination with theories of other phenomena, to explain all the facts about truth.[8]

5. Even if the minimal theory is, in some sense, 'adequate' and 'pure', it is nevertheless unsatisfactory, being so cumbersome that it cannot even be explicitly formulated.

Presented with the minimal theory of truth one's first instinct, no doubt, is to imagine that we can surely improve on it and capture the infinity of instances of the equivalence schema in a compact formulation. However, there does not *have* to be any succinct, explanatorily adequate theory of truth, and I shall be arguing that in fact there isn't one. Such a theory would encapsulate the properties of truth in a *finite* body of principles which would generate everything true of truth, including, at the very least, infinitely many instances of '$\langle p \rangle$ is true iff p'. Moreover, if it is to be explanatorily adequate, the theory would have to subsume all these facts without the use of notions that are themselves mysterious and unexplained. But how might this be done?

One natural suggestion is the single principle

(6) $(x)(x$ is true iff $\{\exists q\}(x = \langle q \rangle \ \& \ q))$,

where the curly brackets indicate *substitutional quantification* over the sentences of English. But this idea fails for a couple of reasons. In the first place, the use of substitutional quantification does not square with the *raison d'être* of our notion of truth, which is to enable us to do *without* substitutional quantification.[9] In the second place, the notion of substitutional quantification would itself require theoretical elucidation. But what kind of elucidation could be given? It would be circular to rely on the standard explanation which is couched in terms of truth.[10] Alternatively, one might try to

---

[8] Although the minimal theory characterizes truth in relation to propositions, it is not a joint theory of these two phenomena; for we can give a prior theory of propositions which makes no reference to truth (see Chapter 6). Thus truth is the only previously unexplained concern of the minimal theory.

[9] See Chapter 1, n. 1, and Question 6 of the present chapter.

[10] '$\{\exists p\}(\ . \ . \ . \ p \ . \ . \ . \ )$' means 'Some sentence formed by replacing the "p" in "$. \ . \ . \ p \ . \ . \ .$" with a sentence of (some extension of) English is true'.

characterize the notion of substitutional quantification via a specification of the rules of inference that govern it: for example, a version of 'universal instantiation', which is the schematic rule

$$(7) \quad \frac{\{q\}(\ldots q \ldots)}{\ldots p \ldots}.$$

However, this cannot be formulated as a generalization over every sentence 'p', viz.

$$(8) \quad \{p\} \frac{\{q\}(\ldots q \ldots)}{\ldots p \ldots}.$$

For there would then be no way of getting from that principle to instances such as:

$$(9) \quad \frac{\{q\}(\ldots q \ldots)}{\ldots \text{snow is white} \ldots}.$$

Nor, again on pain of circularity, could we construe (7) as the claim that every instance of the schema preserves *truth*. The only alternative would be to recognize that the apparently *single* rule (7) is in fact an infinite collection of rules; one for each 'sentence in context' that can be put in place of '$\ldots p \ldots$'. But then we are embracing an unformulable theory after all. Nothing has been gained; yet something (i.e. the use of truth to *dispense* with substitutional quantification) has been lost. Thus it seems that the best overall theory will not involve a definition of truth in terms of substitutional quantification.[11]

Another tempting approach, again designed to avoid the need for infinitely many axioms, is to formulate the theory of truth as the single proposition

(10)   Every instance of $\langle\langle p\rangle$ is true iff $p\rangle$ is true.

It is clear, however, that this will not do. For it would enable us to deduce, for example,

---

[11] For further discussion of the policy of explaining truth in terms of substitutional quantification see Grover, Camp, and Belnap (1975), Baldwin (1989), and Brandom (1994), who embrace it, and Forbes (1986), who rejects it. Similar difficulties beset the prospect of explaining truth in terms of other forms of quantification into sentence positions.

(11) ⟨⟨Snow is white⟩ is true iff snow is white⟩ is true.

But we would have no licence to get from there to the conclusion that

(1) ⟨Snow is white⟩ is true iff snow is white.

To do this we would need the schematic rule of inference

$$\text{(E\#)} \quad \frac{\langle p \rangle \text{ is true}}{\therefore \ p},$$

which, as we have just seen, must be regarded as an infinite collection of separate rules.

Inspired by Tarski (1958), one might think that the solution to our difficulty is to be found by defining the truth of a whole proposition in terms of the reference of its parts and how those parts are put together. But this is a vain hope. Truth and reference are closely affiliated notions, and so a theory that characterized truth in terms of reference but gave no account of reference would be unsatisfactory. But any attempt to provide such a theory re-encounters precisely the problems with which we are now struggling. For just as the theory of truth must subsume everything like

(1) ⟨Snow is white⟩ is true iff snow is white,

so any decent theory of reference would have to subsume the fact that

(12) The propositional constituent expressed by the word 'Aristotle' refers (if at all) to Aristotle,

and so on. It might be thought that in the case of reference this problem may be solved easily—by simply listing the referents of each of the finitely many primitive terms in our language. But this is not so. Just as our understanding of truth goes beyond the list of presently formulable instances of the equivalence schema and tells us that any new sentences could also be instantiated, similarly, our conception of reference goes beyond a knowledge of the referents of our current primitive vocabulary.[12] It covers a potentially infinite

---

[12] This point is stressed by Max Black in his critique of Tarski's theory (Black, 1948).

number of new terms. Consequently, we are pushed into formulations such as

(13) $(x)(y)(x$ refers to y iff $\{\exists d\}(x = \text{`}d\text{'} \ \& \ y = d))$,

where the substitutional variable, d, ranges over singular terms in possible extensions of our language. Thus we find ourselves relying again on substitutional quantification and the need to explain it with an infinite number of rules; so the reduction of truth to reference has turned out to be futile.

Let me make some further points about the search for a finite axiomatization of MT. In the first place, problems exactly like those I have just mentioned arise for the project of explaining truth in terms of *predicate satisfaction*. We would need to add a theory of satisfaction that could encompass all facts like

(14)   The predicate 'blue' is satisfied by blue things;

and once again no finite list could suffice. An adequate theory would have to contain infinitely many propositions of the form

(15)   The propositional constituent associated with the predicate 'F' is satisfied by, and only by, things that are F.

Therefore concern about the infinite character of the minimal theory of truth cannot be assuaged by explaining truth in terms of satisfaction.

Secondly, these conclusions do not tell against Tarski's *own* project, in so far as it aspired merely to explicate a notion of 'true-in-L' for certain highly artificial languages, L. Each of these languages has a fixed stock of primitives, so it is possible to explicate 'refers-in-L' and 'satisfies-in-L' with finite lists of principles. Our project, however, is in certain respects more ambitious than Tarski's. We are aiming for a theory of 'being true'—a property which is attributed to propositions regardless of how or whether they are expressed. Similarly we are looking for a theory of 'expressing truth'—a property we may attribute to an utterance regardless of the language in which it is couched. I have been considering the possibility that someone might hope, in defining 'true', to exploit the strategy that Tarski used in his definition of 'true-in-L'; but this will not work—or so I have argued.[13]

---

[13] For further discussion of Tarski see the answer to Question 38.

Thirdly, it might be thought that the difficulty in obtaining a finite, compositional theory of truth stems from the implicit assumption that propositions are constructed, as Frege said, from *the senses of words*—which are entities that require some theory of reference—and that such problems do not arise if propositions are instead constructed, as Russell proposed, from concrete objects and properties. For in that case we can say

(ET) $(x)(R)(S)[x$ is the proposition consisting of the n-place relation, R, and the sequence, S, of objects $\rightarrow$ (x is true $\leftrightarrow S$ exemplifies $R)]$.

But although this may be fine as far as it goes, it does not go far enough. For it would have to be supplemented with a theory of *exemplification*; and here is where the old troubles will emerge. We might be able to derive from (ET)

(1**)   ⟨Snow is white⟩ is true iff snow exemplifies whiteness.

But in order to get the conclusion that

(1)   ⟨Snow is white⟩ is true iff snow is white

we would need the schema

x exemplifies R_1-ness iff R_1x,

and in general we need schemata for each higher value of n as well, that is,

(Ex) $\langle x, y \rangle$ exemplifies R_2-ness iff R_2xy, $\langle x,y,z \rangle$ exemplifies R_3-ness iff R_3xyz, . . . and so on.

And this apparatus is no less infinitary than the minimal theory. Thus it isn't any easier to give a finite theory of truth if we focus on Russellian, 'concrete' propositions.

An alternative idea involves the conception of a proposition as a set of possible worlds. It may seem that the right definition of truth for such propositions is not the equivalence schema but rather the principle that

(W) x is true = the actual world is a member of x.[14]

---

[14] Suggested by Harty Field (1992).

But again this offers no advantages over the minimal theory. For we would have, for example,

(1***)   ⟨Snow is white⟩ is true iff the actual world is a member of ⟨snow is white⟩.

And we are assuming the schema

(P) $⟨p⟩ = \{w \mid p \text{ in } w\}$.

Therefore we could infer

(1****)   ⟨Snow is white⟩ is true iff snow is white in the actual world.

But at this point we are stuck. In order to derive the MT axiom we would need

          Snow is white in the actual world iff snow is white.

And to get all of the axioms we would need every instance of the schema

(MTA) p in the actual world iff p.

Thus finite axiomatization is not achieved by the explanation of truth in terms of actuality.

Finally, notice that no help is to be found by looking in the direction of traditional theories of truth such as the coherence and pragmatic approaches, or by entertaining some other way of identifying truth with a naturalistic property. For whatever property, F, is associated with truth, we will be able to explain instances of the equivalence schema only to the extent that we can explain instances of the schema

(16) $⟨p⟩$ is F iff p.

And *this* infinite theory will be no easier to encapsulate than the minimal theory.

I conclude that we should not expect to contain all instances of the equivalence schema within a finite formulation: an infinity of axioms is needed. And since this would seem to be an unavoidable feature of any adequate theory of truth, it should not be held against MT. Therefore we must acknowledge that the theory of truth can-

not be explicitly formulated. The best we can do is to give an implicit specification of its basic principles.[15]

> **6.** If there were really no more to a complete theory of truth than a list of biconditionals like 'The proposition *that snow is white* is true if and only if snow is white', then since one could always say '*p*' rather than 'The proposition *that p* is true', it would be inexplicable that our language should contain the word 'true': there would be no point in having such a notion.

This argument has already been dealt with; but it is often raised against what are sometimes called 'redundancy' accounts of truth, so let me repeat my response. First, the fact that the only applications of truth expressly contained in the theory are within propositions of the form

(E*) $\langle\langle p \rangle$ is true iff $p \rangle$

does not imply that the theory covers only those cases in which truth is attributed to an articulated proposition. For suppose 'Einstein's law' refers to the proposition $\langle E = mc^2 \rangle$. In other words,

(17) $\langle E = mc^2 \rangle$ = Einstein's law.

In that case the theory of truth, which applies in the first instance to

(18) $\langle E = mc^2 \rangle$ is true,

must apply indirectly to

(19)   Einstein's law is true,

from which 'is true' cannot be removed. And it is from its role in such sentences that the truth predicate gets its value. To see this, consider how we would manage without it. We would have to put the matter roughly as follows:

---

[15] For further discussion of why we should not expect MT to be explained in terms of more fundamental axioms, see the answer to Question 14.

(20)   (*x*)(If Einstein's law is the proposition *that x*, then *x*).

But this could not be construed in the usual manner. For, given the usual conventions of quantification, that sentence is ill-formed in two distinct ways: the second occurrence of '*x*' is in an opaque context, beyond the reach of normal quantification; and a variable that ranges over *objects* appears in sentential positions. In order to avoid these incoherences it would be necessary to introduce a new form of quantification—substitutional quantification—that *could* legitimately govern opaque contexts and sentence positions. That is to say, we need a quantifier,

(21)   {*p*}(. . . *p* . . .),

whose meaning is not

(22)   Every object, *p*, satisfies '. . . *p* . . .',

but rather

(23)   Every grammatical substitution of a declarative sentence of English in place of '*p*' in '. . . *p* . . .' yields a truth.

But such a quantifier, with its special syntactic and semantic rules, would be a cumbersome addition to our language. The point of our notion of truth is that it provides a simple alternative to this apparatus. For, as I mentioned in Chapter 1, the truth predicate allows any sentence to be reformulated so that its entire content will be expressed by the new subject—a singular term open to normal objectual quantification. In other words, '*p*' becomes '⟨*p*⟩ is true'. Therefore, instead of

(24)   {*p*}(*p* → *p*),

we can say

(25)   (*x*)(If *x* is a proposition of the form '*p* → *p*', then *x* is true).

Instead of

(26)   {*p*}(If Einstein's law is the proposition *that p*, then *p*),

we can say

(27) (x)(If x is a proposition of the form ⟨If Einstein's law is
       the proposition that p, then p⟩, then x is true),

which is logically equivalent to

(28) (x)(If x = Einstein's law, then x is true).

And, in general, instead of the substitutionally quantified

(21) $\{p\}(\ldots p \ldots)$,

we can make do with the ordinary, objectually quantified

(29) (x)(If x is a proposition of the form ⟨ $\ldots p \ldots$ ⟩, then x
       is true).

I am not suggesting, of course, that the truth predicate was introduced *deliberately* to perform this useful function. But I *am* supposing that its usefulness, as just described, is what explains its
presence. For if it were not valuable at all, it would presumably fall
out of use; and as for alternative functions that it might have, there
simply aren't any plausible candidates.

7. The minimal theory fails to specify what are *meant* by
attributions of truth. It fails to provide necessary and
sufficient conditions for the applicability of the truth
predicate.

The second part of this point is quite correct, but does not justify
the initial complaint. For it is not the case that a satisfactory characterization of the meaning of a predicate must take the form of
necessary and sufficient conditions for its correct application—i.e.
an explicit, eliminative *analysis*. A definition of that sort is merely
one particularly simple way of specifying the use of a word; but we
should be open to more complex ways of doing it. So the present
objection presupposes needless restrictions on what sort of
definition of 'true' is needed. Once these implicit constraints are
loosened, the minimalist account will no longer seem inadequate.

   I can perhaps clarify this response by distinguishing some different forms that a definition of 'true' might be thought to take. In the

first place one might offer an *atomic definition*: that is, a definition of the familiar form

(30)   'true' means ' . . . ',

supplying a synonym that would permit us to eliminate the word 'true' in a uniform way from every context in which it appears. An example of an atomic definition is the definition of 'bachelor' as 'unmarried man'. The pragmatists' identification of truth with utility has this character.

In the second place, and a little more modestly, one might offer a *contextual definition*: that is, a set of rules that would allow the conversion of any sentence containing the word 'true' into a synonymous sentence that does not contain it. A well known example of this style of definition is Russell's theory of definite descriptions (Russell, 1905):

(31)   'The F is G' means 'Some G is the same as every F',

which reduces the definite article, 'the', to the notions of predicate logic—specifically, 'some', 'every', and 'the same as'. A partial account of truth along these lines would be contained in the schema

(E!)   'It is true that p' means 'p'.

Thirdly, one might abandon the attempt to provide the sort of account that would enable the word 'true' to be *eliminated*, and aim instead for *implicit definition*: that is, a set of principles involving the truth predicate, our commitment to which fixes its meaning. For example, it is sometimes said that the axioms of any geometry implicitly determine the meanings of the terms 'point' and 'line', at least as they are used when proving theorems of that geometry. An account of the meaning of 'true' along these lines would be given by the substitutionally quantified principle

(E+) $\{p\}(x)(x = \langle p \rangle \rightarrow (x \text{ is true} \leftrightarrow p))$.

Finally, one might deny that the meaning of the truth predicate can be captured in our commitment to any definite body of principles. One might hold that the use—hence the meaning—of 'true' is given by regularities with a more complex structure than simply: '*We accept "A"*.' An example of this sort of *use definition* is the idea that our conception of number is determined by the disposition to

accept Peano's axioms, including infinitely many instance of the induction schema

(32) $\{F(0) \mathrel{\&} (n)[F(n) \to F(n + 1)]\} \to (n)[F(n)]$.

Another example is provided by a certain account of the meaning of counterfactual implication, namely,

(33)   'If p were true, then q would be true' is assertible to degree x iff it is known that x is the empirical tendency for q to be true in circumstances in which p is true and in which all the facts causally and conceptually independent of not-p still obtain.

This rule characterizes a certain sense of 'If . . ., then . . .' by specifying the appropriate level of confidence for any such conditional, and without involving any principles, in the material mode, relating counterfactual dependence to other aspects of reality.[16]

I would suggest that the truth predicate belongs in this final category. Our understanding of 'is true'—our knowledge of its meaning—consists in the fact that the explanatorily basic regularity in our use of it is the inclination to accept instantiations of the schema

(E)   'The proposition *that p* is true if and only if p',

by declarative sentences[17] of English (including any extensions of English). Thus for a normal English speaker it consists (a) in his disposition to accept

(MT)   'The proposition *that snow is white* is true iff snow is white', 'The proposition *that I am hungry* is true iff I am hungry', 'The proposition *that Paris is beautiful* is true iff Paris is beautiful', . . .,

and (b) in the fact that his acceptance of other sentences containing the truth predicate is explained in terms of (a).

---

[16] For elaboration of this account of counterfactual dependence see Horwich, 1987.

[17] Such sentences may be identified by their *meanings*. Moreover it is required that their two tokens express the same proposition. If meaning and propositions were then to be defined in terms of truth we would have a vicious circle. However, I argue in my answers to Questions 22 and 32, and in sect. 3 of the Postscript, that meaning and proposition may be explained in terms of aspects of use (including *acceptance*) that do not presuppose the notion of truth.

The minimalist account of what it is to know the meaning of the truth predicate does not provide an *analysis* and does not enable us to specify in non-circular terms the *content* of attributions of truth. This is precisely what distinguishes it from traditional approaches. But it may be none the less a perfectly acceptable account of what our understanding the truth predicate consists in, just so long as it is capable of explaining all pertinent linguistic behaviour—all our ways of deploying the term 'true'.[18] The question, in other words, is whether we can explain, on the hypothesis that the equivalence schema governs some person's use of the truth predicate, why, for example, that person should endorse an inference from 'What Oscar said is true' and 'What Oscar said is that eels are good' to 'Eels are good'; and, in general, why he uses the truth predicate in just the way that he does. This is the adequacy condition for a theory of the meaning of the truth predicate; and, judging by the examples in the answers to Question 3, the minimalist account would appear to satisfy it.

8. Is the minimalist conception concerned with *truth itself* or with *the word 'true'*?

It is concerned with both—and other intimately related matters as well. However it is important to separate the different questions it addresses. Specifically we should distinguish between:

1. A theory of the *function* of the truth predicate;
2. A theory of what it is for someone to *understand* the word 'true';
3. A theory of the *meaning* of the word 'true';
4. A theory of what it is to have, or grasp, the *concept* of truth;
5. A theory of *truth itself*.

---

[18] See sect. 2 of the Postscript for further discussion. Notice that since minimalism does not provide an explicit definition of truth, it superficially resembles Moore's view that truth is an 'inexplicable quality' (Moore, 1899). The important difference between the two accounts, however, is that minimalism nevertheless purports to give, by means of the equivalence schema, a complete account of truth and of what our grasp of it consists in, whereas on Moore's view it is impossible to shed any light on these matters and truth remains impenetrably mysterious.

According to the minimalist conception, the function of the truth predicate is to enable the explicit formulation of schematic generalizations. Our understanding of the word is constituted by the practice of using it to perform that function—a practice whose basic regularity is an inclination to accept instances of the equivalence schema, '$\langle p \rangle$ is true iff p'. The concept of truth (i.e. what is meant by the word 'true') is that constituent of belief states expressed in uses of the word by those who understand it—i.e. by those whose use of it is governed by the equivalence schema. And the theory of truth itself—specifying the explanatorily fundamental facts about truth—is made up of instances of that schema. Thus, the minimal theory of truth will provide the basis for accounts of the meaning and function of the truth predicate, of our understanding it, of our grasp upon the concept of truth, and of the character of truth itself.[19]

> **9.** Even if we grant that, as predicates go, the *truth* predicate is highly unusual—even if we grant that its special function is to enable us to say certain important things while avoiding new forms of quantification—it surely does not follow that *being true* is not a genuine property.

Quite right. And it is not part of the minimalist conception to maintain that truth is not a property. On the contrary, 'is true' is a perfectly good English predicate—and (leaving aside nominalistic concerns about the very notion of 'property') one might well take this to be a conclusive criterion for standing for a property of *some* sort. What the minimalist wishes to emphasize, however, is that truth is not a *complex* or *naturalistic* property but a property of some other kind. (Hartry Field (1992) suggests the term '*logical* property'.) The point behind this jargon is that different kinds of property correspond to different roles that predicates play in our language, and that unless these differences are appreciated, we will be tempted to raise questions regarding one sort that can legitimately

---

[19] See sect. 5 of the Postscript for more detailed discussion of the answer to this question.

arise only in connection with another sort. A familiar example of this phenomenon derives from the predicate 'exists'. Another, more controversial, case is the conflation of normative and descriptive properties. According to minimalism, we should, for similar reasons, beware of assimilating *being true* to such properties as *being turquoise*, *being a tree*, or *being made of tin*. Otherwise we will find ourselves looking for its constitutive structure, its causal behaviour, and its typical manifestations—features peculiar to what I am calling '*complex*' or '*naturalistic* properties'. We will be puzzled when these expectations are inevitably frustrated, and incline to the conclusion that the nature of truth is profoundly obscure—perhaps even incomprehensible.

As I have indicated, some philosophers hold that *no* predicate refers and that properties do not exist; and, of course, from that nominalistic point of view the particular question 'whether truth is a property' does not arise—at least, in those words. However the underlying issue is still with us in the form of whether or not applications of the truth predicate engender *statements* about the propositions to which it is applied. The thesis that they *do* distinguishes the present view from certain more radical formulations of deflationism—those according to which it is a grammatical illusion to think that

(34) X is true

makes a statement of *any* kind about the proposition X. For example, it was suggested by Frege (1891, 1918), Ramsey (1927), and Ayer (1935, 1936) that the forms

(35) p

and

(36)   It is true *that p*

yield the same sense no matter what declarative English sentence is substituted for 'p'. This is often referred to as 'the redundancy theory of truth' and it evidently conflicts with the view advanced here which associates a definite propositional constituent to the truth predicate—a constituent which is part of one of these propositions but not of the other. Similarly, from the present perspective we are rejecting the idea due to Strawson (1950) and Ayer (1963) that the truth predicate is not used to give descriptions or make statements

about the things to which it is applied, but that it is used instead to perform quite different speech acts: endorsing, agreeing, conceding, etc.

The trouble with the 'redundancy/performative' conception is that it cannot be squared with obvious facts about the character and function of truth. It addresses only cases like

(37)   ⟨Snow is white⟩ is true,

in which the truth predicate is attached to an explicitly articulated proposition. And it maintains, with a certain prima facie plausibility, that the whole sentence has the same sense as the constituent

(38)   Snow is white.

But notice that such uses of truth have no great value: we could easily do without them. And when we turn to genuinely useful attributions, as for example in

(39)   Oscar's claim is true,

the theory has nothing to say about its sense, except that the logical form is supposedly *not* what it would seem to be: i.e. not

(40)   *X* is *F*.

Consequently, the redundancy theory is quite unable to account for the inference from (39) and

(41)   Oscar's claim = the proposition *that snow is white*

to

(37)   The proposition *that snow is white* is true,

and hence to

(38)   Snow is white

—which is precisely the sort of reasoning on which the utility of our concept of truth depends.[20] Thus the redundancy/performative

---

[20] Similar objections to the redundancy theory have been made by Tarski (1943/4), Thomson (1948), Cohen (1950), Ziff (1962), and Ezorsky (1963). A redundancy theorist might attempt to explain the above inference by first analysing '*X* is true' in terms of substitutional quantification as

(SB) $\{\exists p\}(X = \langle p \rangle \; \& \; \langle p \rangle$ is true)   (*cont.*)

theory must be rejected. No doubt we do perform all kinds of speech act (such as *agreeing* and *conceding*) with the truth predicate. But, as Warnock observed (1964), it is best to say that we do so *by* (not *instead of*) making a statement—that is, *by* attributing the property, truth, to the proposition in question. Just as the assertion

(42)   Your article is brilliant

may be intended to achieve an effect beyond conveying accurate information, so one might well have some ulterior purpose in mind in saying

(39)   Oscar's claim is true,

yet none the less be making a statement about Oscar's claim: i.e. attributing a property to it.

> **10.** If the equivalence schema is relied on indiscrimin-
> ately, then the notorious 'liar' paradoxes will result.

Indeed—and for that reason we must conclude that permissible instantiations of the equivalence schema are restricted in some way so as to avoid paradoxical results. To see this, let '#' abbreviate 'THE PROPOSITION FORMULATED IN CAPITAL LETTERS IS NOT TRUE'. Then assuming for the sake of argument that

(43)   ⟨#⟩ is true,

and given the pertinent instance of the equivalence schema, namely

(44)   ⟨#⟩ is true iff #,

we can infer

and then taking this to be synonymous with

(ST) $\{\exists p\}(X = \langle p \rangle \,\&\, p)$.

However, although this strategy is clearly motivated by certain redundancy-theoretic intuitions, it departs radically from that theory in associating a definite content with the truth predicate. We should distinguish between the redundancy theory, according to which 'X is true' says nothing about X, and theories according to which it does say something about X—but something that should be analysed in terms of substitutional quantification. The latter theories face their own set of difficulties, as we saw in the answer to Question 6.

(45)   #,

whose subject, said to be not true, turns out to be the proposition ⟨#⟩. Therefore we have

(46)   ⟨#⟩ is not true,

contradicting our initial assumption. But the alternative fares no better. For if we assume instead that

(47)   ⟨#⟩ is *not* true,

then, given modus tollens applied to the equivalence schema, it follows that

(48)   – #,

which says of ⟨#⟩ that it is *not* not true. Therefore

(49)   ⟨#⟩ *is* true.

Thus we have deduced that

(50)   ⟨#⟩ is true iff ⟨#⟩ is not true.

In order to block the derivation of this contradiction our options are: (1) to deny classical logic—specifically, either modus ponens, modus tollens, double-negation elimination, or Leibniz's law (the indiscernibility of identicals); (2) to deny (à la Tarski) that the concept of truth can be coherently applied to propositions, such as ⟨#⟩, which themselves involve that concept; (3) to deny that the sentence in capital letters succeeds in formulating a proposition; or (4) to reject certain instances of the equivalence schema—including the one obtained by substituting '#' into it.

But (1) cuts too deep; (2) also smacks of overkill; and (3) goes against the fact that, for any condition *C*, one might happen to believe that the proposition meeting that condition is not true—which (since any object of belief is a proposition) would imply that 'The proposition meeting condition *C* is not true' expresses a proposition. And this will be so even if it happens to turn out that the proposition it expresses is the one meeting *C*; so '#' does express a proposition. Therefore the only acceptable solution is (4): only certain instances of the equivalence schema are correct.

We know that this restriction need not be severe. It need have no bearing on the propositions of science—the vast majority of which

do not themselves involve the concept of truth. The problem of giving a *constructive* account of exactly how far one can push the equivalence principle without engendering paradox is the subject of a great deal of contemporary research (e.g. Tarski, 1958; Kripke, 1975; Gupta, 1982; McGee, 1991; Gupta and Belnap, 1993) and will not be addressed in this book. Given our purposes, it suffices for us to concede that certain instances of the equivalence schema are not to be included as axioms of the minimal theory, and to note that the principles governing our selection of excluded instances are, in order of priority: (a) that the minimal theory not engender 'liar-type' contradictions; (b) that the set of excluded instances be as small as possible; and—perhaps just as important as (b)—(c) that there be a constructive specification of the excluded instances that is as simple as possible.[21]

I should emphasize that my intention in these remarks is not to disparage constructive attempts to deal with the paradoxes, or to suggest that our knowledge about truth is not deficient in the absence of such an account. My point is merely that there are manageable and philosophically fruitful problems of truth that are independent of the search for a constructive solution to the paradoxes: first, to *outline* a theory of truth; second, to specify what we mean by the truth predicate; third, to explain its role in our conceptual scheme; and fourth, to say whether there is some theory of the underlying nature of truth. There is no reason to suppose that the minimalist answers that are advanced in this essay could be undermined by any particular constructive solution to the paradoxes—so we can temporarily set those problems aside.

The object of this chapter has been to specify the adequacy conditions for a complete account of truth, to suggest that these desider-

---

[21] Anil Gupta has pointed out to me that the need to restrict instantiation of the equivalence schema is somewhat in tension with the minimalist thesis about the function of our concept of truth—namely that it enables us to capture schematic generalizations. For, in so far as '*p*' is not invariably equivalent to '⟨*p*⟩ is true', then a generalization of the form 'Every instance of schema *S* is true' will not invariably entail every instance of *S*; nor will it always be justified or explained on the basis of those instances. For example, 'Everything he says is true' (i.e. 'Every instance of "If he says that *p*, then *p*" is true') does not entail 'If he says that #, then #', and is not partially justified or explained on the basis of that conditional. However, such problem cases are few and far between; so the utility of truth as a device of generalization is not substantially impaired by their existence.

ata are satisfied by a certain deflationary conception of truth, called 'minimalism', and to make sure that this proposal is not confused with various superficially similar views, such as Tarski's and the redundancy/performative account. The axioms of the minimal theory are all the propositions of the form, $\langle\langle p\rangle$ is true iff $p\rangle$—at least, those that don't fall foul of the 'liar' paradoxes. And (as Anil Gupta has pointed out) there is one further axiom to the effect that propositions are the sole bearers of truth. We found some reason to believe that such a theory—weak as it is—is nevertheless strong enough to account for the conceptual utility of truth, and to explain the facts in which truth is a constituent. And we saw that the single unattractive feature of the theory—its infinite list-like character—is not mitigated by accounts of truth in terms of reference or substitutional quantification. Thus we have gone some way towards justifying the minimalist conception: the view that the minimal theory is *the* theory of truth, to which virtually nothing more should be added.

But many problems remain. For one thing, our entire discussion has taken for granted that truth is a property of *propositions*; and those philosophers suspicious of propositions will find it hard to swallow that aspect of the view. This issue is the focus of Chapter 6. I have placed it towards the end because it is something of a digression, and anyone who is already comfortable with propositions can manage perfectly well without it.

Another widely felt objection to the deflationary view of truth is that it cannot be squared with the *explanatory role* of the notion of truth; and I shall attempt in the next chapter to provide further support for minimalism by showing where this argument goes wrong. The basis for the objection is the idea that any law of nature relating various properties can be explained only by reference to theories that specify the underlying character of the properties involved. For example, in order to say why *all emeralds are green* we need to know what it is to be an emerald and what it is to be green. And similarly, it is argued, in so far as the notion of truth is employed in the formulation of general laws, we are going to need a substantive theory of what truth is in order to explain these laws. I want to suggest, on the contrary, that truth appears in explanatory generalizations in precisely the role identified by the minimalist conception, and that the equivalence axioms are quite sufficient to account for them.

3

The Explanatory Role of the Concept of Truth

11. Truth has certain characteristic effects and causes. For example, true beliefs tend to facilitate the achievement of practical goals. General laws such as this call for explanation in terms of the nature of truth. Therefore there must be some account of what truth is, going beyond the minimalist story, that provides a conceptual or empirical reduction of this property. (Putnam, 1978; Field, 1972, 1986; Devitt and Sterelny, 1989)

As we shall see, truth does indeed enter into explanatory principles, but their validity may be understood from within the minimal theory.

Consider in the first place those of a person's beliefs of the form

(1) ⟨If I perform action A then state of affairs S will be realized⟩.

The psychological role of such beliefs is to motivate the performance of A when S is desired. When this process takes place, and if the belief involved is true, then the desired result will in fact obtain. In other words, if I have belief (1) and desire S, then I will do A. But if my belief is true, then, given merely the equivalence axioms, it follows that if I do A then S will be realized. Therefore, by modus ponens, S *will* be realized; I will get what I wanted. Thus it is easy to see how the truth of beliefs of the kind in question may contribute to the fulfilment of goals. (A formal version of this explanation was given in the answer to Question 3.) Moreover, such beliefs are more likely to be true if they are inferred from true premises; and very little of what we believe can be definitively excluded from the prospect of entering into such inferences as a premise. Therefore it is clear, in

general, how true beliefs contribute to practical success. Nothing beyond the minimal theory is called for to explain this phenomenon.

It is worth noting three features of our argument. In the first place, it does not imply that true beliefs are *always* beneficial. That would be a mistake, since there are obviously circumstances in which a false belief will happen to produce the best outcome and circumstances in which the truth would be too costly to be worth finding out. The argument purports merely to articulate a certain mechanism by which true beliefs engender beneficial results, and does not deny the existence of other mechanisms, which may operate simultaneously, by which a true belief will have bad consequences. In the second place, our explanatory demonstration of the beneficial consequences of true belief is based on facts that are easily and widely recognized. Therefore it may be transformed into an account of why we should want our beliefs to be true—why we aim for the truth. And in the third place, notice that the essential line of explanation is unaffected by the recognition of more complex and realistic patterns of deliberation than those we have been assuming. For example, suppose that we really act according to the principle of utility. In other words, given the choice between actions A and B, we perform the one with the greatest expected value, calculated by means of the formula,

$$(2) \quad V(x) = [V(S_1).B(S_1/x)] + [V(S_2).B(S_2/x)] + \ldots$$

(where $S_1, S_2, \ldots$ is an exhaustive set of mutually exclusive possible outcomes; $V(S_1)$, $V(S_2)$, $\ldots$ are the values that the agent places on them; and $B(S_1/x)$, $B(S_2/x)$, $\ldots$ are his degrees of belief that each outcome will obtain given the performance of action x). To the extent that the degrees of belief are near the truth (i.e. to the extent that $B(S_j/x)$ is high if x would in fact bring about S_j, and low if it wouldn't) then the *expected* value of each action will be close to its *actual* value (i.e. to the value of what would in fact occur if it were performed); and therefore the decision is more likely to be objectively correct.[1] Moreover, as we saw with respect to the simpler

---

[1] To see this, suppose that the actual consequence of act x would be S_1, and consider the possibility of having had degrees of belief, B^*, that were closer to the truth than B are. What this means is that $B^*(S_1/x)$ is closer to 1 than $B(S_1/x)$ is; but for all

model, these beneficial consequences of truth (in the case of beliefs of the specific kind involved in deliberation), indicate that there is value in the truth of any premises from which those beliefs might be inferred.

We may conclude that the explanatory role that the concept of truth plays in the general principle, 'True beliefs facilitate successful behaviour', may be completely understood via the minimal theory. This fact tends against various anti-minimalist positions. First, it should assuage the concerns of philosophers (such as Dummett (1959) and Wright (1988, 1992)) who think that no deflationary conception of truth could do justice to the fact that we *aim* for the truth. Secondly, it shows that the presence of truth in such general principles gives no reason to suppose (with Putnam (1978)) that the property of truth has any sort of underlying structure. And third it undermines the pragmatist's impulse to ensure *by definition* the role of truth in successful activity.

12. Another lawlike generalization is that beliefs obtained as a result of certain methods of inquiry tend to be true. Again this suggests that the minimalist conception overlooks truth's causal/explanatory nature.

We are now turning from the effects of truth to its causes. Beliefs are sometimes reached in such a way as to inspire particular confidence in their truth, and in such cases our confidence is usually vindicated. Consider, for example, observations of the colours of ordinary objects in good light. Reports of such observations are

---

$n \neq 1$, $B^*(S_n/x)$ is closer to 0 than $B(S_n/x)$ is. Thus the expected value of x would have been

$$V^*(x) = V(S_1)(B(S_1/x) + e_2 + e_3 + e_4 + \dots) + V(S_2/x)(B(S_2/x) - e_2) + $$
$$V(S_3/x)(B(S_3/x) - e_3) + \dots$$
$$= V(S_1)B(S_1/x) + (e_1 + e_2 + e_3 + \dots)V(S_1) + V(S_2)B(S_2/x) - e_2V(S_2) + $$
$$V(S_3)B(S_3/x) - e_3V(S_3) + \dots$$

which is *between* $V(S_1)$ (the objective value of x) and $V(x)$ (its expected value relative to the degrees of belief, B). Therefore degrees of belief closer to the truth imply expected values closer to objective ones.

generally correct; and beliefs arrived at by deductive and inductive inference from observational premises are often correct. The question is: why is this so? Why are beliefs regarding certain domains, when resulting from certain methods of inquiry, so strikingly reliable? And won't the answer reveal something important about the nature of truth?

For any observation sentence, 'O', such as 'That's red', 'The needle coincides with the spot marked with a "3"', etc., there are circumstances, $C('O')$, that we take to be particularly conducive to the accurate determination of its truth value, and there are other circumstances in which, though the sentence may nevertheless be asserted or denied, there is thought to be a much higher risk of error. One of the things we are trying to explain is why it is that every instance of

(3)   'O' would be affirmed in $C('O')$ iff% $O*$

is true (where 'p iff% q' means 'The probability of q given p and of p given q are both very high', and where '$O*$' is *our* way of formulating the proposition expressed by 'O').

At the most superficial level the explanation is quite straightforward. It is a biological fact that humans *can be* educated, and a social fact that some of them *are* educated, to say 'That's red' when and only when something red is present, providing the light is good, eyes are open, etc. That is why

(4)   'That's red' would be affirmed in C('That's red') iff% something red is there.

For analogous reasons, though certainly not the *same* reasons,

(5)   'That's green' would be affirmed in C('That's green') iff% something green is there.

And so it will go, for each instance of

(3)   'O' would be affirmed in $C('O')$ iff% $O*$.

Each instance has its own explanation, though some of the instances share some explanatory antecedents. Taken together, these explanations show why *every* instance of (3) is true; or, in other words, why observation reports made in good conditions tend to be true. The minimal theory of truth is perfectly adequate.

Once we have explained why, in certain circumstances, observational beliefs tend to be true, it is not difficult to see why inferred beliefs are also reliable. Truth functions of observational statements (e.g. 'That's red or green') tend to be true because they are reached by deduction from observational beliefs that tend to be true. Beliefs in generalizations of the form

(6)   All As are B

(where 'A' and 'B' are observation terms) are reliable because they are reached on the basis of the observation of many diverse As that are B and of no As that are not B—and it so happens that the world is uniform in this respect. Finally, consider how we might account for the reliability of certain scientific instruments. Suppose a device, I, is designed to discover whether the state of affairs in some domain, S, is $S_1, S_2, \ldots, S_k, \ldots$, or S_n; and suppose this is done by noticing whether the observable output of the instrument is O_1, $O_2, \ldots, O_k, \ldots$, or O_n, and then inferring the presence of the corresponding state—i.e. inferring S_k from O_k. An explanation of the reliability of I might proceed from the following premises:

(7a)   The use of instrument I will give rise, for some k, to the belief that O_k obtains;

(7b)   If we believe that O_k obtains then O_k probably does obtain;

(7c)   There is a high nomological correlation between O_k and S_k;

(7d)   If we believe that O_k obtains, then we infer that S_k obtains.

From these we can infer that instrument I will probably give rise to true beliefs concerning the domain S. Again, nothing beyond the minimal account of truth is needed here.

13. A further explanatory role for truth lies in the fact that the truth of scientific theories accounts for their empirical success. (Putnam, 1978)

No doubt we often do explain the success of a theory by reference to its truth or approximate truth. We say such things as:

(8) The Special Theory enables accurate predictions because it is true,

and

(9) The electron microscope works so well because the theories on which it is based are true.

It remains to be seen, however, whether such explanatory statements provide a reason for thinking that truth has a hidden naturalistic structure, or whether they can be perfectly well accommodated by the equivalence axioms.

Of course, I urge the latter position. Consider the situation in which we know explicitly which theory we are talking about; and suppose its formulation is not very long or complicated. Suppose, for example, that the theory is simply

(10) Nothing goes faster than light.

In that case, rather than saying

(11) The theory that nothing goes faster than light works well because it is true,

we could equally well have said

(12) The theory that nothing goes faster than light works well because nothing goes faster than light.

No further explanatory depth is achieved by putting the matter in terms of truth. None the less, use of the truth predicate in this sort of context will often have a point. What it gives us is a certain economy of expression, and the capacity to make such explanatory claims even when we don't explicitly know what the theory is, or when we wish to generalize, e.g.

(13) True theories yield accurate predictions.

But these are precisely the features of truth that are central to the minimalist conception. Clearly they can provide no reason to go beyond it.

14. Even if all our general beliefs about truth are deducible from the minimal theory (suitably augmented), this does not imply that no deeper analysis of truth is desirable; for one might well hope to find something that will show *why it is* that the equivalence schema holds. (Papineau, 1991)

We can certainly entertain the possibility that the minimal theory is susceptible of explanation via some deeper account of truth. However, there is excellent reason to suppose that *in fact* there is no such deeper theory.

In the first place, the initially tempting strategies for developing such a theory are not workable. In the answer to Question 6 we tried analysing truth in terms of non-standard (sentential) quantification, in terms of reference and satisfaction (à la Tarski), in terms of property exemplification, or in terms of 'actuality', but found it impossible to arrive at an account capable of explaining or improving upon the minimal theory.

In the second place, the equivalence schema is a priori and conceptually basic. In these respects it is analogous to the fundamental laws of logic and arithmetic where there is no expectation of a reductive analysis or any other kind of deeper explanation.

Thirdly, in the domain of a posteriori fact it is reasonable to expect reductive explanations, because the behaviour of a physical system is the causal consequence of the properties of its parts. But such considerations do not apply in the a priori domain. Thus, the minimal theory of truth does not cry out for explanation in the way that a posteriori theories do.[2] Consider, for example, the account of chemical valance (discussed by Field (1972)) which consists in simply listing the valances of each element:

(14) $(x)(y)(x$ is the valence of y iff $x = +1$ and $y =$ potassium, or $x = -2$ and $y =$ sulphur, or . . .).

In this case there *is* a reason to expect further reduction. For there are laws of nature about valence—laws about the relationship between the valences of elements and the proportions in which they combine—that are not explained by the list (14). And any a pos-

---

[2] See Leeds, 1978 for discussion of this point.

teriori lawlike generalization calls for explanation on pain of look-
ing like a sheer coincidence. However the minimal theory does not
itself contain such laws; and it is conceded that *every* general fact
about truth may be explained by the minimal theory. Thus nothing
should lead us either to expect or to desire a deeper explanation.

Fourthly, the equivalence axioms could be explained only by prin-
ciples that are simpler and more unified than they are—principles
concerning propositional elements and the conditions in which truth
emerges from combining them. But the single respect in which the
body of minimal axioms is not already perfectly simple is that there
are so many of them—infinitely many; and no alleged explanation
could improve on this feature. For there are infinitely many funda-
mental propositional constituents to take into account; so any char-
acterization of them will also need infinitely many axioms.
Conceivably, such a theory might none the less obtain unifying
(hence explanatory) credentials by accounting not only for the
equivalence axioms, but also for the phenomena in some other non-
truth-theoretic domain. But this prospect is pure fantasy. We have
no idea of what realm of fact could be related to truth in such a
way.[3]

It has been shown in this chapter that the existence of various sci-
entific, explanatory generalizations, couched in terms of truth, does
not call for an *analysis* of truth—a theory of its underlying struc-
ture. This is because such laws may be wholly understood on the
basis merely of the equivalence axioms, and because any explana-
tion of *these* propositions is highly unlikely. In the next chapter I
turn from the scientific to the philosophical use of the notion of
truth and argue, in a similar vein, that its scope and value are cap-
tured by the minimalist conception.

---

[3] Notice that these considerations relate not merely to *truth*, but also to *reference*
and *satisfaction*. We should expect no deeper analyses of any of these semantic phe-
nomena than are provided by their minimal theories (sketched in the answer to
Question 39).

Methodology and Scientific Realism

A deflationary attitude towards truth is inconsistent with the usual view of it as a deep and vital element of philosophical theory. Consequently the many philosophers who are inclined to give the notion of truth a central role in their reflections on metaphysical, epistemological, and semantic problems must reject the minimalist account of its function. Conversely, those who sympathize with deflationary ideas about truth will not wish to place much theoretical weight on it. They will maintain that philosophy may employ the notion only in its minimalist capacity—that is, as something enabling the formulation of certain generalizations—and that theoretical problems must be resolved without it. The latter point of view is what I will be trying to sustain in the present chapter and in the one immediately following. Here I shall try to show that the realism/anti-realism issue (together with various related questions in the philosophy of science) have nothing at all to do with truth, and that a failure to recognize this fact has stood in the way of clear thinking about those matters. And in the next chapter I shall argue more or less the same point with respect to a range of questions in semantics and in the philosophy of logic.

> **15.** Doesn't the deflationary perspective—the renunciation of a substantive notion of truth—lead inevitably to relativism: to the idea that there is no such thing as objective correctness?

The claim that truth is not a complex or naturalistic property—that it is 'unreal' or 'insubstantial', in the sense advocated by minimalism—must not be confused with the idea that *truths* are unreal, or,

in other words, that no sentence, statement, or belief is ever true. The latter view might arise from an extreme form of relativism in which it is supposed that truth is 'radically perspectival' or 'contextual' or something of the sort. But this kind of theory is not at all affiliated with the minimalist conception of truth. On the contrary, the two philosophical positions tend to be opposed to one another. For it can be precisely the association of the truth predicate with some beefed-up, highly esteemed metaphysical or epistemological property—i.e. the *substantiality* of truth—that leads to the conclusion that nothing ever quite manages to be absolutely true. And conversely, the minimalist position, in so far as it makes it easy to suppose that every proposition, or its denial, is true, implies that relativism (in at least one popular formulation) is untenable. Thus the 'insubstantiality' of truth is in no way tantamount to the nonexistence of truths.[1]

16. Nevertheless, isn't the minimalist perspective in some sense anti-realist? Does it not deny that scientific theories are intended to correspond to a mind-independent world?

Debate over the question 'What is realism?' can easily take on the aspect of an empty, pointless, terminological wrangle. One philosopher will identify the position with, say, an aversion to reductionism; another will complain that certain pro-reductionist positions (e.g. that the mind is merely the brain (materialism), that mathematical facts are merely logical facts (logicism)) are not intuitively anti-realistic and that certain positions that *are* anti-realistic (e.g. that the numbers are a human invention (mathematical intuitionism)) are non-reductionist; the critic may then propose an alternative definition of 'anti-realism'—say, rejection of the principle that every proposition is either true or false; but this again will fail to satisfy everyone's intuitions about when to apply the label. Thus the process continues interminably, so that one is left wondering—since after all

---

[1] An argument that the minimalist conception of truth leads to relativism has been given by Putnam (1981). However, objections that I consider to be conclusive have been levelled against it by Michael Williams (1986).

the words 'realism' and 'anti-realism' are terms of art—whether there is any genuine problem here at all. Wouldn't it be best to distinguish explicitly various senses of 'realism' corresponding to each of the alternative proposals, so we can begin to focus our attention on the real questions: namely, which of these so-called 'realist' positions are correct and which are not?

I think that this tempting point of view is in fact mistaken. Realism and anti-realism are definite and interesting philosophical stances, and the issue of what exactly they are is a substantial one. Our impression to the contrary comes partly from the fact that the answers that are usually suggested bear such little resemblance to one another, and partly from the fact that even with respect to individual philosophical positions (for example, behaviourism and intuitionism) there is no consensus about whether they should be counted as realist or not. But the explanation of all this divergence of opinion, it seems to me, is not that there is no correct definition of realism to be found, but rather that the definitions usually proposed are of completely the wrong sort.

What I have in mind can be brought out by reviewing some well-known facts about how natural kinds—for example, *diseases*—are properly defined. Notorious difficulties arise if one tries to characterize a disease in terms of observable symptoms, for there will always be some that are not manifest in a few people who none the less have the disease, and other characteristic conditions that sometimes occur in the absense of the disease. The familiar moral is that the right way to specify the criterion for having a disease is to identify the underlying causes of its symptoms, rather than the symptoms themselves. It seems to me that this moral applies pretty well to our questions about the nature of the realism/anti-realism issue. What accounts for the endless squabbling about it is that we have been focusing our attention on the symptoms of realism and anti-realism. Not surprisingly, no definition in those terms can work. What we must do instead is think about what it is that leads people to adopt the positions we are inclined to regard as realist or anti-realist. Thus we will see what the basic conflict between realism and anti-realism really is. And the positions that we intuitively classify as realist or anti-realist—e.g. that physical objects are constructs from experience (phenomenalism), that electrons are fictional entities (instrumentalism), etc.—will qualify as such in virtue of being adopted as a consequence of taking one side or the other on

the basic issue. Notice that this analysis of the situation will imme-
diately account for the divergence of opinion over whether certain
specific theories, such as behaviourism and intuitionism, are anti-
realist or not. As in the case of a disease, where a given condition
may in some cases be a symptom and in another case not, a given
philosophical thesis will be an anti-realist move in the case of those
philosophers who embrace it for certain reasons, but not for those
whose motives are quite different. Thus we might distinguish, say,
between a '*realist* behaviourist' and an '*anti-realist* behaviourist':
both have exactly the same view about the reducibility of mental
facts to behaviour, but they diverge in their reasons for holding it.

What, then, *is* the essence of realism? The answer is very simple.
There is a question about how it is possible for us to know of the
existence of certain facts given our ordinary conception of their
nature. This is because there can seem to be a tension in ordinary
thinking between the metaphysical autonomy of the world (its inde-
pendence of us) and its epistemological accessibility (our capacity
to find out about it). The difference between a realist and an anti-
realist, in a nutshell, is that the realist decides on reflection that there
is actually no difficulty here—so our ordinary ideas about what we
know can stand; whereas the anti-realist decides, on the contrary,
that the alleged conflict is genuine and that it has certain
ramifications for what we can take ourselves to know. The alterna-
tive symptoms of anti-realism are alternative views about what these
ramifications are—alternative modifications of our naive view of the
world and our capacity to comprehend it. Thus it is not unusual for
an anti-realist with respect to some domain to embrace one of the
following strategies:

(1) Deny that there are any facts of the sort at issue (e.g.
    formalism, instrumentalism, emotivism, relativism, non-
    factualism);
(2) Deny that we have the capacity to know such facts (e.g.
    scepticism, constructive empiricism, error theory);
(3) Reduce the facts in question to other facts whose epistemo-
    logical status seems unproblematic (e.g. phenomenalism,
    behaviourism, logicism).

The central point, once again, is that none of these doctrines, nor
any collection of them, is either necessary or sufficient for anti-
realism. Rather, anti-realism is the view that our common-sense

conception of what we know is incoherent: the supposed character of facts of a certain type cannot be reconciled with our capacity to discover them. It is in response to this view that one or another of the above doctrines may well be espoused and in such case the adoption of the doctrine qualifies as an anti-realist move.

According to this way of thinking, *any* position whatsoever might count as an anti-realist move for *some* philosopher, providing that this philosopher regards the position as the proper solution to the anti-realist dilemma. Thus the class of 'possible anti-realist positions' (like the class of 'possible symptoms of diabetes') is completely uninteresting. What is of interest, however, is something we might call the class of '*natural* anti-realist positions': that is, positions which really would remove the alleged tension in our naive world view. In other words, it is worth distinguishing, from amongst the more or less irrelevant things an anti-realist might be inclined to say, those positions that really would, if they could be adequately sustained, address the alleged dilemma.

The three strategies just mentioned clearly have this character: they are natural anti-realist positions. However, what is glaringly absent from this group is any particular thesis about the nature of *truth*. This is not of course to deny that someone might, as a matter of biographical fact, feel forced into some account of truth by the anti-realist dilemma. My claim, rather, is that any such response would be irrational. If the dilemma is real, then no theory of truth could help to resolve it.

> **17.** But this conclusion is extremely counterintuitive. It seems obvious that the nature of truth bears directly on the structure of reality and the conditions for comprehending it. Surely, 'truth' and 'reality' are semantically inextricable from one another; so how could one's position in the realism debate be divorced from one's conception of truth?

The term 'realism' is an over-used, under-constrained piece of philosophical jargon, and one can no doubt invent senses of it such that the minimalist approach qualifies either as 'realist' or 'anti-realist'. However, the substantial question here, as we have just seen, con-

cerns the relation, if any, between our conception of truth and the justifiability of believing in facts that exist independently of thought or experience. And there is no relation—or so I shall argue. On the contrary, a significant source of confusion in the debates about scientific realism is the tendency to assume that the problem of truth is fundamentally involved.[2]

Anti-realism, as we have seen, consists in a perceived tension between the realist's twin, common-sense commitments to credibility and autonomy. Some anti-realist philosophers have supposed that since the facts are independent of experience they are non-existent, or at least epistemologically inaccessible to us. Thus we arrive at the sort of instrumentalism or theoretical scepticism advocated, for example, by Duhem (1954), Popper (1962), and van Fraassen (1980). Other philosophers have supposed, conversely, that since the facts are verifiable they must reduce to observation. Thus we get the sort of reductive empiricism characteristic of phenomenalism and the Vienna Circle. But none of these natural anti-realist positions is in any way affiliated with the minimalist conception of truth. In the first place, minimalism gives no reason to think that theories are constructions out of data, and is quite at home with the holistic considerations that have led most philosophers to reject that aspect of logical positivism. Secondly, it is perfectly consistent with the minimal account of truth to suppose that the scientific method provides us with theories that we should believe to be *true* and not merely observationally adequate. According to the deflationary picture, believing that a theory is true is a trivial step beyond believing the theory; and the justifiability of this attitude is certainly not precluded by minimalism.[3]

Not only is the minimalist conception of truth quite neutral with respect to the two central aspects of realism (namely, the questions of justified belief and empirical reducibility), but the same can be said of alternative conceptions. As we shall now see, the *choice* of a theory of truth is orthogonal to the issues surrounding realism. The

---

[2] The independence of questions about truth from the traditional issues of realism was urged by Tarski (1943/4), and has been emphasized by Michael Devitt (1984). For further discussion see Horwich, 1996.

[3] In support of the anti-sceptical component of realism I have argued elsewhere (1991) that there is really no difference between *believing* a total theory and the apparently less-committed attitude of *instrumental acceptance*.

theory of truth can have no definite implications for either the epis-
temological or the semantic component of the problem.

Consider, for example, the constructivist account, which identifies
truth with a kind of demonstrability or verifiability. There is an incli-
nation to suppose that this conception of truth immediately entails
the falsity of certain forms of scepticism, and that it thereby sup-
ports the epistemological aspect of realism. At the same time it is
also thought that the meanings of sentences would be given by their
'truth conditions' (in the constructivist sense); and so it seems that
the content of a claim such as

(4)   There are electrons

could then be nothing more than

(5)   'There are electrons' is demonstrable.

Thus the autonomy of theoretical facts would be lost, and we would
have semantic anti-realism.

But both of these arguments are fallacious. In order to combat
scepticism regarding some theory, T, we must be able to argue (in
the face of the underdetermination of theory by all possible data)
that we are justified in believing T. Constructivism (according to
Peirce (1932/3)) tells us that we can infer

(6) T is true

from the premise

(7) T will eventually be demonstrated (i.e. verified) in the limit
      of scientific investigation.

In Putnam's (1981) version of the doctrine, the premise should be

(8) T would be demonstrated in the course of an ideal inquiry.

However, it is not at all obvious, from the sceptic's viewpoint, that
we are entitled to believe either of these premises. Moreover, even if
that source of scepticism were removed, it would still be unclear why
a justified belief *that T is true* should carry with it the justification
to believe T. No doubt a *minimalist* can assume this; for he regards
the equivalence of such beliefs is as fundamental. But a construc-
tivist, on the contrary, defines truth in terms of demonstrability. For
him, the equivalence schema is a substantive claim that must be sup-
ported on the basis of this fundamental assumption about truth.

And the possibility of such an argument is precisely what the sceptic will deny. Thus constructivism gives no easy proof of the ordinary claims to knowledge constitutive of scientific realism.

Nor does it amount to a form of anti-realism, as many writers, following Dummett (1977), have assumed. There is a tendency to confuse the following three theses:

(9)   The meaning of '*p*' consists in the fact that '*p*' is regarded as demonstrated in such and such circumstances;

(10)   ' "*p*" is true' means ' "*p*" is demonstrable';

and

(11)   '*p*' means ' "*p*" is demonstrable'.

The final thesis evidently conflicts with realism; for it reduces facts about external reality to facts about our thought and experience. However, constructivism gives us the right to nothing more than (10), and arguably (9). And these premises provide no basis for denying that scientific theories describe a mind-independent reality.[4]

A second account of truth with merely *apparent* implications for realism is the view that truth is a primitive, non-epistemic property that is grasped independently of the equivalence schema. This view seems to wear on its face the radical autonomy of theoretical facts. And, as a consequence, scepticism can appear to be unavoidable. For if the property of truth is primitive and wholly unexplainable, then we can surely have no reason to suppose that the propositions we regard as confirmed tend to have this property. But, once again, these attempts to link the theory of truth with realist and anti-realist theses are misconceived. For the meaning of the word 'true' is one thing, and the meanings of theoretical terms like 'electron' and 'super-ego' are quite separate. Whatever we say about 'true' cannot determine our view on the question of whether our theoretical terminology is reducible to observational terminology. Similarly, however mysterious and inaccessible we think the property of *truth*—however hard we suppose it is to assess '*T is true*'—we will not necessarily be saddled with scepticism; for we need not also

---

[4] Putnam (1983: 280) appears to go wrong in this way when he argues that anyone who adopts the combination of a redundancy theory of truth and an assertibility condition conception of meaning will is 'perilously close to being a solipsist of the present instant'.

suppose that the property of (say) *being an electron* is mysterious and inaccessible; there need be no scepticism about T itself. As before, the essential point is that any theory that explains truth *independently* of the equivalence schema, loses the right to assume without further ado that the schema holds. Therefore, relative to the conception of truth in question, problems regarding the justification of 'T *is true*' are not automatically linked with problems regarding the justification of T.

18. If, as the minimal theory implies, 'truth' is not *defined* as the product of ideal inquiry, why should we believe that an ideal inquiry would provide the truth?

To regard a certain inquiry as *ideal* is to suppose that one should not question its outcome. So if an inquiry into whether or not there is life on Mars yields the result that there isn't, and if the inquiry is taken to be ideal, then we should be absolutely confident that there is no life on Mars. Moreover, given the equivalence schema, one should also be confident in the *truth* of the proposition that there is no life on Mars. Similarly, one should believe of *any* ideal inquiry that it provides the truth.[5] This is a trivial consequence of the minimal theory and the meaning of 'ideal inquiry'. Therefore constructivism is unmotivated; for what it feels the need to guarantee by *definition* may in fact be derived from the minimal theory.

Notice that there is no presumption here that *every* hypothesis is susceptible to some idealized inquiry. Therefore, although we have grounds for the schematic thesis

(12)   If $\langle p \rangle$ is the product of an ideal inquiry, then $\langle p \rangle$ is true,

the converse claim—and, a fortiori, the identification of truth with idealized verification—has not been supported. Indeed, this identification—the constructivist theory of truth—greatly overesti-

---

[5] In the terminology of Putnam (1981), we are here rebutting the 'metaphysical realist' thesis that 'truth is radically non-epistemic'. Fine; but this does not mean that we agree to incorporate epistemic ideas into the very notion of truth. Peter van Inwagen (1988) points out in a similar vein that Putnam's argument that 'a fair amount of what we believe must be true' also does not imply that truth is an epistemic concept.

mates our epistemological power. For there are truths beyond the reach of even an ideal investigation. Consider the phenomenon of vagueness (e.g. 'is bald' applied to a borderline case); or underdetermination of theory by data; or sentences with assertibility conditions that don't allow for conclusive verification (e.g. 'The probability that drug X will cure disease Y is 0.3'). Any of these phenomena might involve a proposition which is such that no ideal investigation would engender either its assertion or its denial. In that case we have

(13)   It is not the case that $\langle p \rangle$ is verifiable and it is not the case that $\langle -p \rangle$ is verifiable.

But then, according to the constructivist's definition of truth, we can infer

(14)   It is not the case that $\langle p \rangle$ is true and it is not the case that $\langle -p \rangle$ is true.

And by the equivalence principle,

(15) $p \leftrightarrow \langle p \rangle$ is true,

we get

(16) $-p$ and $--p$,

a contradiction! Thus not only is constructivism unmotivated (i.e. not needed to account for whatever correlation exists between truth and verification), it is extensionally false since it cannot acknowledge the existence of truths that are not conclusively demonstrable. One might try to argue, in response, that given a sufficiently pumped-up construal of 'ideal inquiry', there is really no need to acknowledge unverifiable truths. For it could be supposed that for every proposition, including the problematic cases just mentioned, a *sufficiently* ideal investigation would decide its truth. But although the constructivist principle may be protected against counterexample by this manœuvre, it will then be even less appropriate than before to regard it as our basic theory of truth. For the notion of 'sufficiently ideal inquiry' will now be one that is most naturally explicated in terms of the concept of truth. Thus the constructivist principle, once it is cast into a plausible form, clearly becomes something to be derived

from the minimal theory of truth, and not to be taken as explanatorily basic.

19. How is it possible, given the minimal theory, for truth to be something of intrinsic value, desirable independently of its practical utility?

To value truth is, very roughly speaking, to wish for satisfaction of the schema 'p iff I believe p', and therefore to be committed to the techniques of investigation that will apparently achieve this result.[6] To value truth for its own sake is to desire it to some extent regardless of its compatibility with other goals. Such a value may be ethical or aesthetic, or it may be something one simply wants. Which of these alternative possibilities is correct is left open by the minimal theory, as is the possibility that truth *not* be valued for its own sake.

It might be thought that if truth *is* intrinsically valuable, then minimalism is in trouble, since it surely lacks the resources to explain that value.[7] But this criticism is unjust. For the difficulty that attaches to explaining why true belief is intrinsically good is no more or less than the difficulty of explaining, for any other particular thing (e.g. kindness, happiness, etc.), why it is intrinsically good. The problem stems from our failure to understand the concept of intrinsic goodness, rather than from our adoption of the minimalist conception of truth. More specifically, in so far as we don't know what it is for something to possess the quality of intrinsic goodness, no explanation of why truth possesses it will suggest itself, regardless of which theory of truth is adopted. On the other hand, if some

---

[6] This way of articulating what it is to value truth involves a couple of oversimplifications. First, it suggests that we wish to believe every true proposition; whereas, in fact, the truth values of most of them are of no interest to us. And second, it does not allow for *degrees* of belief. Really, what we want is to 'minimize', in a certain sense, the error in our plausibility judgements. If these are represented by numbers between 0 and 1, then what we want is to minimize the square difference between the probability assigned to each proposition $\langle p \rangle$, and either 1 (if $\langle p \rangle$ is true) or 0 (if $\langle p \rangle$ is false). See Horwich, 1982*b* for discussion of this matter, together with an argument that these desires are accommodated by the acquisition of new data.

[7] Bernard Williams (1996) expresses scepticism about minimalism on these grounds.

account of intrinsic goodness is assumed, then it is far from obvious that a minimalistically acceptable explanation of why truth has that quality could not be based upon the assumed account. Consider, for example, an analysis of 'being intrinsically good' roughly along the lines of 'being normally conducive to human welfare'. On the basis of this sort of theory the minimalistic account of truth's pragmatic value (given in response to Question 11) might well be developed into an account of its intrinsic value. Note that such an account would not undermine the *intrinsicality* of the value. For to recognize that truth is valuable 'for its own sake'—i.e. to suppose that truth is good even in those cases where it will have no practical benefits—is not to suppose that the explanation of this value is independent of the *normal* advantages of truth.

20. How can minimalism accommodate the idea of science progressing towards the truth?

It suffices to imagine a temporal sequence of total theories $T(1)$, $T(2), \ldots, T(k), \ldots, T$(final), becoming gradually (but not necessarily monotonically) more similar to T(final)—where T(final) is true. T(final) is a conjunction of unknown and presently inexpressible propositions. However, as we saw in the answer to Question 2, this is no obstacle to applying the truth predicate. The notion of 'theoretical similarity' remains to be explained; but there is no reason to expect that this can or should be done with a high degree of precision. We can get by with our ordinary crude intuitions of the extent to which two bodies of claims are similar to one another. Of course, in order to make these comparisons it is necessary that the theory formulations be to a fair degree intertranslatable (or 'commensurable' in Kuhn's terminology (Kuhn, 1962)). I have not tried to show that this would be so. However, I am not arguing here that there actually exists progress in science; but only that a minimalist conception of truth does not stand in the way of such a thesis.[8]

---

[8] Even if we give up the idea that there exists a 'final true theory' we could still make sense of a weaker version of the view that science progresses with respect to truth. We might suppose (a) that a proposition, $\langle p \rangle$, is *roughly true* just in case the proposition $\langle$Roughly, $p\rangle$ is true (for example, it is roughly true that John is six feet tall when it is true that John is roughly six feet tall); (b) that later members of the

21. From the perspective of the minimalist conception of truth it is impossible to produce an adequate justification
of scientific methods. (Friedman, 1979)

An argument that our ways of acquiring beliefs take us in the direction of *truth* might proceed on the basis of assumptions that are themselves the results of methods we are trying to justify. That is to say, an explanation of the fact that a certain scientific method M is reliable might proceed from premises of which some are believed as a consequence of employing M itself. According to Michael Friedman, this sort of derivation will sometimes constitute a *justification* of the method M. He argues that—given a sufficiently substantive conception of truth—the *circularity* that is evidently involved need not be vicious. It need not render the derivation so easy to provide that it isn't worth having. After all, he says, there can be no prior guarantee that the products of M will not suggest a theory that implies M's *un*reliability. Therefore, it can be a pleasantly surprising and epistemologically significant fact regarding M if it turns out to be demonstrably reliable—even if the demonstration employs the results of M itself.

Friedman's objection to deflationism is that, from the perspective of such an *in*substantive conception of truth, it would be a *trivial* matter (available regardless of what one's beliefs and methods actually are) to produce demonstrations of reliability. As a consequence, these demonstrations would be quite devoid of explanatory or epistemological importance, and we would be left with no grounds for confidence in the reliability of scientific inquiry.

He reasons as follows. Suppose

(17)   Method M has engendered beliefs
$p_1, p_2, \ldots,$ and p_n.

In order to derive the reliability of M it suffices to combine premise (17) and the additional premises, $p_1, p_2, \ldots,$ and p_n, that were obtained from M. From these additional premises (and the equivalence schema) we infer that

possibly endless sequence of total theories tend to contain a greater number of roughly true, basic theoretical claims than earlier members; and (c) that most members of the sequence are such that the earlier members tend to be increasingly theoretically similar to it.

(18) 'p_1' is true,

   'p_2' is true,

   . . .,

   and 'p_n' is true.

And then, from premise (17), we get the result that M engenders true beliefs.

This pseudo-explanation is indeed worthless and provides no support whatsoever for method M. But before casting aspersions on minimalism we should consider a series of further questions:

   (19) Does minimalism play any role in the above pseudo-explanation?

   (20) Could there not be—consistent with minimalism—a more substantive demonstration of M's reliability?

   (21) Even from the perspective of a *non*-minimalist theory of truth, could there be a telling demonstration of M's reliability?

If the answers to these questions were, respectively, Yes, No, and Yes, then we might indeed have to concede that minimalism has unwelcome epistemological consequences. But in fact I believe that the answer to all three questions is No. In the first place, the above pseudo-justification relies merely on the equivalences that are common to all reasonable accounts of truth. Minimalism—the thesis that such biconditionals exhaust the theory of truth—plays no role and cannot be blamed. Secondly, the reliability of certain methods way well be demonstrable on the assumption of facts discovered by other methods—but *their* reliability would then be at issue. Eventually the question would arise as to whether some method (or collection of methods) is capable of justifying itself in the manner that Friedman proposes. And it seems to me that, regardless of one's theory of truth, no such justification can be given.

Here I am questioning whether it really is possible for science to undermine itself in a thoroughgoing way. I would suggest that the circular procedure envisaged by Friedman could not go *badly* wrong; and that this has nothing to do with the account of truth that is employed. I am not denying that the theories resulting from method M might fail to provide an explanation of M's reliability.

For perhaps further theories are needed; or perhaps the reliability of M is extremely hard to explain. What I question is that the theories resulting from M might imply that M is not reliable. Consequently, since I agree that something is supported by its successful predictions only to the extent that they *might* have been mistaken, I also question whether the success of Friedman's circular *explanatory* procedure could constitute any sort of *justification* for relying on M.[9]

In order to motivate this scepticism, let me consider a couple of examples. Suppose that, on the basis of observational beliefs O_1, $O_2, \ldots$, and O_n, we were to postulate a theory T that entails that our methods of observation were very unreliable. In that case we would have a theory T that, on the one hand, is confirmed by the fact that it entails $O_1, O_2, \ldots$, and O_n; yet, on the other hand, entails that a high proportion of $O_1, O_2, \ldots$, and O_n are false. Therefore T would have to be internally inconsistent, and we shouldn't have postulated it in the first place. Similarly, suppose that our theory T^*, justified by induction, entails that inductive inference is unreliable. In other words:

(22) $T^* \to$ Although data have conformed to T^* in the past, they will not conform to T^* in the future.

Again, such a theory is internally inconsistent and should not have been taken seriously in the first place.

Thus there are limits to the extent that science, rationally pursued, can invalidate itself. But this has nothing to do with the nature of truth and, in particular, is not a consequence of minimalism. Friedman suggests, on the contrary, that a naturalistic (causal) theory of truth would leave more room for scientific self-criticism and self-validation, for he supposes that it would open up the possibility that most of the beliefs we regard as verified will turn out not to have the naturalistic property of truth. But this is an illusion. If there were such a property—a naturalistic reduction of truth—it

---

[9] Perhaps Friedman's concern is merely with the *explanation* of M's success, and not with providing reasons for confidence in it. But this interpretation is hard to reconcile with various facts: (1) that no reason is given for thinking that a non-trivial, minimalistically acceptable *explanation* of M's reliability cannot be found; (2) that it is thought necessary to maintain that M might undermine itself; and (3) that it is thought necessary to have a conception of truth according to which it is conceivable that most of our beliefs are false.

could be recognized as such only by means of the assumption that it *does* tend to be present in circumstances that we regard as instances of verification. If we don't impose this constraint then we violate not merely the *minimal* theory of truth, but any theory that respects the equivalence of '*p*' and '⟨*p*⟩ is true'. Thus no remotely plausible account of truth could make it conceivable that our beliefs are predominantly false and our methods of arriving at them unreliable. The limited applicability of Friedman's epistemological strategy will not be expanded by rejecting the minimal theory.[10]

Having attempted in earlier chapters to make clear and plausible the minimalist point of view, I have here begun to explore its philosophical implications, specifically with regard to the debate over realism. Not surprisingly, what I have tried to show is that our notion of truth does not occupy the central theoretical position that philosophers often assume it must occupy. Indeed, many problems are exacerbated by the conviction that truth is essentially involved and that their solutions depend on finding out more about its underlying nature. We have seen on the contrary that in so far as the notion of truth is properly employed in the philosophy of science it displays no more than its minimalistic function. And in the next chapter I argue the same point with respect to a broad range of semantic questions. I shall take up (a) the nature of understanding, (b) the basis of logic, (c) empty names, (d) vagueness, and (e) the status of ethical assertion; and I shall indicate in each case how the problems becomes much simpler once it is acknowledged that the concept of truth should not be relied on to solve them.

[10] For some further criticisms of Friedman's line of thought see Williams, 1986.

5

Meaning and Logic

> **22.** As Davidson (1967) has argued, understanding a sentence, say, 'Tachyons can travel back in time', is a matter of appreciating what must be the case for the sentence to be true—knowing its *truth condition*. That is to say, one must be aware that 'Tachyons can travel back in time' is true iff tachyons can travel back in time. Therefore it is not possible to agree with the minimalist claim that this knowledge also helps to constitute our grasp of 'is true'. For in that case we would be faced with something like a single equation and two unknowns. Rather, if knowledge of the truth condition of 'Tachyons can travel back in time' is to constitute our understanding of that sentence then this knowledge would presuppose some pre-existing conception of truth.
> (Dummett, 1959; Davidson, 1984)

What is right in this point is that knowledge of the truth condition of a sentence cannot simultaneously constitute *both* our knowledge of its meaning *and* our grasp of truth for the sentence. What is wrong about it is its choice of the first of these options. For, on such a view, how might we come to know that

(M)   'Tachyons can travel back in time' is true iff tachyons can travel back in time?

How could this feat be accomplished? The picture that comes to mind is that we deliberately associate the sentence 'Tachyons can travel back in time' with a possible state of affairs, where the form of association is our decision to count the sentence 'Tachyons can travel back in time' as true if and only if the state of affairs obtains. But there are considerable difficulties in this position. It is surely impos-

sible for an individual to conceive of such an explicit association unless he employs some sort of mental event—call it '*R*'—to represent the possible state of affairs. And in that case two problems emerge. In the first place, it is more straightforward to represent the association of our sentence with the possible state of affairs by means of the definition

(M*) 'Tachyons can travel back in time' is to have the same meaning as *R*

rather than by means of (M); so the notion of truth need not be involved. And, in the second place, we must raise the question of what provides representation *R* with *its* meaning. In order to avoid an infinite regress it must be conceded that certain representational entities obtain their content by means other than having been explicitly associated with possible states of affairs. Yet the truth-conditional approach provides no place for such alternative means. Moreover, if some cases of meaning do not arise from explicit associations with truth conditions, then why should we assume that understanding 'Tachyons can travel back in time' must have that character?

The way to avoid this mess is to recognize that while understanding a sentence does indeed usually *coincide* with an explicit knowledge of its truth condition, understanding does not *consist* in such knowledge. It consists, rather, in appreciating the sentence's syntactic structure and understanding its constituent words, which, in turn, consists in knowing the basic regularities in their use—those regularities that will explain the overall use (including the assertibility conditions) of all the sentences in which those words occur. Once 'Tachyons can travel back in time' is understood in this way by someone with a conception of truth, then the minimalist account entails that he knows that 'Tachyons can travel back in time' is true iff tachyons can travel back in time—i.e. it entails that he knows the truth condition of 'Tachyons can travel back in time'. Moreover, such knowledge can usually be attributed only to those who understand 'Tachyons can travel back in time'.[1] Thus anyone with a conception of truth who understands 'Tachyons can travel back in time' will

---

[1] But not always. Someone who does not understand German and who is told that 'Schnee ist weiß' is true iff frozen H_2O is white, does not understand the German sentence, even though he knows its truth condition, because he does not know whether 'Schnee' means 'snow' or 'frozen H_2O'.

indeed come to know its truth condition. However, contrary to what is assumed in the objection, the understanding does not derive from this knowledge.[2]

Let me stress some aspects of this position. First, the advantage of relying on *use* in giving a naturalistic characterization of understanding is that it is obvious how knowledge of the use of a word is manifested (namely, by accepting certain sentences containing it in certain conditions), whereas it is not at all clear how knowledge of truth conditions is manifested; that is, unless such knowledge is construed, in the way that I have suggested, as the *product* of a knowledge of meaning (which is in turn explained in terms of use) and a grasp on the concept of truth. Secondly, it is no objection to either the coherence, or the preferability, of the notion of *the use* of a word that it might sometimes include accepting certain sentences when and only when their *truth* conditions are satisfied. For example, it might be that the use of 'blue' involves accepting 'That is blue' when and only when the designated object is blue. This does not alter the fact that we can see how knowledge of the word's use may be manifested, whereas we cannot see how knowledge of its contribution to the truth condition may be manifested. Thirdly, an analysis of *meaning* (and *truth condition*) in terms of *use* does not imply that a sentence cannot be true without being assertible. In the first place, we are not simply identifying the meaning of a sentence with its assertibility conditions. And in the second place, even if we did, the assertibility of a sentence would not follow from its truth.[3] Finally, it is sometimes claimed to be a special advantage of the 'truth-conditional' analysis of meaning that it enables us to see how the meanings of composite expressions depend on the meanings of their parts, and to see therefore how it is possible for us, with our finite minds, to understand a potential infinity of compound expressions. But this alleged advantage is an illusion; for the compositionality of meaning can equally well be accommodated within the use conception of meaning. In so far as the fact that a word has the meaning it has consists in a certain fact about its use, then the meaning of a complex expression will consist in the fact that it has a certain structure together with the facts

---

[2] See Harman, 1974, 1982, for a statement of this position. There is more on the use theory of meaning in the answer to Question 32. See also my *Meaning* (1998).

[3] For an elaboration of this point, see the discussion of scientific realism in sections 17 and 18.

underlying the meanings of its constituents. The main idea behind this approach to compositionality is that understanding a complex expression is, by definition, nothing over and above understanding its constituents and appreciating how they have been combined with one another. If this is so, then compositionality can put no constraint on how the meanings of words are constituted. For whatever are the underlying characteristics that do this, complexes will mean what they do in virtue of being composed as they are from words with specific characteristics of that sort. In particular, if a word's meaning derives from its use, then a complex's meaning consists in its being the result of combining, in a certain way, words with certain uses.[4]

23. What about falsity and negation?

The simplest deflationary strategy is to define falsity as the absence of truth, as follows:

(1) $(x)(x$ is false $\leftrightarrow x$ is a proposition & x is not true),

or, in other words,

(2) $\langle p \rangle$ is false $\leftrightarrow \langle p \rangle$ is not true.

As for the word 'not', it is traditionally supposed that the best way to explain both it and the other logical constants is by means of truth tables. In the case of negation, this is generally taken to be

(3)

p	$\underline{\text{not } p}$
T	F
F	T

However, in the present context such an account is unacceptably circular—we have defined falsity in terms of negation, and would now be defining negation in terms of falsity. Moreover, the logician's term 'not' (underlined), which is supposedly defined by the truth table, does not mean the English 'not', but rather 'It is not the case that', or in other words 'It is not true that'—which isn't what we needed to define.

---

[4] For elaboration of this view of compositionality see my 1997*a*, reprinted as ch. 7 of my *Meaning* (1998).

In order to deal with the second of these problems, and to bring it about that an account of the logician's 'not' will engage with a minimalistic account of falsity, we must adopt a slight modification of that account—we must replace (2) with the definition

(2*) $\langle p \rangle$ is false ↔ not [$\langle p \rangle$ is true],

or equivalently

(2**) $\langle p \rangle$ is false ↔ not p.

And in order to deal with the first problem, and remove the circularity involved in combining (2*) with the truth table definition of 'not', we must eliminate the notions of truth and falsity which occur in the truth table by explicating them in accordance with the minimalist proposals. In that case the lines of the truth table, which initially say

(N)   (i) $\langle p \rangle$ is true → $\langle$ not $p \rangle$ is false
         (ii) $\langle p \rangle$ is false → $\langle$ not $p \rangle$ is true,

and the implicit claim that one of the two lines applies, i.e.

         (iii) $\langle p \rangle$ is true or $\langle p \rangle$ is false,

are transformed into the following theses:

(N*)   (i*) $p \to$ not not p
         (ii*) not $p \to$ not p
         (iii*) p or not p.

In taking 'not' to be implicitly defined by (N*), rather than by (3) or (N), we are no longer attempting to explain it in terms of falsity; so the threat of circularity has been defused. However these three theses, though to some extent constraining the meaning of 'not', are not enough to fix it completely. A complete account of the meaning of 'not' must contain those fundamental facts about its use that suffice to explain our entire employment of the term. Such basic regularities of use might well include acceptance of the theorems of deductive logic—which include the laws implicit in (N*). But a further pattern of usage, not implied by (N*), must be recognized: namely, that which is characterized by the principle

(K)   'not p' is acceptable to the degree that 'p' is unacceptable.

Perhaps the combination of (N*) and (K), when conjoined with facts about the use of other terms, will be capable of explaining all our

ways of deploying '<u>not</u>'. If so, then its meaning will be fixed and we can proceed to define falsity in terms of it by means of definition (2*) and without fear of circularity.

24. As Frege (1918) said, logic is the science of truth; so surely our accounts of truth and logic should be, if not identical, at least bound up with one another. Yet the minimal theory does not even enable one to prove that the principle of non-contradiction is true.

The concept of truth is involved in stating laws of logic and meta-logic; for example,

(4) Every proposition of the form '$p \rightarrow p$' is true.
(5) If a conditional and its antecedent are true, then so is its consequent.
(6) The laws of logical inference preserve truth.

Thus one easily gains the impression that truth and logic bear a peculiarly intimate relationship to one another. Indeed Frege maintained that, just as biology is the science of living things and astronomy the science of stars, so logic is the science of truth. And it is not uncommon (e.g. Dummett, 1977; Putnam, 1978) to see the competition between alternative logics described as a choice between conceptions of truth.

From the perspective of minimalism, this way of thinking is incorrect. The reason that the notion of truth is heavily involved in logic is not that logic is *about* truth, but simply that logic makes precisely the sort of generalization that the truth predicate enables us to formulate.

Moreover, just as we should, if possible, prise apart our theories of truth and reference, so, for the same reason, we should distinguish logic from the theory of truth. One and the same theory of truth can be combined with classical logic to demonstrate, for example, that the distributive law is true, or combined with quantum logic to show that it isn't. And just as our belief that the Principle of Relativity is true requires for its derivation assumptions from physics as well as from the theory of truth, similarly it should be quite unsurprising that in order to prove the truth of the principle of non-contradiction we need to invoke logic, and not simply the theory of truth. This means

that the issue between, for example, classical and intuitionistic logic has as little to do with truth as has the issue between Newtonian and Einsteinian physics. It would be absurd to describe the latter as a conflict between Newtonian and Einsteinian conceptions of truth; and for the same reason it is wrong to think of the logical dispute as a competition between 'the classical conception of truth' and 'the intuitionistic conception of truth'.

Although I use classical logic throughout this book, none of my claims about truth presupposes it. Indeed, as I have just emphasized, a central tenet of the point of view advanced here is that the theory of truth and the theory of logic have nothing to do with one another. Thus minimalism is the proper conception of truth even in the context of deviant logics such as intuitionism or quantum logic, and would not be undermined by any arguments demonstrating the preferability of non-classical rules of inference.

> **25.** Perhaps minimalism can be squared with alternative
> logics; but it cannot be squared with the role that truth
> must play in the *foundations* of logic—in justifying one
> logic over another.

But this suspicion is also mistaken; for the concept of truth plays no substantial role in the justification of logic. To see this consider how a system of basic rules of inference, L, could conceivably be justified. One strategy, it might be thought, is to specify the meanings of the logical constants by principles in which truth is the primary semantic notion—by means of truth tables—and then to show on the basis of these principles that the rules of L preserve truth. The trouble with this strategy is that it is blatantly circular; for the principles specifying the meanings of the constants are just trivial reformulations of the very rules we want to justify. For example, the classical truth table for 'and' is

(7)

p	q	$p \& q$
T	T	T
T	F	F
F	T	F
F	F	F

which is tantamount to the rules of inference

(7*)

$$\frac{p, q}{\therefore p \,\&\, q} \qquad \frac{-p, q}{\therefore -(p \,\&\, q)} \qquad \frac{p, -q}{\therefore -(p \,\&\, q)} \qquad \frac{-p, -q}{\therefore -(p \,\&\, q)}$$

$$\therefore \; (p \,\&\, q) \lor (-p \,\&\, q) \lor (p \,\&\, -q) \lor (-p \,\&\, -q)$$

for the difference between them is nothing more than the equivalence principles of truth and falsity. When these rules are supplemented with the rules underlying the other constants we obtain the classical propositional calculus. Thus the truth tables are not substantially different from the rules of inference that we are trying to support, and so do not form a suitable starting-point.

A second strategy for the justification of logic *L* would begin with the specification of the meanings of the logical constants by means of principles whose main semantic notion is something other than truth—e.g. assertibility or proof. Consider for example the intuitionistic account of disjunction:

(8)   Something is a proof of '*A* ∨ *B*' just in case we can see that
      it is (or will yield) a proof of '*A*' or a proof of '*B*'.

One might then hope to justify logic *L* by showing that only proofs in *L* would accord with these principles. But we should appreciate two facts about this strategy.

In the first place, no account of truth is involved. So there is certainly no reason here to think that something beyond the minimal theory is required in the foundations of logic. It might be thought that assumptions about truth must play a role at some stage; for isn't the whole point to show that the rules of *L* preserve *truth*? But even given that characterization of our goal, minimalism will do. If we can show that *L* provides the only system of rules of inference that accord with the meanings of the logical constants, then we have thereby justified the rules of *L*. Then, if we want, we can introduce the equivalence schema to derive the conclusion that if the premises of an argument of *L* are true then so is the conclusion. But this final step is not needed and does not invoke anything beyond the minimal theory.

The second point about this strategy is that it suffers from the same flaw as the first strategy, in that it does not really offer a noncircular justification of logic. The principles that, for each logical

constant, constrain the assertibility conditions of sentences containing them, have no epistemological priority over the rules of inference one might derive from them. Both the assertibility conditions and the basic rules of inference each determine regularities in the use of the logical constants. So, in so far as meaning is constituted by use, they could equally well be regarded as specifying the meanings of the constants. And if it is legitimate to accept the principles that specify assertibility conditions without further justification, then it should be legitimate to accept the basic rules of inference in the first place.

My inclination is to think that the principles of deductive logic are neither susceptible of, nor in need of, demonstrative justification. Instead, I would follow Quine (1953) in supposing that the logical principles relied upon in scientific theorizing are justified in so far as they are members of the simplest way of accommodating experience. I would say, moreover, that the logical principles deployed *outside* science in ordinary life—which may or may not differ from those used in science—are not subject to revision in light of experience. Hence these commitments are a priori; their justification is pragmatic; purely practical goals are furthered by adopting them.[5] I have not argued for these conclusions here. All I have tried to show is that alleged justifications of logic from assumptions about the meanings of the logical constants are not substantive, and anyway do not threaten the minimalist conception of truth.

26. How can truth-value gaps be admitted?
(Dummett, 1959)

They can't be. Begin with the minimalist view of falsity formulated either as

(2*) $\langle p \rangle$ is false $\leftrightarrow$ <u>not</u> [$\langle p \rangle$ is true]

or as

(2**) $\langle p \rangle$ is false $\leftrightarrow$ <u>not</u> p.

Either of these definitions, in conjunction with the equivalence schema

---

[5] For further discussion see my 1997*e*, reprinted as ch. 6 of *Meaning* (1998).

(E) $\langle p \rangle$ is true $\leftrightarrow p$

yields the conclusion

(9) $\langle p \rangle$ is not true and not false $\rightarrow \underline{\text{not}}\ p\ \&\ \underline{\text{not}}\ \underline{\text{not}}\ p$.

Thus we cannot claim of some proposition that it has no truth value, for that would imply a contradiction. Moreover, given classical logic (in particular, the law of excluded middle, '$(x)(Fx \vee -Fx)$'), we can go even further, not merely refraining from the claim that some propositions are neither true nor not true, but asserting positively that every proposition is true or not—i.e. true or false.

Admittedly these results do not derive solely from the minimal theory of truth, but depend also on our having defined falsity as the absence of truth. So one might conceivably make room for propositions that are neither true nor 'false' by constructing a narrower conception of 'falsity' than the one I have endorsed; one might hope to distinguish various different ways for a proposition to be untrue— only one of which is 'being false'. The reasons for supposing, on the contrary, that the minimalist characterization of falsity is the right one are as follows:

(a)   The account reflects our pre-theoretical intuition that if a proposition is *not* true then it is false, and that if something is *not* the case then the claim that it *is* the case would be false.

(b)   No reasonably plausible alternative characterization of falsity is able to accommodate these features of the concept.

(c)   The minimalist picture of truth encourages a parallel account of falsity—such as (2**)—according to which its attributions are similarly equivalent to non-semantic propositions.

(d)   The spirit of minimalism precludes accounts of truth and falsity which would equip them for theoretical work in semantics.

Thus we have ample reason to accept the minimalist view of falsity— embodied in principles (2*) or (2**)—and to embrace the conclusion that no proposition can be neither true nor false.

27. But doesn't philosophy require truth-value gaps in
order to accommodate such phenomena as non-referring
names, vagueness, the emotivist conception of
ethics, etc.?

These are all cases in which the misimpression that truth is a funda-
mental ingredient of reality has fostered the idea that it should play a
substantive role in philosophical theorizing. The minimalist position,
on the other hand, is that truth can play no such role, and that when-
ever we are tempted to give it one, the philosophical issues will be
clarified and resolved by recognizing that this is a mistake.

In the first place, as Russell (1905) argued quite convincingly
against Frege (1891), an atomic proposition entails that the referents
of its singular terms exist: *a is F* entails that *a exists*; and in that case,
it is natural to allow that if it is false that *a exists* then it is false that *a
is F*. Therefore atomic propositions containing vacuous singular
terms may very plausibly be regarded as *false* and don't call for truth-
value gaps. One way of ensuring this result is to combine Russell's
theory of definite descriptions (Russell, 1905) and Quine's strategy
for predicatizing names (Quine, 1953), to produce logical forms that
are free of singular terms. For example,

(10)   Everyone has an ancestor from Atlantis,

which contains the empty name 'Atlantis', becomes

(10*)   There is a place with the property of *being-Atlantis*, and
        everyone has an ancestor from there,

which is uncontroversially false.

28. It is obvious that many predicates—for example, 'blue',
'small', 'bald', 'heap'—do not have definite extensions;
and when such predicates are applied to certain objects
the result will surely be propositions with no truth value.

We can grant that there are indeed propositions—notably those
attributing vague properties to borderline cases—whose truth values
can never be determined and which might well be described as

'having no determinate truth value' or 'being objectively neither true nor false'. But we are not forced, on that account, to give up the classical law of excluded middle: $(x)(Fx \vee -Fx)$. And this is just as well. After all, *classical* logic is attractively simple and familiar. Furthermore, it is the logic that has been extracted from ordinary linguistic practice, and ordinary language is predominantly vague; so it would be remarkable if the phenomenon of vagueness were to dictate any departure from classical logic.[6] Moreover, once we have steeled ourselves to apply the law of excluded middle come what may, it is not *additionally* counterintuitive to retain also the principle of bivalence: that every proposition is true or false. On the contrary, its rejection leads quickly to contradiction, as we saw in the answer to Question 26. Thus we should retain the classical principles if at all possible. But the urge to back away from them will not be completely relieved until we have a satisfactory treatment of vagueness that will plainly allow them to be preserved.

In order to obtain such an account, a useful step is to recognize a distinction between ordinary *truth* and *determinate truth* (Field, 1986; Wright, 1987), enabling us to assuage our intuitions by saying that a proposition in which a vague predicate is applied to a borderline case is *not determinately true* but might nonetheless be *true*. A natural explication of the needed notion of determinacy may be given by reference to the way in which the meaning of a vague predicate prevents us from finding out whether its application in certain cases would be correct or incorrect: we can say that an object with the property F-ness is *determinately F* when there is no such semantic obstacle to discovering that it is F, and that it is *not determinately F* when there is such an obstacle—i.e. when the meaning of the predicate 'F' precludes the prospect of our arriving at a stable conclusion about whether or not it applies. Consider, for example, the ascription of 'heap' to an unclear case—a little collection of sand particles.

---

[6] It would not be appropriate to restrict the application of classical logic to propositions with a determinate truth value. In the first place, such a strategy would not generally be workable. For sometimes the use of logic is required to see what entails a given proposition and thereby to discover whether or not it has a determinate truth value. And in that case, the proposition could be tested for its conformity with logic only by already presupposing that it will pass the test. And secondly, any restriction of logic to propositions that are definitely true or definitely false would conflict with the a priori character of logical laws, since their applicability would become contingent on the favourable outcome of empirical investigation—the determination of whether we are, or are not, dealing with a borderline case.

Given the meaning of 'heap', and given the number, size, and arrangement of the grains, we cannot become inclined to attribute either 'heap' or 'not heap' to this collection. Therefore

(11)   It is neither determinately a heap nor determinately not a heap,

even though, assuming classical logic,

(12)   It is, or is not, a heap.

Therefore, in light of the minimalist accounts of truth and falsity, we may conclude that

(13)   ⟨It's a heap⟩ is not determinately true and not determinately false,

although

(14)   ⟨It's a heap⟩ is true or false.

Thus we are able to accommodate the phenomenon of vagueness without questioning the law of excluded middle, the principle of bivalence, or the minimalist conception of truth.

It remains, however, to say what it is about a predicate that makes it vague. What character does the meaning-constituting property of a vague predicate have to have in order to explain the indeterminacy that is characteristic of vagueness—the semantically induced impossibility of coming to know, in borderline cases, whether or not it applies? The answer, I would suggest, is that the vagueness of predicate, 'F', consists in the distinctively 'gappy' character of the basic regularity governing its use. Very roughly speaking such regularities have the form

(V)   'F' is applied to things which have underlying property # to a degree greater than y, and 'not F' is applied to things that have # to a degree less that n, and neither is applied to things that have # to an intermediate degree,

where y is greater than n.[7] For in so far as 'F' is deployed according

---

[7] The basic use regularities of vague predicates are not quite as simple as (V), since (a) they usually involve *several* underlying parameters whose values fix whether the predicate applies, and (b) they engender higher-order indeterminacy (i.e. the boundaries of the 'clearly F' and the 'clearly clearly F', etc. are also unclear).
The present account is set within a use theory of meaning (sketched in Chapter 6)

to such a basic regularity, we will indeed be unwilling to apply either it or its negation to the things that have # to some degree between y and n, and we will indeed be confident that no further discovery could help matters. This account of vagueness explains its main symptom—namely, our irremediable inability to decide in certain cases whether or not a term applies—without suggesting any departure from excluded middle or bivalence.

Any decent theory of vagueness must also have something to say about the notorious sorites paradox:

(Sor)   0 grains cannot make a heap;
         For any n, if n grains cannot make a heap, then $n + 1$
            grains cannot make a heap either;
         Therefore: For any n, n grains cannot make a heap,

which seems to show, on the basis of incontrovertible premises and impeccable reasoning, that there are no heaps. In so far as we want to give up neither classical logic nor the view that some things are heaps and others are not, then the only remaining option—which I think we must learn to happily embrace—is to deny the second premise. That is, we must allow that there is some unknown (indeed unknowable) number, h, such that h grains cannot make a heap but $h + 1$ grains can. Thus we are allowing that the predicate 'is a heap' has an extension—albeit an indeterminate one. True, we could not, even in principle, discover the extension. In particular, we could never know the fact of the matter as to whether our little pile is a heap. For, as we have seen, such knowledge is precluded by the very meaning of the word—by its being vague. But why should this be thought odd or implausible? It is surely only the lingering seductiveness of verificationism—an inclination to hold that the existence of a fact requires the conceivability of knowing it—that gives rise to discomfort with this situation.[8]

according to which the overall use of each word is explained by means of some *basic* regularity in its use, and the word means what it does in virtue of being governed by that basic regularity. For a more thorough treatment of vagueness from this point of view see Horwich, 1997*b*. See also Tim Williamson (1994), who also retains excluded middle and bivalence in the context of vagueness, but who offers a very different account from mine of why knowledge is impossible in borderline cases.

[8] A further source of discomfort with the idea that a vague predicate might apply (albeit indeterminately) in a borderline case is that there could be no explanation, in terms of its use, of *why* it applies. However (as I argue in *Meaning*), it is a mistake to expect any such explanation. Even though the meaning of each word is constituted by

I should stress that the above treatment of vagueness and of the sorites paradox has substantial merits, and should not be viewed merely as a necessary, but rather unwelcome, outgrowth of minimalism. In the first place, it is not the case that minimalism *dictates* our solution. The situation, rather, is that minimalism primes us to expect—what could well be recognized independently—that the problems of vagueness are simply not addressed by an account of truth. Once this is appreciated, we can see that the real choice is between the abandonment of classical logic and the abandonment of verificationism. My sentiments in favour of the latter option are independent of minimalism. That response succeeds in preserving the simple and well-entrenched principles of excluded middle and bivalence, offers a positive analysis of vagueness, proposes a way out of the paradox, and gives an explanation of why that way out might have seemed so counterintuitive. Thus our approach to indeterminacy is not just a byproduct of minimalism, and has a great deal of plausibility in its own right.

The rejection of minimalism would be of no help whatsoever in dealing with these matters. For, as we have seen, the only genuine alternative to what is proposed here involves the rejection of classical logic—which is a move not made any less difficult by the adoption of a non-minimalist theory of truth.[9] Moreover, the semantic principle

its use, there need be no explanation of why a given use should provide a predicate with the particular meaning (and hence extension) that it has.

Frank Jackson has expressed a related concern. He thinks it would be a violation of our sense of symmetry (i.e. the principle of sufficient reason) to suppose that one of the pair of propositions 'It is a heap' and 'It is not a heap' is a fact and the other isn't, when what we have is clearly a borderline case. However, although I don't doubt that symmetry is appealing, it seems to me that we would rather accept that the world is disappointingly asymmetrical, than abandon classical logic.

[9] It is far from clear what the new logic would be. Putnam (1983) suggests intuitionistic logic. However, a 'vagueness logic', in so far as it is motivated by the idea that 'indeterminate truth' is impossible, would need to renounce the inference from the conditional

k is $F \rightarrow k$ is determinately F

to its contrapositive,

k is not determinately $F \rightarrow k$ is not F.

Otherwise, in a case where k is evidently a borderline case, i.e.

k is not determinately F and not determinately not F,

it could be inferred that

(E) $\langle p \rangle$ is true iff p,

which binds our logic and metalogic, and which prevents us from combining classical logic with the thesis that neither $\langle A \rangle$ nor $\langle \text{not } A \rangle$ is true, is by no means peculiar to minimalism. Just about all accounts of truth subscribe to it. So it certainly cannot be suggested that by adopting minimalism we are depriving ourselves of an easy solution to the problems of indeterminacy.

Finally, it is perhaps worth noting how certain superficially different approaches to vagueness might be partially accommodated to the present point of view. Consider, for instance, an idea of Kit Fine's (1975), endorsed by Dummett (1978), that a sentence containing vague expressions is absolutely TRUE just in case it is true relative to every admissible way of making the vague terms precise. If we take 'TRUE' to mean 'Determinately true' then this idea becomes a perfectly consistent and illuminating elaboration of the minimalist view sketched above. But if it is assumed that nothing can be true unless TRUE, as Fine and Dummett are inclined to say, then the result is disaster. For we can infer that neither $\langle A \rangle$ nor $\langle \text{not } A \rangle$ is true, which implies a contradiction. So the former interpretation is clearly preferable. Another strategy for dealing with vagueness has been to invoke infinitely many TRUTH VALUES: all the real numbers from 0 to 1. The idea is that when a vague predicate, 'F', is applied to a borderline case, n, the resulting TRUTH VALUE is somewhere in the middle between 0 and 1; that the conditional '$F(n) \rightarrow F(n + 1)$' has a TRUTH VALUE just less than 1; and that the TRUTH VALUE of the conjunction of many such propositions gradually decreases as the number of conjuncts increases—reaching 0 when we get to '$(n)[(F(n) \rightarrow F(n + 1)]$'. As before, these ideas can be reconciled with the minimalist viewpoint. It suffices to identify 'TRUTH VALUE = 1' with 'Determinate truth', 'TRUTH VALUE = 0' with 'Determinate falsity', and (perhaps) 'TRUTH VALUE of $\langle p \rangle = x$' with something like

k is not F and k is not not F

—a contradiction. Thus 'vagueness logic' must renounce *contraposition*—a principle which is perfectly acceptable from the intuitionistic point of view. Moreover, there is no reason for a 'vagueness logic' to quarrel with the principle of double negation elimination,

k is not not $F \rightarrow k$ is F,

which is intuitionistically *un*acceptable. Thus intuitionism is not the logic of vagueness.

'The probability of $\langle p \rangle$ is x'. So far, so good. Difficulties arise only if the new TRUTH VALUES, 1 and 0, are identified with truth and falsity. But this step can and should be resisted. In order to eliminate the paradoxes of vagueness it suffices to appreciate, first, that a notion of *determinacy* should be invoked to articulate the nature of vagueness; and second, that this will make it unnecessary to tamper with any of the entrenched laws of logic and semantics, and will enable us to accept with equanimity the falsity of the sorites' main premise.

29. There is a substantive issue in meta-ethics as to whether evaluative utterances purport to assert truths or whether they are merely expressions of feeling; but this question would be trivialized by minimalism.

There has indeed been a tendency for ethical emotivists (also known as 'non-cognitivists' and 'expressivists'), to want to use the notion of truth value to distinguish 'genuine descriptions' from syntactically similar sentences whose linguistic role is arguably non-descriptive. And this practice certainly is at odds with a minimalist perspective from which (1) ethical pronouncements express genuine *propositions* (as we shall see in the next section), for they form 'that-clauses' (e.g. 'that honesty is good') which function in the normal way to designate objects of assertion, belief, etc.; and (2) ethical propositions provide perfectly good and useful instances of the equivalence schema— instances which are needed to formulate generalizations (e.g. in logic) that cover such propositions.

However, the moral here is not that minimalism and emotivism are incompatible, but that emotivism should be reformulated. For a minimalist might happen to accept the emotivists' central insights: namely, that the function and assertibility conditions of certain ethical claims are fundamentally different from those of empirical descriptions, and that an appreciation of the difference will help to resolve philosophical problems surrounding the notion of an ethical fact. My point is neither that this position is entailed by minimalism; nor that it is correct; but that it need not, and should not, be formulated in such a way as to preclude the minimalist conception of truth. It should be articulated, rather, as a view about the unique character

of ethical propositions. More specifically, the emotivist might attempt to characterize the nature of ethical propositions by maintaining, *very* roughly speaking, that the meaning of 'X is good' consists in the fact that it is asserted by someone when and only when he wants X (which is *not* to say that 'X is good' means 'I want X'). This sort of account (suitably refined) has some claim to being able to dissolve some of the epistemological problems of ethics and explain why certain ethical beliefs have motivational force. Thus the essential character of emotivism might be captured without having to question the existence of ethical propositions, beliefs, assertions, etc., and without having to deny that they satisfy the usual logical and metalogical principles.[10]

[10] The bearing of minimalism on the formulation of emotivism is debated in a series of papers by myself (1993, 1994), Michael Smith (1994*a*, 1994*b*), and Jackson, Oppy, and Smith (1994).

6

Propositions and Utterances

30. Propositions are highly dubious entities. It is unclear what they are supposed to be, and their very existence is controversial. Would it not be better, therefore, to develop a theory of truth that does not presuppose them—by assuming, for example, that *utterances* are the primary bearers of truth?

According to the advocate of propositions, whenever anyone has a belief, a desire, a hope, or any of the so-called propositional attitudes, then his state of mind consists in there being some relation between him and a special kind of entity: namely, the *thing* that is believed, desired to be the case, hoped for, etc. Thus if Oscar believes that dogs bite, this is alleged to be so in virtue of the obtaining of a relation, *believing*, between Oscar and a certain proposition: namely, *that dogs bite*.

The considerable merit of this theory is that it appears to provide an adequate account of the logical properties of belief attributions and the like. We are inclined to infer from

   (1)   Oscar believes that dogs bite

that

   (2)   There is something Oscar believes,

and from

   (3)   Oscar doubts that it will rain

and

   (4)   Barnaby is saying that it will rain

that

(5)   Oscar doubts what Barnaby is saying.

Such inferences may be subsumed under familiar logical rules provided that apparently singular terms such as

(6)   what Barnaby is saying

and

(7)   that it will rain

are taken at face value to refer to a species of object—to be called 'propositions'.

It might well be countered, however, that such evidence should not be taken as conclusive. After all, syntax is not an infallible guide to semantic structure. Indeed, one of the central concerns of this book—whether the truth predicate does or does not stand for some sort of property—derives from this well-known fact. And there are many less controversial examples: e.g. 'Jones's sake', 'The average man', 'Smith's appearance'. The last of these examples is particularly close to the issue at hand. There is a relation of *resemblance* between certain things, but this relation is not transitive: it may be that A looks like B and B looks like C, but A does not look like C. Given this failure of transitivity, we cannot suppose that there is a realm of entities, *appearances*, such that every object possesses exactly one of them, which it shares with all the objects it resembles; for that supposition would imply transitivity. Thus although we do express resemblance claims by saying that certain things have the same appearance, it turns out on closer scrutiny of the logical properties of such claims that their structure cannot be what would seem most natural given their syntactic form: namely,

(8) $(\exists x)(x$ is the appearance of A and x is the appearance of $B)$.

Similarly, it should not be taken for granted that our ordinary talk of 'what Oscar believes' is to be construed at face value as referring to a special kind of entity.[1]

---

[1] Particular grounds for doubt lie in the possibility that *translation* is affiliated with the notion of *resemblance* (roughly, via the identification of intertranslatability with resemblance in use), and is infected with its intransitivity. If this were so, then we could not suppose there to be a realm of entities such that every declarative utterance

In order to clarify this issue, let us step back for a moment from the case of propositions and consider more generally why it is that we impute logical form in the way that we do. The logical forms of the sentences in a language are those aspects of their meanings that determine the relations of deductive entailment holding amongst them. Let us imagine a body of sentences characterized by their concern with a certain range of phenomena; and suppose that we have mounted an investigation into the relations of deductive entailment that hold amongst these sentences. Suppose that the results of our investigation suggest an attribution of logical forms having the implication that some of the sentences will clearly entail the existence of entities of a certain type—call them 'Ks'. Suppose, finally, that we believe that some of those sentences express truths. Taken together, these considerations would provide a basis for thinking that things of type K exist. But how powerful are these reasons? Under what circumstances, if any, would it be right to resist them— to maintain that despite the aforementioned evidence there are really no such entities as Ks?

Of course, one thing that would justify resistance would be the discovery that our assignment of logical forms is unsatisfactory. Suppose we found another way of doing it which gave a more complete representation of our inferential practice, and which involved no commitment to Ks. In that case our prior ground for believing in them would be completely undermined.

Another possible source of resistance would be the discovery of non-philosophical arguments for the conclusion that Ks don't exist. In calling these arguments 'non-philosophical' I have in mind that they would come from *inside* the field to which the statements in question belong. Consider, for example, arguments within physics that the aether does not exist, or arguments within zoology against the Loch Ness Monster. These are evaluated with respect to the same canons of justification that govern the original body of statements. If they are deemed acceptable then the result is a revision in our scientific beliefs; but no change is called for in the logical forms that we attribute to them.

is associated with exactly one of these entities, sharing it with all the other utterances with which it is intertranslatable. Thus the existence of propositions is put in question. This problem is addressed directly in the answer to Question 32.

Finally, the most philosophically interesting case is that in which general philosophical considerations motivate a disinclination to postulate Ks. It might be argued, for example, that only *material* objects exist and Ks would not be material; or that sheer ontological parsimony requires us to do without Ks if at all possible; or that, given the nature of Ks, knowledge about them could not be squared with otherwise attractive epistemological theories; or that we would be unable to answer basic questions, such as 'What are Ks?', 'When is K_1 the same as K_2?', 'Where are Ks?', to which we would have every right to expect answers if Ks really existed.

There are three types of response to any such argument: (a) we may regard it as fallacious, and proceed to explain how this is so; (b) we may find it persuasive, accept that Ks don't exist, and conclude that a certain body of what we used to believe is mistaken; or (c) in order to preserve our earlier scientific beliefs and still be able to accept the conclusion that Ks don't exist, we might abandon the account of logical forms that involves commitment to Ks and replace it with one that doesn't.

Of these alternatives it seems clear that option (a) is always best; for the arguments that it asks us to reject are extremely weak in the first place. Often they involve barefaced overgeneralization of the following sort. First, material objects are taken to be paradigm examples of what exists; secondly, certain prominent properties of such objects are identified; thirdly, it is inferred that only entities with these properties could exist; fourthly, it is noticed that Ks would not have them; and finally, the conclusion is drawn that Ks cannot exist. Evidently no great conceptual strain is involved in rejecting such arguments, which beg the whole question in their first premise. So option (a) is quite acceptable. The other alternatives, however, exhibit some highly undesirable features. Option (b) implies that we must start denying certain things that we presently regard as certainly true. Option (c) involves the idea that the correct logical forms are not those that provide a perfectly adequate account of inferential practice. We are to reject a certain way of articulating our beliefs solely because it has consequences that are irrationally regarded as unwelcome, and even though it accords precisely with our entrenched criteria for being a good logical articulation.

It is easy to see how these general conclusions will apply to the case of propositions. In the first place, we can suppose that an adequate account of the logical forms of belief attribution (and other

so-called propositional attitude statements) involves the supposition that 'that *p*' is a singular term. Thus from 'Oscar believes that dogs bite' we can infer 'Oscar believes something', and this is assimilated to the rule of existential generalization. In the second place, we may assume that some propositional attitude statements are certainly true. For example, it is certainly true that Einstein claimed that matter is a form of energy. Thirdly, we should take these assumptions to entail that there is an entity, *what Einstein claimed*, or in other words, *the proposition that matter is a form of energy*. And fourthly, we should not be troubled by an inability to say *where* this entity is, or *what* it is made of. For such questions are not appropriately asked of things like propositions. Similarly, it should not trouble us that propositions will not enter into causal relations with us, so that our knowing anything about them will violate the 'causal theory of knowledge'. For this theory derives its entire prima facie plausibility from the sort of blatantly question-begging overgeneralization mentioned above—in this case, from the unjustified presupposition that everything is known in the same way that physical objects are known.

I conclude that a compelling argument for the existence of propositions may be built on the premise that they participate in an adequate account of the logical forms of belief attributions and similar constructions. Moreover, the required premise appears to be correct. Therefore, despite their peculiarities, we should not balk at propositions and should not object to their use in a theory of truth.

31. The case for propositions assumes the adequacy of a certain logical analysis of belief—one that construes the state of belief as a relation between a person and a kind of entity, the *content* of the belief. But this assumption is plagued with familiar difficulties and appears to be mistaken.

The main alleged difficulties have to do with a dispute about the nature of propositions—a dispute which goes back to Frege and Russell. Russell (1903) claimed that a proposition consists of the very objects it is about. For example,

(9) the proposition *that Hesperus is visible*

would be made up of the object, Hesperus, and the property of *being visible*. If Russell was right, then, since 'Hesperus' and 'Phosphorus' have turned out to be two names for the same planet, then

(10) the proposition *that Phosphorus is visible*

is the very same proposition as (9). But how, in that case, do we account for the fact that someone may be aware—as we say, *de dicto*—that Hesperus is visible and not that Phosphorus is? Frege (1892) had already posed this question, and answered it by maintaining that the proposition expressed by a sentence (what he called a 'thought') is composed of the *senses*, rather than the *referents*, of its constituent words. So (9) and (10) would be different from one another. But if he was right then so-called *de re* beliefs become problematic. For if, in that sense, someone believes, regarding Phosphorus, that it is visible, it does follow, given the identity of Phosphorus and Hesperus, that he has that belief about Hesperus too.

One way of dealing with these problems is to acknowledge the existence of *both* Fregean *and* Russellian propositions—the first being the objects of *de dicto* belief, and the second of *de re* belief. The idea is that all propositions have some compositional structure (e.g. $\$(x)$, $@(x,y)$, $(x)(\exists y)(\$(x) \rightarrow @(x,y))$, etc.), and that each singular position in such a structure may be filled either by a Fregean sense or by the referent of such a sense. Thus there are pure, Fregean, *abstract propositions*, in which a compositional structure is filled only with senses; there are pure, Russellian, *concrete propositions*, in which each singular position in a compositional structure is filled only with referents; and there are *mixed propositions*, in which some of the singular locations are occupied by senses and the others by objects. For example, corresponding to the sentence

(11) Hesperus is identical to Phosphorus

there is an abstract proposition which consists of the senses of 'is identical to', 'Hesperus', and 'Phosphorus' embedded, in that order, in the structure '$@(x,y)$'; the concrete proposition consisting of the referents of 'Hesperus' and of 'Phosphorus', and the property of identity, embedded in that structure, and two different mixed propositions. Therefore,

(12)   Raphael believes that Hesperus is Phosphorus

has four readings[2]—one for each of the propositions to which one might be saying that Raphael is related. Disambiguation is achieved (in English) by use of the qualifier '*of X*' for any of the propositional constituents that are intended to be referents rather than senses. Thus, if one says

(13)   Raphael believes *of Hesperus* that it is Phosphorus,

the object of belief is the mixed proposition in which the '*x*'-position in '$@(x,y)$' is filled by the object, Hesperus, and the other positions are filled with senses. And if one attributes the belief without any such qualification, what is conveyed is a *de dicto* belief—a belief whose object is the purely abstract proposition.

Moreover, it seems plausible to suppose that in order to have a *de re* belief *of* something, it is necessary to have in mind a special way of conceiving of the thing—a way that is sufficiently intimate to constitute 'knowing what the thing is'.[3] Therefore a certain relationship exists between *de dicto* and *de re* belief. In order for someone to believe a proposition that is not entirely abstract (i.e. in order for him to have a *de re* belief), two things are required. First, he must believe some wholly abstract proposition from which the other may be derived by replacing some of the senses with their referents. And, secondly, the believer's conception of these referents must be such as to qualify him as 'knowing what they are'. Admittedly the second condition is vague and context-dependent. But this is entirely appropriate, since the set of circumstances in which *de re* belief may properly be attributed is also vague and context-dependent.

It seems, therefore, that prima facie difficulties in what would otherwise appear to be the most straightforward construal of belief attributions may be overcome. Beliefs should indeed be analysed as relations between persons and propositions. Thus our practice of belief attribution commits us to the existence of propositions.

[2] I am assuming (a) that the referents of only the *singular* terms can be what *de re* beliefs are about, and (b) that the Russellian property of *F*-ness is the same thing as the Fregean sense of '*F*'. For elaboration of this approach—including discussion of *de se* attitudes and complications arising from context-sensitivity—see my *Meaning* (1998), ch. 3.

[3] See Hintikka, 1962, Kaplan, 1969, and Quine, 1977, for discussion of this sort of requirement for *de re* belief.

32. 'The proposition *that p* is true iff *p*' may be thought to capture our conception of truth only if truth is not already presupposed in the very idea of a proposition. But this requirement may well be violated. For a central component of the notion of a *proposition* is lodged in the statement of identity conditions for propositions—the conditions for two utterances to express the same proposition. But this is an idea one might plausibly explain in terms of the *inter-translatability* of the utterances, which, in turn, must be construed as their having the same *truth* conditions. And if the concept of truth is needed to say what propositions are, then a theory of truth cannot take propositions for granted.

A complete treatment of this issue would go beyond the scope of this essay. However, I agree that our having the concept of proposition presupposes that we can make sense of intertranslatability (interpretation) claims of the form

(14)   *u* expresses the same proposition as my present utterance of '*p*',

which, taking for granted that

(15)   My present utterance of '*p*' expresses the proposition *that p*,

will yield

(16)   *u* expresses the proposition *that p*.

Therefore, it is incumbent on me to say something about the notion of *translation* on which I am relying—at least enough to quell the suspicion of circularity.

The general conception of meaning and translation to which I wish to appeal is the so-called 'use theory' articulated in various forms by Wittgenstein (1953), Sellars (1954), Quine (1960), Harman (1982, 1987), Peacocke (1986, 1992), and many others. Roughly speaking, the right translation between the words of two languages is the mapping that preserves basic patterns of usage—where usage is characterized non-semantically, in terms of circumstances of application, contribution to the assertibility conditions and inferential role of containing sentences, etc. What counts as the 'basic

patterns' are those regularities that (in conjunction with other factors) best explain overall usage. For example, the basic use of the word 'red' in English is arguably that it is applied to observably red things and unapplied to observably not-red things; the basic use regularity governing the word 'and' is arguably that we are disposed to infer 'p and q' from 'p' and 'q', and vice versa; and the basic use of 'true', as we have seen, is that we are inclined to accept instances of 'The proposition *that p* is true iff p'.[4]

Thus the concept of translation may be explicated, as required by minimalism, in entirely non-truth-theoretic terms. However, this will not quite suffice for an account of propositional identity. For in order to determine which of my present utterances expresses the same proposition as an utterance of u, we must not only translate u into English but also take into account any differences between the contexts in which u and my utterance are produced. For example 'J'ai faim' uttered by Pierre at noon yesterday expresses the same proposition as my now saying 'Pierre was hungry at noon yesterday'. This is so in virtue of the fact that 'J'ai faim' translates into 'I am hungry', which is then adjusted for contextual variation by recourse to a fixed set of rules:

'I' becomes a term 'a', such that a is the speaker;

'now' becomes a term 't' such that t is the time of the utterance;

'here' becomes some term 'l' such that l is the place of the utterance;

---

[4] For a full articulation and defence of the use theory of meaning see my *Meaning* (1998). This account bears some resemblance to the 'conceptual role semantics' articulated by Field (1977) and Block (1986). A vital difference between them, however, is that according to their view the meaning of a word is identified with its *internal* role, whereas according to the use theory, as I understand it, the meaning of a word also includes its use in relation to the external world.

For certain terms it may be objectively unclear precisely which regularities for their use are explanatorily fundamental. However, this is by no means fatal as far as propositions are concerned. For if there is no determinate fact of the matter as to whether two utterances are intertranslatable, then there is no determinate fact of the matter as to whether they express the same proposition. But, as we saw in the answer to Question 28, such facts may be indeterminate but none the less exist. Indeterminacy is found throughout our conceptual scheme—whether we are talking about material objects, persons, societies, species, or whatever—and so its presence in the domain of propositions too should be no cause for sceptical concern.

'that *F*' becomes a term '*d*' such that *d* is the *F* to which the
speaker was attending;

... etc.

Note that the rules of context adjustment do not invoke the concept
of truth. Moreover, as we have seen, the primary determinant of
propositional identity—namely, translation—is explicable in terms
of 'identity of use', where 'use' is characterized in terms that do not
cite truth. Consequently, propositions are not grasped via the notion
of truth; and so we are free to suppose that this notion is entirely
captured by the equivalence principle.[5]

33. But the 'use theory' of meaning implies that proposi-
tions don't exist. For if *translation* is a matter of resem-
blance in use, then it is not a transitive relation, and so
there can be no such things as 'what intertranslatable
utterances have in common'. (Harman, 1973)

Formulated more carefully, this argument against propositions pro-
ceeds by comparing the following three principles:

(A) Two utterances express the same propositions just in case
they are intertranslatable.

(B) Utterances are intertranslatable just in case they have corre-
sponding constituents with a similar use.

(C) There are words, *x*, *y*, and *z*, such that the use of *x* resem-
bles the use of *y*, which resembles the use of *z*; but the uses
of *x* and *z* do not resemble one another.

These principles are inconsistent with one another. For the intransi-
tivity of resemblance expressed in (C), combined with the use theory
of translation expressed in (B), together entail the intransitivity of
translation. But, combining this with (A), which says that inter-
translatable utterances express the same propositions, we can infer

---

[5] For further discussion of this point—specifically the relationship between *under-standing a sentence* and *knowing its truth conditions*—see the answer to Question 21.

the possibility that u and v express the same propositions, v and w express the same propositions, but u and w do not express the same propositions—which is absurd. Thus, given that relations of resemblance clearly are intransitive, the use theory of translation (as articulated in (B)) entails that the identity conditions for propositions are incoherent, and therefore that propositions don't exist.

My response is to modify premise (B). The use theory of meaning is indeed correct. But the argument shows that premise (B) gives a misleading rendering of it. What we should say instead is that two words are intertranslatable only when they have *the same basic use*— and not merely *similar* uses. But we must acknowledge that the observable facts of linguistic behaviour, although providing the only possible guide to the existence of this relation of sameness, may not always be sufficient to establish it objectively. For there may be alternative but equally good ways of explaining the overall deployment of a word in terms of alleged basic regularities in its use. In that case it will be indeterminate which of these patterns of use is really basic. Consequently it may be indeterminate whether or not two words conform to the same basic patterns of use. Thus claims of intertranslatability based on sameness of use are not always determinate. Two people could have equally good grounds for divergent claims and there be no way to settle the matter. So the relation between translation (or 'same use') and observed linguistic behaviour is like the relation between baldness and amount of hair, or between poverty and net assets. In each case attribution of the former notion can be based only on information about the latter; however, there is often a degree of indeterminacy—a semantically caused, irremediable ignorance regarding the precise conditions in which the former concept applies. None the less, there does exist a fact of the matter. Either two words *are* properly intertranslatable, or they are not— even though it may be impossible to say which is so.[6]

I have been trying to rebut an argument designed to show that translation is intransitive—a conclusion that would undermine the existence of propositions. However this rebuttal does not prove that translation *is* transitive. Of course, one way of showing this is to derive it from the existence of propositions. But if one is nervous about propositions precisely on the grounds that translation might

---

[6] See the answer to Question 28 for a discussion of how vagueness and indeterminacy are treated from the minimalist point of view.

be intransitive, then this strategy is no good. An alternative line of thought offering some reassurance goes as follows.

It is implicit in the existence of propositional attitude constructions that a person may adopt a given attitude no matter what the natural language is in which he would express it. We do not regard it as a necessary condition for the truth of

(17)   Florence believes that snow is white

that Florence has our language available to her. Moreover, we do not suppose that this fact about what Florence believes, if it is a fact, depends on our having characterized it, or on our being able to characterize it in the particular way that we did—i.e. in English. One and the same propositional attitude of Florence's may be equally well described in other languages: for example, with

(17-f)   Florence croit que la neige est blanche.

But this sort of translatability of propositional attitude attributions implies that translation is transitive. For suppose Florence asserted in a third language some words whose translation into English is

(18)   Snow is white,

and suppose that on that basis we English speakers assert (17). If translation were intransitive then there would be no guarantee that

(18-f)   La neige est blanche

—which is the translation into French of the translation into English of Florence's words—is the translation into French of Florence's words. So there would be no guarantee that (17-f) manages to characterize the same propositional attitude as (17) does. Thus the assumption that our attributions of propositional attitudes are translatable in the normal way presupposes that translation is transitive.

More formally, suppose u and A are particular utterances in different languages, $L(u)$ and $L(A)$, and that A attributes a certain content to u by means of the sentence 'v'. Utterance A says in effect that 'v' expresses in the language $L(A)$ what u expresses in $L(u)$. Or, in other words:

(I) A is true iff the translation of u into $L(A)$ is 'v'.

Suppose, moreover, that *we* would translate the sentence 'v'—as it is

used within $L(A)$—as the sentence 'p'. Then, from our point of view, A expresses the statement

(B) u expresses the statement that p.

Therefore, A is true iff u expresses the statement that p. In other words,

(II) A is true iff our translation of u is the same as our translation of 'v' from $L(A)$.

Combining (I) and (II)—the two statements of A's truth conditions—we can infer that u and 'v'-in-$L(A)$ are intertranslatable just in case their translations into our language are the same. Therefore, translation is transitive. It can indeed define the equivalence classes with which propositions may be correlated.

> **34.** Many philosophers would agree that if propositions
> exist then *propositional* truth would be covered by
> something like the equivalence schema. But they might
> still maintain that the truth of an *utterance* consists in its
> 'correspondence with reality', or some other substantive
> thing. Thus, it is for *utterances* that the deflationary
> account is controversial, and this position has received
> no elaboration or defence. (Field, 1986)

Fair enough. So let me now indicate how minimalism deals with the truth of utterances. The initial deflationary impulse is to say that

(D?)   Any utterance of the sentence-type 'p', is true iff p.

But the trouble with this idea is that different utterances of the same type (e.g. 'I am hungry', 'Banks have money', 'Mary's book is on that table') have different truth values depending on the circumstances in which they are produced and on what is meant by them. Consequently, it is simply false that

(19)   *All* utterances of 'I am hungry now' are true iff *I* am hungry *now*,

and similarly for the other examples.

To avoid this difficulty we need a restricted form of the disquotational principle—one in which the way that '*p*' is construed when it is *mentioned* on the left-hand side is the same as the way it is construed when it is *used* on the right-hand side. In other words, we can endorse instances of the schema

(D′) $(u \in 'p') \rightarrow (u \text{ is true} \leftrightarrow p)$,

provided that they—specifically the right-hand sides of the biconditionals—are construed with respect to the same pertinent contextual variables as *u* itself. In other words, an instance of the disquotational schema holds if it is asserted in a context that is not relevantly different from the context of the utterance whose truth is in question.

Implicitly attempting to ensure that this condition is satisfied, Quine (1970) suggests that we restrict instantiation of the disquotational schema to so-called *eternal sentences*—sentences whose tokens exhibit no context-sensitivity because they don't contain elements like indexicals, demonstratives, and ambiguous words. As for utterances that *are* context-sensitive, Quine proposes to derive each of their truth values from the truth values of 'equivalent' eternal sentences. For example, a particular instance of

(20) I am hungry

may be true because of its 'equivalence' to the eternal sentence,

(21) Albert Einstein is hungry at noon on 1 January 1947.

However, there are a couple of difficulties with this strategy. In the first place, the problem of context-sensitivity is not really solved because the same string of words can appear in several languages and for that reason have different truth values on different occasions. Thus, strictly speaking, there are no eternal sentences. Moreover, as soon as one tries to patch things up by inserting reference to a *language* in the disquotation schema, i.e.

(D+) '*p*' is true in *L* iff *p*,

then the theory is undermined. For unless something is said about what it is for an utterance to be 'in language *L*', the schema will become a definition of *that* notion rather than a definition of truth. But if we do fill out the modified schema with an account of what it is for an utterance to be 'in language *L*', then, in effect, the theory will be explaining the truth of utterances in terms of the truth of the

propositions they express, and so its distinctively disquotational character will be lost.[7] In the second place, even if we allow Quine's language-relative notion of 'eternal sentence', it is unclear that there would exist enough of them to provide a sufficient basis for the truth values of all instances of non-eternal sentences. Ambiguity and context-sensitivity are by no means restricted to indexicals and demonstratives. Most names, predicates, and quantifiers can also be construed in various alternative ways. So much so that it is not easy to find English sentences whose instances must all have the same truth value.[8]

For these reasons I do not follow Quine's approach, but adopt an alternative way of making sure that instances of the disquotation schema are construed in the same way as the utterances whose truth condition they specify. My strategy is to introduce a kind of quote-name designed to pick out an expression-type on the basis, not only of syntactic form (i.e. physical character), but also on the basis of meaning (or, more specifically, the propositional constituent expressed). Thus '*dog*' picks out the word-type with a certain shape and a certain meaning; '*bank*' is ambiguous, since there are two different types (with the same shape but different meanings) that it can designate; similarly, '*I am hungry*' is multiply ambiguous, designating a separate expression-type for each of the limitless number of different propositions it may be used to express (depending on the speaker and time of utterance). Thus we can accept

(22) $(u \in $ *I am hungry*$) \rightarrow (u$ is true iff I am hungry),

provided the quote-name is understood appropriately. This will be so if the instance of the syntactic form 'I am hungry' in the consequent of (22) is a member of the type designated in the antecedent. And this is achieved by understanding the two tokens in the same way. Generalizing, we have an inclination to accept any instance of

(D) $(u \in $ *p*$) \rightarrow (u$ is true iff p),

when what is put in place of 'p' is given a uniform interpretation.

However, this is not quite enough to capture our concept of truth for utterances; for it leaves unspecified how we attribute truth to

---

[7] See Putnam, 1988, for a forceful presentation of this point.

[8] This point was impressed on me by Dan Sperber, Pierre Jacob, and François Recanati. See Searle, 1978, and Barwise and Perry, 1983.

utterances in foreign languages and, in general, to utterances whose types are not exemplified in our own current discourse (e.g. John's utterance yesterday of 'I am hungry'). In order to cover these cases, we must add a further principle enabling us to 'project' attributions of truth from our own current utterances onto other, 'equivalent' ones—where u and our 'p' are 'equivalent' when 'p' is the correct *interpretation* of u (i.e. $\text{Int}(u) \in$ 'p'), and where interpretation is a matter of translation plus context adjustment, as discussed in the answer to Question 32. Thus we suppose that

(23) $(\text{Int}(u) = v) \rightarrow (u$ is true iff v is true),

which, together with the disquotation schema (D), yields the more general schema

(DT) $(\text{Int}(u) \in {}^*p^*) \rightarrow (u$ is true iff $p)$.

Our grasp of truth for utterances is constituted by our inclination to accept instances of this schema. Moreover, these instances are the axioms of the minimal theory of that property.

The minimal theory of truth for utterances is equivalent—modulo two further principles—to the minimal theory for propositions. The auxiliary assumptions needed to derive either theory from the other are, first, a specification of the conditions for an arbitrary utterance, u, to *express* a given proposition,

(24) u expresses the proposition *that p* $\leftrightarrow \text{Int}(u) \in {}^*p^*$,

and, second, a specification of the relationship between truth for propositions and truth for utterances,

(25) u expresses the proposition *that p* $\rightarrow$ (u is true $\leftrightarrow$ the proposition *that p* is true).

These two schematic assumptions entail

(26) $\text{Int}(u) \in {}^*p^* \rightarrow$ (u is true $\leftrightarrow$ the proposition *that p* is true),

which allows us to derive the equivalence schema for propositions,

(E)   The proposition *that p* is true iff p,

from the disquotation schema (DT) for utterances, and also to derive (DT) from the equivalence schema. Thus, the accounts of truth for propositions and for utterances are unified and equally deflationary.

Notice, moreover, that a slight modification of the preceding

argument allows us to derive a minimal theory of truth for beliefs, claims, insinuations, etc. Instead of the preceding auxiliary principles governing the relationship between utterances and propositions, we invoke analogous principles relating utterances and propositional attitudes, namely,

(24*) u expresses the statement *that p* $\leftrightarrow$ Int$(u) \in *p*$

and

(25*) u expresses the statement *that p* $\rightarrow$ (the belief (claim, suggestion, . . .) *that p* is true $\leftrightarrow u$ is true).

These principles imply:

(26*)   Int$(u) \in *p* \rightarrow (u$ is true $\leftrightarrow$ the belief (etc.) *that p* is true).

And, as before, given the interpretation of u we can now go easily either from the disquotation principle (DT) to

(EB)   The belief (etc.) *that p* is true iff *p*,

or the other way round.

Ordinary language suggests that truth is a property of propositions, and that utterances, beliefs, assertions, etc., inherit their truth-like character from their relationship to propositions. However, the above derivations show that this way of seeing things has no particular explanatory merit. The truth-like conception for each type of entity is equally minimalistic. And by assuming any one of them we can easily derive the others.

In so far as our aim is merely to understand *our* conception of truth, and not to promote some allegedly better one, then I think we have no choice but to acknowledge that truth is primarily attributed to what we believe, question, suppose, etc. That is to say, truth is applied to the objects of our so-called propositional attitudes. However, there are a couple of influential sources of scepticism about this article of common sense—an article wholly embraced by minimalism—and it has been my purpose in this chapter to respond to them. First, many philosophers feel a certain queasiness about recognizing that propositions exist. Therefore I sketched the case in favour of them (based on the logical form of belief attributions) and tried to undermine the usual counterarguments (which derive from

their supposed weirdness and the alleged intransitivity of translation). In the second place, there is a tendency to think that our conception of 'proposition' *presupposes* the notion of truth, so that the minimalist order of explanation would seem to be the wrong way round. And to this I replied that we might employ an account of meaning—hence 'proposition'—in terms of *use*, which would not require a prior grasp of truth. This idea was already mentioned in the answer to Question 22. Finally, for those who are not convinced by these arguments, I showed that minimalist theories of 'truth' for utterances and belief-states can be given without making a commitment to propositions.

7

The 'Correspondence' Intuition

35. Is it not patently obvious that the truth or falsity of
a statement is something that grows out of its relations
to external aspects of reality?

This sentiment is what lies behind so-called 'correspondence theories
of truth'. According to one such theory (Wittgenstein, 1922) the
alleged correspondence is between representations and *facts*: the
truth of a sentence or proposition is said to derive from the existence
of whatever fact it '*depicts*'. Alternatively, there are accounts (e.g.
Tarski, 1958; Davidson, 1969) in which the ontological category of
fact is eschewed, and the truth of a whole sentence or proposition is
built directly out of the relations of reference and satisfaction
between its parts and various external objects. Here, the truth of a
thing is engendered by the objects to which its constituents corres-
pond.

Admittedly minimalism does not *explain what truth is* in any such
way. But it does not deny that truths *do* correspond—in *some*
sense—to the facts; it acknowledges that statements owe their truth
to the nature of reality; and it does not dispute the existence of rela-
tionships between truth, reference, and predicate satisfaction. Thus
we might hope to accommodate much of what the correspondence
theorist wishes to say without retreating an inch from our deflation-
ary position. Let us see how this can be done.

It is indeed undeniable that whenever a proposition or an utter-
ance is true, it is true *because* something in the world is a certain
way—something typically external to the proposition or utterance.
For example,

(1) ⟨Snow is white⟩'s being true *is explained by* snow's being
     white.

That is to say,

   (2)   ⟨Snow is white⟩ is true *because* snow is white.

But these intuitions are perfectly consistent with minimalism. In mapping out the relations of explanatory dependence between phenomena, we naturally and properly grant ultimate explanatory priority to such things as the basic laws of nature and the initial conditions of the universe. From these facts we attempt to deduce, and thereby explain, why, for example,

   (3)   Snow is white.

And only then, invoking the minimal theory, do we deduce, and thereby explain, why

   (4)   ⟨Snow is white⟩ is true.

Therefore, from the minimalistic point of view, (3) is indeed explanatorily prior to (4), and so (1) and (2) are fine. Thus we can be perfectly comfortable with the idea that truths are made true by elements of reality. Since this follows from the minimal theory (given certain further facts), it need not be an explicitly stated part of it.

36. Is it not equally clear that, contrary to minimalism, statements are made true by the existence of facts to which they correspond?

From the just-acknowledged explanatory relation

   (2)   ⟨Snow is white⟩ is true *because* snow is white.

it is not at all obvious that we can infer

   (5)   ⟨Snow is white⟩ is true *because* there exists the fact that snow is white.

After all, the two suggested explanations,

   (3)   Snow is white

and

   (6)   The fact that snow is white exists,

report different (though intimately related) states of affairs—and so something might well be due to one but not the other.

Moreover, there are positive grounds for resisting (5). For consider, to begin with, a Russellian proposition: namely, an arrangement of objects and properties.[1] It is hard to see how there could be any difference at all between such a proposition and the fact corresponding to it, since each of these entities would involve precisely the same arrangement of precisely the same objects and properties. Thus Russellian facts are just those Russellian propositions that are true. In other words, Russellian propositions, when true, are *identical* to facts, and not merely 'similar' to them; there is no difference between the truth of the proposition that snow is white and the existence of the fact that snow is white. Consequently there can be no explanatory relationship between these things, and they can correspond to one another only in so far as 'correspondence' is construed as identity.[2]

Turning to Fregean propositions (which are arrangements of the *senses* of words), there are two ways to go. One option is to maintain that they also, when true, are identical to facts. This idea may well come from the feeling that, for example, the fact that Phosphorus is Hesperus is *not* identical to the fact that Phosphorus is Phosphorus. Thus we might countenance both Fregean and Russellian facts— and it will be as hard to distinguish Fregean facts from Fregean propositions as it was to distinguish Russellian facts from Russellian propositions. Once again we would be excluding the possibility of a substantive correspondence theory.

Alternatively one might suppose that any Fregean proposition— and a sentence would presumably be treated in the same way—is made true by the existence of the corresponding *Russellian* fact. This, in effect, is Wittgenstein's (1922) 'picture theory'.[3] The rough idea here is that a sentence-token (or a Fregean proposition) consists of elements arranged in a certain logical form, and the fact that this is so *depicts* that there is in reality a Russellian fact consisting of the

---

[1] See the answer to Question 31 for a brief discussion of the distinction between Russellian and Fregean propositions.

[2] The identity of facts and true propositions was urged by Moore (1927) and has been recently defended by Jennifer Hornsby (1997).

[3] What follows is just *one* way of reading Wittgenstein's cryptic remarks. Arguably, the Tarskian approach, discussed in the answer to Question 38, is closer to what he had in mind.

referents of the elements arranged in the same logical form as in the sentence. The sentence is true if there is such a fact and false if there isn't. Formulated more precisely, the theory might look something like this:

(PT) I. A sequence of terms (or senses), S, *refers* to a sequence of entities, O, when S and O have the same length and the nth member of S refers to the nth member of O.

II. A sentence (or Fregean proposition), x, *corresponds* to a Russellian proposition, y, $\leftrightarrow$ there exists an S and an O such that
—x is composed of the members of S;
—y is composed of the members of O;
—x and y have the same logical form;
—S refers to O.

III. y = the Russellian proposition that p →
(y is a *fact* $\leftrightarrow p$).

IV. x is *true* $\leftrightarrow$ there exists a Russellian proposition, y, such that
—x corresponds to y;
—y is a fact.

But although these principles are quite plausible and certainly worth noting, they do not make a good alternative to the minimal theory. The single respect in which the minimal theory can seem unattractive is its infinite, list-like character. Sellars (1962) compares it to a telephone directory. And precisely this feature is preserved in the picture theory's principle III, which specifies, for each Russellian proposition, the circumstances in which it will qualify as a fact. As we saw in the answer to Question 5, there can be no hope of encapsulating this schema within a finite body of axioms. In addition, a theory of reference is required; and, as we also saw in the answer to Question 5, there is reason to think that this too would require infinitely many axioms. And we have not yet mentioned the need for theories of what it is for a sentence and a proposition to have the same logical form, and for a proposition to be composed of a certain sequence of entities. Thus if we were to trade minimalism for the picture theory we would sacrifice purity and simplicity for absolutely no benefit. This implies that, although its principles may well be correct,

the picture theory of truth should not be thought to qualify as our *basic* theory of truth.

Thus we have failed to find any correspondence account of truth possessing methodological virtues lacked by the minimal theory. No doubt one may formulate some interesting, plausible schemata that relate the concepts of truth, fact, and correspondence. But the conjunction of such schematic principles is best viewed as yielding a legitimate extension of our theory of truth; it does not provide a tempting alternative.

37. Certain cases of representation (e.g. by maps) clearly involve a correspondence—a structural resemblance— to what is represented. So is it not reasonable to expect some such relation in linguistic representation also?

Perhaps so. However it remains to be shown, for either type of representation, that such structural similarities are in any way *constitutive* of truth. The minimalist view is that they are not.

The difference between a map (or any so-called 'realistic' representation) and a sentence of a natural language is that the interpretation of maps is more 'natural'—or in other words, less 'conventional'—than the interpretation of sentences. Let me explain what I mean by this. In the case of certain maps, a representation that consists in some set of objects (symbols) standing in some relation to one another is supposed to be interpreted as saying that the referents of those objects stand in that very same relation. For example,

(7)   The fact that *point y* is on a straight line between *points x* and *z*

expresses

(8)   The fact that *the place represented by y* is on a straight line between *the places represented by x* and *z*.

Thus what is special about 'realistic', 'natural', 'non-conventional' systems of representation is that they have syntactic features (e.g. one thing being on a line between two other things) that refer to

themselves. In other words, facts (7) and (8) share, not merely logical form, but also a certain pictorial form: namely,

(9)   That . . . is on a straight line between . . . and . . ..

This suggests the following account of pictorial truth:

(P) x is a true picture $\leftrightarrow$ the referents of x's pictorial elements exemplify the very relation exemplified by those elements.

But the then needed accounts of *exemplification* and *reference* will be far from unproblematic. For the former would involve the schema

(Ex*)   The sequence of objects, $\langle a_1, a_2, \ldots, a_n \rangle$, exemplifies the relation R-ness iff $R(a_1, a_2, \ldots, a_n)$.

And, in so far as there is no limit to the number of possible pictorial elements, the latter would also call for infinitely many axioms. Thus the proposed theory of pictorial truth is not especially attractive in comparison with minimalism.

An alternative strategy would be to construct a theory of truth for pictorial representations that would parallel the theory of truth for sentences, (PT), given above—the only difference being that in principle PT-II, we would speak of pictorial forms instead of logical forms. However, although this is indeed possible, the result would be just as cumbersome as (PT) turned out to be, and just as unsatisfactory compared to the minimalist theory.

So, even for pictorial representations, the best theory of truth is the minimal theory. This assumes, of course, that the minimal theory can be applied to pictures; and let us quickly confirm that this is so. As we saw in the answer to Question 24, the minimalist axioms of the theory of truth for *expressions* (as opposed to propositions) are instances of

(DT) $\text{Int}(u) \in {}^*p^* \rightarrow (u$ is true iff $p)$.

Therefore the axiom-schema concerning the map

(M)

$$\boxed{\text{A}\!-\!\!-\!\!-\!\!-\!\!-\!\!-\!\text{B}\!-\!\!-\!\!-\!\!-\!\!-\!\!-\!\text{C}}$$
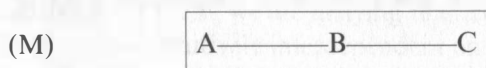
is

(10) $\text{Int}(\text{M}) \in {}^*p^* \rightarrow \text{M}$ is true iff p.

But in fact the English interpretation of M happens to be the sentence

(11)   B is between A and C.

Consequently

(12)   M is true iff B is between A and C.

Thus we are able to articulate the difference and similarity between linguistic and pictorial representation, and show that in neither case is there any reason to go beyond the minimal theory of truth.

38. The minimal theory fails to show how the truth of a sentence depends on the referential properties of its parts.

Another substantial account of truth that is sometimes regarded as an elaboration of the 'correspondence theory', is the Tarski-inspired compositional approach.[4] Tarski's account resembles the picture theory in that the truth of a sentence is explained in terms of the semantic properties of its constituents. But it differs from the picture theory in not explicitly referring to facts or to logical structures. Most importantly, it does not call for a schema, such as PT-III, to specify for each proposition when it is a fact; instead it offers a non-schematic, finite theory by deducing the truth condition of each statement from the referents of its parts. Thus, roughly speaking and simplifying enormously, we get:

   (TT) u is true in language L iff

      (i) u has the form '$\$(x)$' and there is an object, x, such that the predicate of u is true-in-L of x and the subject of u refers-in-L to x;

   or (ii) u has the form 'p & q' and each conjunct of u is true-in-L;

---

[4] It is unclear that Tarski himself would endorse this construal of his work. Davidson (1969), however, has argued that it should be seen as vindicating the correspondence intuition. Note also that Tarski's own initial theory of truth (Tarski, 1958) was not given for languages containing names, and does not reduce truth to reference but solely to predicate-satisfaction. Nevertheless, the following points apply to it.

or   (iii) *u* has the form '*–p*' and what is negated in *u* is not true;
or   (iv) *u* = . . .;

which can ultimately be turned into an explicit definition of truth.

Tarski himself was content to show that for certain simple languages, *L*, he could define a predicate, 'true-in-*L*', whose extension would be the truths of *L*. But someone (e.g. Davidson, 1969) might hope to go further and arrive at a general theory of truth by showing, even for complex, natural language such as English, that 'true-in-*L*' could be defined in terms of 'refers-in-*L*' and 'satisfies-in-*L*'.

However, there are two basic objections to any such approach. In the first place, as Davidson concedes, the Tarskian strategy applies only to those sentences whose logical forms may be represented in first order logic. Or, in other words, it applies only to those sentences whose truth values are determined by the truth values of atomic sentences. But there is no reason to assume that every conceivable truth has such a structure. Consider, for example, counterfactual conditionals, probability claims, laws of nature, and modal assertions of various kinds. All of these constructions resist formalization in the language of predicate logic, and so it is not clear that Tarski's theory can be made to cover them.

In the second place, as we saw in the answer to Question 5, such a theory of truth will be adequate only if it is supplemented with theories of reference and satisfaction; and these require either substitutional quantification, or an unformulable collection of axioms. So no simplification is achieved by reducing truth to reference and satisfaction. On the contrary, the extreme conceptual purity and simplicity of minimalism is thrown away for nothing. The correspondence approach promised to rectify a certain perceived defect of minimalism: namely its infinite, list-like character—its failure to say what all truths have in common. But it turns out that some of the terms to which truth is to be reduced are similarly intractable; and so the initial promise of correspondence is not fulfilled.

In rejecting the idea that the correspondence theory describes the basic nature of truth, we are denying that truth, reference, and satisfaction are constitutively interdependent on one another. But we are certainly not trying to suggest that the theoretical principles relating those properties are incorrect. We are supposing, rather, that such principles should not be treated as explanatorily basic, but should each be explained in terms of simple, separate, minimal theories of

truth, reference, and satisfaction. And in fact this can be done very easily. For example, it is not a part of our basic theory of truth that we should accept instances of

> (13)   'p and q' is true iff 'p' is true and 'q' is true.

However, this is trivially deducible from

> (i)   'p and q' is true iff p and q,
> (ii) 'p' is true iff p,

and

> (iii) 'q' is true iff q.

Similarly, we do not define truth in terms of reference and satisfaction using such principles as

> (T/R)   'Fa' is true iff $(\exists x)$('a' refers to x & 'F' is satisfied by x).

However, this schema is easily explained. For we have

> (S') $(x)(x$ satisfies 'F' iff $Fx)$

and

> (R') (x)('a' refers to x iff $a = x)$.

Therefore the right-hand side of (T/R) is equivalent to

> (14) Fa,

which, given the disquotational principle, is equivalent to the left-hand side of (T/R),

> (15)   'Fa' is true.

Thus, the minimal theory does not preclude the possibility of showing how the truth value of a sentence is related to the referential properties of its parts.

The best theory of truth will be the smallest, simplest collection of statements that, in conjunction with theories of other matters such as reference, will enable us to derive everything we believe about truth. This desideratum points us towards minimalism and away from theories, such as Tarski's, which are unnecessarily complex, explanatorily misguided, and foster the misleading impression that truth, reference, and satisfaction are inextricably intertwined with one another.

39. The great virtue of defining truth in terms of reference
is that the account may be supplemented with a naturalistic
(causal) theory of the reference relation to yield, in the end,
a naturalistic and scientifically respectable theory of truth.
(Field, 1972; Devitt, 1984)

The usual line on singular reference goes roughly as follows. Once
upon a time we believed the description theory of Frege (1892) and
Russell (1905). It was held that every time a name is used it is
intended to be synonymous with some definite description. Thus
'Socrates' may, in some idiolect, mean 'the teacher of Plato'. This
idea was refined by Wittgenstein (1953) and Searle (1958), who
argued that a name, though perhaps not synonymous with any par-
ticular definite description, is none the less associated with a certain
cluster of definite descriptions which jointly determine what it
stands for. But, so the story goes, this whole picture of reference has
been decisively refuted by Kripke (1972). Suppose, to give one of his
examples, an obscure mathematician, Schmidt, was really the person
who discovered the incompleteness of arithmetic, although Gödel
has been unfairly given the credit. In that case, even if someone's
only belief expressed by means of the name 'Gödel' were that Gödel
discovered the incompleteness of arithmetic, we would none the less
say that his uses of the name do not refer to Schmidt, but to Gödel.
Kripke used such counterexamples to the description theory in order
to motivate an alternative 'causal' picture of the reference relation.
The idea here is that there is a certain type of causal chain connect-
ing the uses of a name and the thing it stands for. What remains,
according to this line of thought, is to formulate a definite theory of
reference that would manage to capture this causal picture—but
unfortunately no-one has been able to do it yet.

From the minimalist perspective this failure comes as no surprise.
For truth and reference are so intimately related that the rationale
for a minimal account of truth will equally well motivate a minimal
account of reference. Reference, on this view, is not a complex rela-
tion; a naturalistic or conceptual reduction is not needed and should
not be expected.

Notice, moreover, that Kripke's own model contains nothing to
preclude some role for *representation* in the fixing of a name's refer-
ent; and the introduction of names by means of initial modes of

presentation (demonstrative or descriptive) is not given a non-semantic analysis. His central claim is merely that there need be no reference-fixing characterization in the minds of users of the name. For as use of the name spreads 'causally' through a linguistic community, the initial modes of presentation might well be lost. But this is merely an instance of a quite general point about meaning, applying to all words, and not just to names. The point is that the meaning of an expression is not intrinsic to the minds of individual language users, but resides in the practices of the linguistic community as a whole and is properly attributed to an individual on the basis of his interactive relations with the community (see Burge, 1979). Thus what is 'causal' in Kripke's theory has nothing specifically to do with names, and the term 'causal theory of names' is quite misleading in connection with his discussion. One shouldn't think that he gives even so much as a *crude* version of a causal theory of reference.

Such a conception becomes even less plausible when it is combined, as it often is, with the thesis that names have no meaning (i.e. no Fregean sense) and that their only possible contribution to the propositions they help to express is their referent. As we saw in the answer to Question 31, this position makes it impossible to give a satisfactory account of *de dicto* propositional attitudes (which at least the description theory was able to do). Moreover, it appears to be the product of an overreaction to Kripke's arguments. Granted they show that names do not have the same meanings as definite descriptions; but this provides no reason to conclude that names have no meanings at all. We might suppose, rather, that they are *primitive* expressions. And there would be no mystery about this. After all, most predicates are primitive; so why shouldn't names be as well?

As in the case of singular terms, the search for a causal theory of predicate satisfaction (for example, by Stampe (1977), Dretske (1981), and Fodor (1987)) is conducted as if it were perfectly obvious that there is such a thing and all that remains is to work out the details. But this obviousness, this sense that from a naturalistic, scientifically respectable viewpoint such an account *must* be right, has its origin in the same linguistic illusion that motivates substantive theories of truth. Therefore the continual failure of its advocates to produce an adequate theory is only to be expected.[5]

---

[5] See T. Blackburn (1988) for a discussion of the failed attempts. As he says, 'it is

Paralleling our approach to truth, minimalist accounts of satisfaction and reference would begin with a story about the *function* of those notions. Roughly and briefly, one might suppose that just as truth helps us to talk about propositions that are *wholly* unarticulated, so satisfaction and reference help us to talk about propositions of which certain *parts*—those corresponding, respectively, to predicates and singular terms—are unarticulated. Just as truth performs its function by providing a sentence of *our* language,

(16)   '*X*' is true,

which is equivalent to the utterance

(17)   *X*,

which might be in *any* language, so satisfaction and reference do their jobs because

(18)   satisfies '*F*'

and

(19)   the referent of '*D*'

provide *home language* equivalents of any predicate, '*F*', or singular term, '*D*', whatever their language, and whether or not their translations are available. Moreover, just as truth is a convenient alternative to substitutional quantification into sentence positions, so satisfaction and reference enable us to do without substitutional quantification into predicate and singular term positions. Thus instead of

(20) $\{\#\}(x)(\#x \lor -\#x)$

one can say

(20*)   Given any predicate and any object, either the object
        satisfies the predicate or it doesn't,

and instead of

(21)   {∃d}(Raphael's belief contains ⟨d⟩ and d = the moon)

one can say

(21*)   The referent of a constituent of Raphael's thought is the moon,

or

(21**)   Raphael is thinking about the moon.

Having characterized the non-descriptive role of 'satisfies' and 'refers', the second stage of a minimalist account of satisfaction and reference is to specify the theories that would suffice for the performance of these functions. Paralleling the theory of truth, these would be, respectively, every proposition of the form

(S) $(x)(x$ satisfies ⟨F⟩ ↔ Fx),

and every proposition of the form

(R) $(x)(⟨d⟩$ refers to x ↔ $d = x$).

Given such an explanation of the behaviour and purpose of 'refers' and 'satisfies', it would be a very surprising coincidence if there were any unified conceptual or naturalistic reduction of the reference relation.[6]

The correspondence conception of truth involves two claims: (1) that truths correspond to reality; and (2) that such correspondence is what truth essentially is. And the minimalist response, urged in this chapter, is to concede the first of these theses (properly understood) but to deny the second. The rationale for this response is that the minimalistic equivalence biconditionals can easily be supplemented with characterizations of *correspondence* and *fact* to show that, indeed, for any true proposition or sentence, there is a corresponding fact. However, we have also seen that there are no advantages—and substantial disadvantages—in supposing that this entire construction constitutes the basic theory of truth. One merely imag-

---

[6] Arguments in favour of the conclusion that these minimalist axioms of reference and satisfaction are indeed explanatorily basic may be constructed in parallel to the arguments, given in the answer to Question 14, that MT is explanatorily basic. For a more extensive deflationary treatment of reference and names see my *Meaning* (1998), ch. 5.

inary benefit is that a correspondence account, by reducing truth to reference and reference to causation, would leave us with a finite, naturalistic model—and thereby make the concept of truth scientifically respectable. I argued, however, that *reference, satisfaction, correspondence*, and *fact* are just as non-naturalistic, and in need of infinite, deflationary theories, as truth is. Moreover, such theories are *already* perfectly legitimate from a scientific point of view, without any additional naturalistic reduction. Therefore the extra notions employed in the correspondence theory do not earn their keep, and merely introduce unnecessary complexity. Ordinary canons of explanatory priority dictate that the equivalence biconditionals be taken as fundamental, and that further characteristics of truth, including its correspondence properties, be accounted for on the basis of that theory.

Conclusion

Although the minimalist point of view has been subjected to many objections, we have seen that it is by no means incapable of dealing with them. Indeed, from this trial it emerges even stronger than before. For we now have a clear formulation of the doctrine, a coherent set of replies to a very broad range of criticisms, a picture of the alternative accounts and what is wrong with them, and a sense of the theory's significant philosophical implications.

Let me end by summarizing the line of thought that has been pursued here. We began with just about the only uncontroversial fact to be found in this domain: namely, that the proposition *that snow is white* is true iff snow is white, the proposition *that tachyons exist* is true iff tachyons exist, and so on. We then posed, and attempted to answer, two questions about this general fact. First, what precisely is it? Can we provide a clear and logically respectable characterization of it—one which does not rely on improper locutions, like 'and so on'? Second, is there any further, deeper, non-trivial theory of truth—some account going beneath or beyond instances of the equivalence schema? With respect to the descriptive question I argued that a theory containing all the equivalence theses cannot be formulated—unless notions are employed that themselves require unformulable theories. This is because there are too many such theses, and also because some of them concern propositions that we cannot yet express. On the question of whether any further theory of truth remains to be found my answer was a categorical No. Hence the name 'minimalism'. The justification for this answer fell into a couple of parts. I argued, in the first place, that all uses of the truth predicate are explained by the hypothesis that its entire *raison d'être* is to help us say things about unarticulated propositions, and in particular to express generalizations about them. It transpired in the course of the book that our apparently deep uses of truth in logic, semantics, and

the philosophy of science are simply displays of this role. Secondly, I showed that the performance of this function requires nothing more or less than the truth of the equivalence axioms. Thus minimalism perfectly explains all the pertinent facts; and that is its justification. As for the philosophical import of minimalism, this should no longer be in doubt. We have seen that many controversies—regarding, for example, scientific realism, meaning, vagueness, normative emotivism, and the foundations of logic—are standardly assumed to interact essentially with the nature of truth. To the extent that the notion of truth is clarified and its *independence* of these problems established, they can be certain to receive clearer formulation and be more amenable to resolution.

Postscript

In the years since the first edition of this book was published[1] several critical reactions to it have appeared. As far as I can tell, none of these discussions succeeds in undermining the minimalist perspective. Some of them show, however, that various points were poorly formulated and stand in need of clarification, and some raise new objections which should be addressed. I would like to take this opportunity to deal with the most serious of these issues. I will begin by repeating what I regard as the essence of the minimalist conception of truth, and then go on to discuss some of the recent objections to it. I shall focus particularly on questions raised by Anil Gupta, Hartry Field, Mark Richard, Donald Davidson, Crispin Wright, Scott Soames, Michael Dummett, Paul Boghossian, and Michael Devitt.

1. What is minimalism?

The deflationary attitude toward truth—and the particular variant of it that I call 'minimalism'—are a reaction against the natural and widespread idea that the property of truth has some sort of underlying nature and that our problem as philosophers is to say what that nature is, to analyse truth either conceptually or substantively, to specify, at least roughly, the conditions necessary and sufficient for something to be true. Amongst the products of this traditional point of view there is the correspondence theory (x is true iff x corresponds to a fact), the coherence theory (x is true iff x is a member of a coherent set of beliefs), the verificationist theory (x is true iff x is

---

[1] *Truth* (Oxford: Blackwell, 1990).

provable, or verifiable in ideal conditions), and the pragmatist theory (*x* is true iff *x* is useful to believe). But nothing of this sort has ever survived serious scrutiny—which comes as no surprise to the deflationist, who denies that there is any prospect of an explicit definition or reductive analysis of truth, even a very approximate one.[2]

Deflationism begins by emphasizing the fact that no matter what theory of truth we might espouse professionally, we are all prepared to infer

> The belief *that snow is white* is true

from

> Snow is white

and vice versa. And, more generally, we all accept instances of the 'truth schemata'

> The belief (conjecture, assertion, supposition, . . .) *that p* is
>   true iff *p*.

But instead of taking the traditional view that an *analysis* of truth still needs to be given—a reductive account, deeper than the truth schemata, which will explain why we accept their instances—the deflationist maintains that, since our commitment to these schemata accounts for everything we do with the truth predicate, we can suppose that they implicitly define it. Our brute acceptance of their instances constitutes our grasp of the notion of truth. No conceptual analysis is called for—no definition of the form

> 'true' means '*F*',

where '*F*' is some expression composed of terms that are more basic than the truth predicate. Moreover there is going to be no non-definitional analysis of truth either, however rough and ready—no substantive discovery of the form

> *x*'s being true *consists in x*'s having property *F*.

---

[2] For a discussion of what is right and wrong about these traditional analyses of truth, see Chapter 1, and Kirkham, 1992. It may be felt that since almost no concepts are susceptible to exact reductive analysis (not even 'table' or 'house'), it cannot be of much interest to deny that truth is. The alleged peculiarity of truth, however, is that there is nothing to be said—*not even very roughly speaking*—about what it consists in.

Hence the term 'deflationism'.

This, of course, is highly reminiscent of the old 'redundancy theory' of Frege (1891), Ramsey (1927, 1991), Ayer (1935), and Strawson (1964): the idea that

The proposition *that p* is true

means no more and no less than simply

p.

And the redundancy theory is indeed an early form of deflationism. However, there are various respects in which we have been able to improve upon this original articulation of the view—leading to the formulation of deflationism presented and defended in the preceding chapters, namely, *minimalism*.

In the first place, as the name given to their doctrine suggests, the redundancy theorists had nothing much to say about the *function* of our concept of truth. But if it really is redundant, why on earth do we have such a notion? A virtue of minimalism is that it contains a satisfying response to this question—one that was first proposed by Quine (1970)[3]—namely, that the truth predicate plays a vital role in enabling us to capture certain generalizations. For we can generalize

The moon is subject to gravity

by saying

Every physical object is subject to gravity.

And similarly we can always obtain a generalization from a statement about a particular object by first selecting some kind or type *G* (e.g. 'physical object'), and then replacing the term referring to the object with the universal quantifier 'Every *G*'. However there is an important class of generalization that cannot be constructed in this way: for example, the one whose instances include

If Florence is smiling, then Florence is smiling.

How can we extract the law of logic it instantiates? Or consider

---

[3] The redundancy theorists appreciated the need for some account of the function of the truth predicate. They tended to maintain that it serves various merely *pragmatic* purposes (e.g. to emphasize, as in 'It's true that I didn't see anything'; or to concede, as in 'It's true that they are a lousy team; but they will win anyway'). But the Quinean account is clearly superior.

Physicists would like to *believe* that there are black holes only
if there *are* black holes.

How are we to generalize this to obtain the epistemological policy it
exemplifies? The solution provided by our concept of truth is to
convert each such proposition into an obviously equivalent one—
but one that *can* be generalized in the normal way. Thus, given the
equivalence of

p

and

The statement *that p* is true

and

The belief *that p* is true,

we get, respectively

The statement *that if Florence is smiling then Florence is
smiling* is true

and

Physicists would like to believe *that there are black holes* only
if the belief *that there are black holes* is true,

each of which generalizes in the standard fashion, yielding

Every statement of the form 'If p, then p' is true[4]

and

Physicists would like to believe only what is true.

From these we can derive (given the truth schemata) all the state-
ments we initially wished to generalize—and nothing logically
weaker would suffice. They may therefore be regarded as generaliza-
tions of the initial statements. Thus it is indeed useful to have a term
that is governed by the truth schemata—despite their triviality.
There is a clear *raison d'être* for a concept having precisely the char-
acteristics that the minimalist attributes to truth.

---

[4] 'Every statement of the form " . . . "' can be read as 'Every statement expressed
by a sentence of the form " . . . "'.

A second and related respect in which minimalism improves on the original redundancy theory is in not claiming that '*p*' and 'The statement (belief, . . .) *that p* is true' have exactly the same meaning. This claim is implausibly strong; for after all, the words 'true' and 'statement' do have meanings, and those meanings would appear to be, in some sense, 'components' of the meaning of 'The statement *that p* is true' but not of '*p*'. Moreover, there is no need to take such an extreme position. The generalizing function of truth is perfectly well fulfilled as long as instances of the truth schemata, understood as merely *material* biconditionals, are accepted. Relative to that body of assumptions, any generalization of the form 'All instances of schema *S* are true' will entail all the instances of *S*—which are precisely the statements we needed to generalize.

Anil Gupta (1993*a*) rightly notes that the instances of the generalizations that we use the concept of truth to formulate (e.g. instances such as 'The statement *that if Florence is smiling then Florence is smiling* is true') will not say exactly the same thing as what we wished to generalize ('If Florence is smiling then Florence is smiling') unless corresponding instances of 'The statement *that p* is true' and '*p*' express the very same proposition—which, as we have just conceded, is not especially plausible. But this point does not undermine the minimalist story about the function of truth; for, as just mentioned, that function requires merely that the generalizations permit us to *derive* the statements to be generalized—which requires merely that the truth schemata provide material equivalences. This isn't to deny that the instances so understood are not only true but necessarily true (and a priori). The point is that their mere truth is enough to account for the generalizing function of truth.

It was perhaps an exaggeration to have suggested that the concept of truth is *needed* for this generalizing purpose. An alternative strategy would be to introduce some form of non-standard (e.g. substitutional) quantification, by means of which we could say

$(p)(\text{If } p \text{ then } p)$

and

Physicists would like it that
$(p)(\text{they believe that } p, \text{ only if } p).$

But in that case there would be required a battery of extra syntactic and semantic rules to govern the new type of quantifier. Therefore,

we might consider the value of our concept of truth to be that it provides, not the only way, but a relatively 'cheap' way of obtaining the problematic generalizations—the way actually chosen in natural language.

It is emphasized by some philosophers that the truth predicate is a device for forming *prosentences*: just as one might use the pronoun 'he' instead of repeating a name (as in 'John said he was happy'), so one might say 'That's true' instead of repeating the sentence just asserted. Evidently this is a perfectly correct observation as far as it goes. However, there remains the issue of how best to explain it; and here is where minimalists such as myself part company with self-styled 'prosententialists' such as Dorothy Grover, Joseph Camp, and Nuel Belnap (1975), and Robert Brandom (1994). The latter tend to suppose that 'is true' should be analysed using substitutional quantification—roughly speaking, that 'x is true' means '$(p)(x = \langle p \rangle \rightarrow p)$'. In contrast, the former contend that the overall use of the truth predicate (including its use as a prosentence forming device) is best explained by supposing that 'is true' is an unanalysable predicate governed by the equivalence schema; and that the virtue of such a predicate is that it allows us to avoid the complexities and obscurities of substitutional quantification.

A third defect of the original redundancy theory lies in its implication that truth is *not a property*—or, in other words, that attributions of truth (as in 'The proposition *that dogs bark* is true') do not have the logical form, 'x is F', that is characteristic of attributions of properties to objects. A better form of deflationism will reject this thesis as inconsistent with the logical role of the truth predicate—in particular with the inferences we make from '$x = that\ p$' and 'x is true' to '*that p* is true', and hence to 'p'. No doubt truth is very different from most properties in so far as it has no underlying nature; but, in light of the inferential role of 'true' as a logical predicate, it is nonetheless a 'property', at least in some sense of the term. This is a point to which I return in section 8.

My defence of the minimalist perspective will proceed as follows. In the next section I consider various criticisms of the idea that the truth schemata provide an implicit definition of the truth predicate. Then, in sections 3 and 4, I take up the question of what sort of entity we apply it to, and restate the case in favour of propositions. Turning, in section 5, from the question of how the word 'true' means what it does to the question of what explains the

characteristics of truth, I reinforce the claim that truth has no under-
lying nature and that the basic theory of that property consists in
instances of the equivalence schema, 'The proposition *that p* is true
iff *p*'. In sections 6 and 7 I address the objection that the *normative*
status of truth implies that there must be something more to it than
is appreciated by minimalists. And I conclude by looking back at the
issue of whether or not truth is, in some sense, a *property*.

2. Does minimalism give an adequate account of our concept of truth?

A central component of the minimalistic picture is its claim about
the particular non-semantic fact that provides the word 'true' with
its peculiar meaning. This is said to be the fact that a certain basic
regularity governs the overall use of the truth predicate—that a cer-
tain pattern of behaviour provides the best explanation of our total
linguistic practice with the term. Now, as we have just seen, the reg-
ularity which plays that explanatory role is the regularity in virtue of
which the term is useful: namely, our disposition to accept instances
of the truth schemata. Therefore, this is the fact about the truth
predicate that constitutes our meaning by it what we do.[5] More
specifically, the meaning-constituting fact is this: that the explanato-
rily basic fact about our use of the truth predicate is our tendency to
infer instances of 'The proposition *that p* is true' from corresponding
instances of '*p*', and vice versa, whenever (a) each '*p*' is replaced with
tokens of an English sentence, (b) these tokens are given the same
interpretation as one another, (c) under that interpretation they
express the content of a statement (a proposition), and (d) the terms
'that' and 'proposition' are given their English meanings.

It might be objected that this condition is not strong enough to
distinguish truth from various other concepts, since there are innu-

---

[5] This claim about the meaning of 'true' is the product of two distinct ideas: first,
the fundamental thesis of minimalism with respect to truth (namely, that the truth
schemata are the explanatory basis of our overall use of the truth predicate); and sec-
ond, the use theory of meaning (namely, that a word means what it does in virtue of
the basic regularity governing its deployment). Many of the radical consequences of
minimalism depend only on its basic tenet (that the schemata are conceptually basic)
and do not depend upon the further idea that they constitute our concept of truth.

merable predicates, '*F*', other than 'true', for which we would accept instantiations of (F*):

(F*)   The proposition *that p* is *F* iff *p*.

Imagine, for example, that '*F*' is the predicate 'is true and 1 + 1 = 2'. However, this objection overlooks the fact that, although there are many predicative expressions that obviously satisfy schema (F*), it is not generally true, for each such expression, that this is the *full* basic regularity governing its use. For example, our knowledge that 'is true and 1 + 1 = 2' satisfies schema (F*) does not explain our inclination to deduce, from the premise that this predicate applies to an object, the conclusion that 1 + 1 = 2. And, in general, if a complex predicate satisfies schema (F*), then it will exhibit various use relations to its constituents; and these facts will not be explained by our reliance on the schema. Thus, in the case of the truth predicate, the schema provides the *complete* basic regularity for its use. In contrast, the predicates other than 'true' that satisfy the schema are governed by additional regularities. And this is what distinguishes their meanings from that of 'true'.[6]

A separate objection, made by Hartry Field (1992),[7] is that the minimalist account could capture the meaning of 'true' only in its application to statements that we are able to *formulate*; for only in those cases can we supply the relevant instance of the truth schemata. The idea is that if we take a statement, *s*, to say, for example, *that snow is white*, we might then suppose that the content of

   s is true

is more or less

   snow is white,

whereas, if we don't know what *s* means, then we don't know the content of '*s* is true'. Now this would indeed count against a theory that aimed, by means of the truth schemata, to provide a *reductive analysis* of each utterance containing the word 'true'. However one need not, and should not, promise any such reductive analysis;

---

[6] This potential difficulty was brought to my attention by Anil Gupta in October 1992.

[7] Although Field is critical of my particular brand of deflationism, he is one of the strongest advocates of the deflationary point of view. See, for example, Field, 1994.

indeed it is a central tenet of minimalism that there is no such thing. On the minimalist view, we aim to define the truth predicate, not by providing another expression with the same meaning (nor even by providing a rule transforming every sentence containing the word 'true' into a content-equivalent sentence without it), but rather by specifying which property of the truth predicate constitutes its having the meaning it has; and to that end we must identify the property that best explains our overall use of the term. In particular, the minimalist thesis is that the meaning of 'true' is constituted by our disposition to accept those instances of the truth schemata that we *can* formulate. In that way, the word is provided with a constant meaning wherever it appears—even when ascribed to untranslatable statements. And the justification for this thesis is that such a pattern of acceptance accounts for our entire use of the term—including its application to untranslatable utterances. Thus we might attribute truth to an untranslatable statement on the basis, for example, of a belief in the reliability of the person who made it; and we do, as Field suggests, need an account of truth that will explain this sort of attribution; but the minimalist proposal would do so perfectly well.

Mark Richard (1997) has noted that some instances of the truth schemata will be found, by some people, to be less obvious than other instances. In particular, propositions involving empty names ('Atlantis is in the Atlantic'), vague attributions ('Smith is bald'), ethical pronouncements ('Killing is wrong'), and future contingents ('The sea-battle will occur tomorrow') may be affirmed, yet none the less be regarded as 'non-factual' hence incapable of truth or falsity. In which case '*p*' is not taken to imply the corresponding 'The statement (belief, conjecture, . . .) *that p* is true'. Consequently, argues Richard, minimalism is mistaken in maintaining that the concept of truth is fixed by our disposition to accept instances of the truth schemata.

However, the essence of minimalism is that certain uncontroversial schemata suffice to fix our concept of truth and that there is no deeper definition taking the traditional, explicit form. Therefore even if, in light of the above examples, we were to reformulate minimalism using the qualified schemata

> '*p*' has 'factual content' →
> (The proposition *that p* is true iff *p*)

(where the notion of 'factual content' is explicated non-truth-

theoretically, in terms of 'use' or 'conceptual role'), we would leave intact the core of the minimalistic perspective.

Perhaps some such qualified schemata do give a better account of certain philosophers' concept of truth; but for the rest of us the unqualified schemata are accurate. Thus one might be inclined to suppose that there are *two* concepts of truth: that the original truth schemata provide a minimalist account of one of them, and the qualified schemata a minimalist account of the other. However, the truth predicate is needed as a device of generalization in *all* domains of discourse, including those in which certain philosophers are wary of attributing truth. Thus the original, unqualified schemata are what capture the concept we need. It seems reasonable, therefore, to blame any anxiety regarding those schemata on philosophical confusion rather than on deployment of a more restricted concept. Thus, despite the intuitions cited by Richard, I believe we can stand by the view that the meaning-constituting use of 'true'—the 'expert' use—is the explanatory role of our disposition to accept instances of the unqualified truth schemata.

3. What are the bearers of truth?

In ordinary language what are said to be true are the *things* that we believe and that our utterances express—so-called *propositions*. Thus, on the face of it, propositions exist, some of them (presumably, half) are true, and the correlated truth-like attributes of utterances and of acts of believing, asserting, etc. are the complex, derivative properties, '*u* expresses a true proposition' and 'the object of act *x* is a true proposition'.

However, it remains to be seen (1) whether this first impression is correct—specifically, whether there really are such things as propositions; (2) whether, if correct, this naive picture is consistent with minimalism—specifically, with minimalism's requirement that propositions be conceptually prior to truth; and (3) which of the three truth concepts is really fundamental—specifically, whether it is not explanatorily preferable to go against the current of ordinary language and to explicate propositional truth and act truth in terms of a previously developed account of utterance truth.

On the basis of the inferential behaviour of that-clause

(propositional attitude) constructions (which are sentences like 'John believes that God exists', 'Mary said that lying is wrong', and 'Charles imagines that he is Napoleon'), a good case can be made for concluding that they articulate *relations* (such as believing, asserting, and hoping) between people and whatever are designated by the constituent that-clauses—i.e. propositions.[8] Assuming, as I shall, that this conclusion is right, there none the less might be no *nominalistic* answer to the question 'But what are propositions?'—no theory specifying, for example, some species of physical or mental entity with which propositions are to be identified. An inability to give such an answer does not justify scepticism about the existence of propositions, for the nominalistic presupposition of any such sceptical argument can easily be resisted. It may be supposed, rather, that the conception we have of any given proposition lies in the circumstances in which we would take it to be expressed. In other words, it is by making explicit the grounds on which we maintain such things as 'John's utterance, *u*, expresses the proposition that dogs bark' that we will articulate our concept of proposition.

The minimalist view of truth has two important consequences regarding the grounds for such interpretive claims. One is that they not involve truth-theoretic notions (i.e. 'true', 'is true of' or 'refers'). For minimalism is primarily the view that the equivalence schema is conceptually basic vis-à-vis the truth predicate—and, analogously, that parallel schemata are conceptually basic vis-à-vis 'is true of' and 'refers'. This implies that the various truth-theoretic concepts be posterior to the concept of proposition, and therefore that it be possible to possess the concept of proposition without possessing the concepts of truth, being true of, and reference. And this implies that that-clause attributions not rest on grounds involving truth-theoretic notions. For example, minimalism implies that nothing along the lines of

$$u \text{ expresses the proposition } that \ p = u \text{ is true iff } p$$

could be what articulates our conception of proposition.

---

[8] I have in mind the need to accommodate such inferences as: 'Oscar believes that dogs bark and Barny denies that dogs bark; therefore there is something which Oscar believes and Barny denies.' This sort of argument in favour of propositions is developed in Chapter 6 and in Schiffer, 1994. For the sake of simplicity of exposition I am ignoring the various other ways of referring to propositions (e.g. 'John *wonders whether May loves him*') and am writing as though propositions were invariably designated by that-clauses.

The second important consequence of minimalism is that the grounds for that-clause attributions need not even involve anything that might *constitute* the truth-theoretic properties. In other words, it is not merely that the evidence for that-clause attributions does not explicitly advert to truth and reference; it need not even advert to anything non-semantic that might underly truth and reference. For example, it need not be that 'dog' expresses the propositional constituent that it does in virtue of standing in some non-semantic relation R to dogs—where R would constitute the relation 'x is true of y'. For if truth is fully explained by the schema 'The proposition *that p* is true iff p', then the capacity of propositions and utterances to be true is guaranteed regardless of how it comes about that an utterance expresses the particular proposition it does.

Both of these consequences of the minimalist perspective are vindicated by the use theory of meaning. To see this, let us, for simplicity, restrict our attention to context-insensitive sentences (such as 'dogs bark' and 'snow is white')—sentences not involving indexical or demonstrative terms. In such cases it is plausible to identify the proposition expressed by an utterance with the meaning of the sentence-type to which the utterance belongs;[9] and it is plausible to suppose that the meaning of a sentence-type is determined by the meanings of its component words and by the procedure by which those words were combined. Thus the pair of minimalist implications about propositions amounts to the claim that word meanings are constituted neither in terms of truth-theoretic properties nor by anything to which such properties might reduce. Now according to the use theory of meaning, each word-type means what it does in virtue of a certain non-semantic regularity regarding its tokens: more specifically, that the occurrences of its tokens derive from the acceptance of certain specified sentences (or inference rules) containing the word.[10] On this view there is no reason to expect that the property in virtue of which a predicate 'f' means what it does will consist in that predicate's standing in some non-semantic relation to fs. Consequently, there is no reason to think that 'f's' meaning is constituted in terms of a non-semantic reduction of the 'true-of'

---

[9] For simplicity, I am not only ignoring context-sensitivity but also the distinction between *de dicto*, *de re*, and *de se* propositional attitudes. For a discussion of these complexities, see my *Meaning* (1998), ch. 3, Objection 12.

[10] For elaboration and defence of this position see *Meaning*, ch. 3.

relation.[11] Thus, if the use theory of meaning is correct, as I believe it is, then the two predictions of minimalism are verified.

We are now in a position briefly to address Donald Davidson's dissatisfaction with the minimalist point of view (Davidson, 1996). His critique is founded on two objections to the thesis that truth is implicitly defined by the equivalence schema. The first of them is that, contrary to minimalism, truth is conceptually prior to meaning and proposition. This claim is an immediate consequence of Davidson's own theory of meaning, according to which the meaning of a declarative sentence is constituted by its having a certain truth condition. However, Davidson's theory, although widely accepted, is far from unproblematic. Amongst the reasons for scepticism about it are:

(1) After thirty years there still remains the notorious unsolved problem of how to articulate a conception of truth condition ('*u* is true iff *p*') that is strong enough to constitute the corresponding attribution of meaning ('*u* means that *p*').

(2) Davidson's truth-conditional theory of meaning is largely motivated by the desire to explain how the meanings of sentences depend upon the meanings of their constituent words. The proposed explanations piggy-back on Tarski-style derivations of the truth conditions of sentences from the reference conditions of words. But (as we saw in the answer to Question 38) there are long-standing questions as to whether *all* the constructions of a natural language can be forced into the mould of a Tarski-style truth theory.

(3) Moreover, Davidson's truth-theoretic approach to the explanation of the compositionality of meaning is not at all compulsory. One can 'deflate' the issue by supposing that understanding a complex expression is nothing over and above understanding its component words and appreciating how they are combined.[12]

---

[11] For a detailed discussion of this point see Horwich, 1997*c*, reprinted (in a revised form) as ch. 4 of *Meaning*.

[12] This idea is sketched above in the answer to Question 22. For a full discussion of it, see Horwich, 1997*a*, reprinted as ch. 7 of *Meaning* (1998).

(4) We have just seen how it is possible to give a use-theoretic account of proposition that does not presuppose the concept of truth.

It is therefore far from evident that meaning and proposition should be explicated in terms of truth.

Davidson's second objection to the brand of deflationism presented here is that expressions like 'The proposition that dogs bark', construed as singular terms, are unintelligible. However, this rather counterintuitive claim is entirely theory-driven: it is derived from Davidson's inability to find any account (of the sort required by his truth-theoretic paradigm) of how the referents of such expressions could be determined by the referents of their parts. Therefore, given the above-mentioned doubts already shrouding that paradigm, this particular inability to accommodate it would merely constitute one more reason to give it up.[13]

4. Which is the basic truth concept—propositional truth, utterance truth, or act truth?

Ordinary language suggests that propositional truth is fundamental and that the notion of an utterance 'expressing a true proposition' and the notion of a belief 'being directed at a true proposition' are understood in terms of it. From this point of view the order of explanation will go as follows:

First: begin to characterize propositions by means of the principle

p expresses the proposition *that p*,

where the sentence-type, *p*, is individuated semantically as well as physically, and where the two tokens of '*p*' are understood in the same way.

Second: explicate propositional truth via the equivalence schema

The proposition *that p* is true iff *p*.

---

[13] For a more detailed response to Davidson's critique of minimalism, see Horwich, 1999.

Third: complete the account of propositions (and propositional truth) by adding

> u and v express the same proposition iff u and v have the same use-theoretic construction property.

Fourth: introduce utterance truth ('*truth*') and act truth ('truth*') by means of the linking principles

> Utterance u expresses $x \rightarrow$ (u is *true* iff x is true);
> Act a is directed at $x \rightarrow$ (a is true* iff x is true).

Fifth: prove the adequacy of these notions by deducing the corresponding equivalence schemata

> u expresses the proposition *that p* $\rightarrow$ (u is *true* iff p)

(which in turn implies the disquotation schema: *$*p*$ is *true* iff p)

and

> Act a, of believing (conjecturing, etc.) *that p*, is true* iff p.

However, despite its conformity with ordinary language, this order of explanation is not at all compulsory: we can equally well base everything on one of the other truth concepts. For example, we might start by characterizing utterance truth (for our own utterances) by means of the disquotation schema

> $*p*$ is *true* iff p;

then add a projection principle to cover *truth* for other utterances:

> u and $*p*$ have the same use-theoretic construction property
> $\rightarrow$ (u is *true* iff $*p*$ is *true*);

then bring in the above account of propositions:

> $*p*$ expresses the proposition *that p*;

> u and v express the same proposition $\rightarrow$
> u and v have the same use-theoretic construction property,

together with the above linking principle

> u expresses $x \rightarrow$ (u is *true* iff x is true);

from which we can easily derive the equivalence schema

The proposition *that p* is true iff *p*.

The fact that the ordinary language truth predicate typically expresses propositional truth makes the first approach more natural. And this is my rationale for taking truth, rather than *truth* or truth*, to be the conceptually basic truth concept and for supposing that its bearers are propositions; and in the same vein I take it that our concept of truth is engendered by our disposition to accept instances of the equivalence schema and that this is what explains our confidence in the truth schemata and in the disquotation schema. However, we have seen that this conceptual ordering is not especially profound or important. Our language certainly could have had a word expressing a primitive concept of utterance *truth* or act truth*; and in that case it would have been a simple matter to construct the notion of propositional truth.

5. What does minimalism have to say about the nature of truth itself, as opposed to the nature of our concept of truth?

We do not hesitate to distinguish a theory about the concept of water (or the meaning of the word 'water') from a theory of the phenomenon of water itself. The first would be supplied by a semanticist and the second by a physicist; the first would be designed to account for our thoughts and statements about water and the second for the properties of water itself; and the first need not say anything about H_2O whereas the second would have to. Similarly, it is prima facie reasonable to distinguish between an account of our concept of truth and an account of truth itself. The former purports to specify the conditions in which someone uses the word 'true' with a certain meaning; the latter purports to specify the fundamental facts about what that word stands for—about the phenomenon, truth. Consequently, the former—the meaning of 'true'—is specified by a generalization about the *word* 'true', a generalization that will explain all our uses of it; whereas the latter—the theory of truth itself—consists in principles about the *property* of truth on the basis of which all the facts about truth are to be explained. And parallel distinctions should be drawn between our concepts of truth expression and act truth, and the phenomena themselves.

What, then, from a minimalist point of view, is the basic theory of truth itself? Which body of *fundamental* facts about truth provides the best explanation of all the further facts about it? A natural conjecture is that, although there is always *in principle* a difference between the theory of the concept of X and the theory of X-ness itself, perhaps, in the case of truth, these theories more or less coincide. Perhaps the axioms of the theory of truth are instances of the equivalence schema

    The proposition *that p* is true iff *p*,

comprising what I have been calling the Minimal Theory of truth.

An immediate problem for this proposal is created by the *liar* paradoxes: the existence of contradiction-implying instances of the schema. Notice that this is not an overwhelming difficulty for the supposition that what we *mean* by 'true' is captured by the equivalence schema. For although certain instances yield contradictions, it might be argued that anyone who means what we do by 'true' has a certain inclination to accept even those instances—an inclination that is overridden by the discovery that they lead to contradiction. Indeed, one might suppose that it is only because we have such an inclination that the 'liar' sentences present us with a *paradox*! But no such manœuvre is available to protect the schema if it is proposed as a theory of truth itself. All that can be said (as we saw in the answer to Question 10) is that the theory may contain only a restricted set of the instances of the schema. But further theory will have to be deployed in order to specify that restriction.

A second deficiency of the proposed Minimal Theory, one might think, is that it has infinitely many axioms. Leaving aside the banned, contradiction-inducing instances of the equivalence schema, there is a separate axiom for each proposition. So would it not be better to pursue a Tarski-like strategy of explaining the truth of the infinitely many propositions in terms of the referents of their finitely many constituents? The answer, it seems to me, is No. The Tarski-style approach offers false hope (1) because, as is well known, there are many kinds of proposition (e.g. statements of probability, counterfactual conditionals, etc.) whose truth we have no reason to believe can be explained on the basis of the referents of their parts; and, more importantly, (2) because such a strategy would miss those propositions that are constructed from the primitive concepts that are not expressed in our language. If *all* propositions are to be cov-

ered, then there would have to be axioms specifying the referents of all the infinitely many possible primitives. So the Tarskian approach would turn out to need no fewer axioms than the Minimal Theory.

A third objection to the Minimal Theory (emphasized by Anil Gupta (1993*a*) and Scott Soames (1997)) is that it is too weak to explain any *general* facts about truth. Consider, for example, the fact that

> Every proposition of the form '*p* → *p*' is true.

We might try to explain this by assembling particular explanatory inferences that invoke nothing more than axioms of the Minimal Theory. For example, given that if snow is white, then snow is white, and given the axiom

> The proposition *that if snow is white then snow is white* is true iff if snow is white then snow is white,

we may explain why it is that

> The proposition *that if snow is white then snow is white* is true.

But in order to arrive at the *general* fact to be explained we need to collect all these conclusions together; and there is no logically valid rule that would enable us to do so. Clearly, a set of premises attributing some property to each object of a certain kind does not entail that everything of that kind has the property. We would need a further premise specifying that we have a premise for every object of that kind—and this would be tantamount to our conclusion.

However, it seems to me that in the present case, where the topic is *propositions*, we can find a solution to this problem. For it is plausible to suppose that there is a truth-preserving rule of inference that will take us from a set of premises attributing to each proposition some property, *F*, to the conclusion that all propositions have *F*. No doubt this rule is not *logically* valid, for its reliability hinges not merely on the meanings of the logical constants, but also on the nature of propositions. But it is a principle we do find plausible. We commit ourselves to it, implicitly, in moving from the disposition to accept any proposition of the form '*x* is *F*' (where *x* is a proposition) to the conclusion 'All propositions are *F*'. So we can suppose that this rule is what sustains the explanations of the generalizations about truth with which we are concerned. Thus we can, after all,

defend the thesis that the basic theory of truth consists in some sub-
set of the instances of the equivalence schema.

But there remains one further question. Even if all the familiar
characteristics of truth do indeed stem from the equivalence facts,
might not these facts in turn be explained by some more fundamen-
tal theory—perhaps a theory that attributes to truth some specific
underlying nature? The answer, it seems clear, is No. For the equiva-
lence axioms are *conceptually* basic and a priori. In these respects
they are on a par with instances of fundamental logical laws, such as
the law of excluded middle, and with the basic principles of arith-
metic. And in no such case should we anticipate theoretical reduc-
tion, as we should in the empirical domain. For in the area of a
posteriori facts, the behaviour of systems can be expected to be
caused by the properties of their parts and the way those parts are
combined with one another; therefore one might well anticipate
some reductive explanation of this behaviour. In the a priori
domain, however, there is no such reason to anticipate a reduction.
And in the particular case of the equivalence axioms there are par-
ticular grounds for scepticism. For an explanation of them would
have to provide a simple body of facts about truth and other matters,
from which they could all be deduced. But that body of facts would
itself need to be infinite—since the behaviour of infinitely many dif-
ferent propositional constituents must be accounted for. Con-
sequently, the unification and gain of simplicity required of a decent
explanation could be achieved only if the new theory were to explain
not merely the equivalence axioms but various other phenomena as
well. But we have absolutely no reason at all to think that there exist
any other phenomena whose properties might be explained by the
same theory that explains the equivalence axioms. It would be very
surprising if there were. Thus it is indeed reasonable to regard the
equivalence axioms as explanatorily basic; hence to suppose that
truth has no underlying nature.

Thus minimalism is not merely the view that truth has no *concep-
tual* analysis. It involves the stronger thesis that *no* sort of reduction
can be expected. And it bases that claim on a certain line of thought:
namely, that the utility of the concept is non-descriptive and is
explained by our acceptance of the truth schemata, that such accep-
tance constitutes our grasp of truth, and that the property so char-
acterized will not be susceptible to reductive analysis. This is why we
should regard neither Moore (1899) nor Davidson (1990) as

minimalists. For despite their insistence that truth is an indefinable primitive, they don't acknowledge the non-descriptive function of the concept, they don't take it to be implicitly defined by the truth schemata, and they do not oppose the prospect of a substantive reduction.

6. What does minimalism have to say about the normative significance of truth?

In Michael Dummett's famous paper of 1959 he argues against the redundancy theory that it fails to capture the *desirability* of truth.[14] The redundancy theory tells us *when* our beliefs are true: it says that our belief *that p* will be true if and only if *p*. But it leaves out the important fact that we *want* our beliefs to be true; we aim for the truth. And the same point could be made against minimalism.

In response, however, it can be said that in order for an account of truth to be adequate it suffices that it be able to *explain* the desirability of truth—it is not required that the desirability of truth be an integral *part* of the account. Moreover, such an explanation (or, at least, a sketch of one) can indeed be given.

Roughly speaking, it is easily seen why I should want it to be the case, for example, that I *believe* that if I run I will escape, only if I *will* escape if I run. I want this because, given a desire to escape, that belief would lead to a certain action (running), and that action would satisfy my desire if indeed it implies escape. This is why I would like it to be that I believe that I will escape if I run, only if I will indeed escape if I run.

Clearly, one has parallel reasons for wishing that

> If I believe that *D*, then *D*

whenever '*D*' expresses what we might call 'a directly action-guiding proposition'—namely, a conditional of the form 'If I do *A*, then I will get *G*', whose antecedent describes a possible action under one's

---

[14] Dummett's view is elaborated by Crispin Wright (1992) (to whom I respond in section 8 of this Postscript) and by Bernard Williams (1996) (to whom I responded in the answer to Question 19).

control, and whose consequent describes a state of affairs that one might wish to be realized.

Now recall the use of our concept of truth to capture such generalizations. We can reformulate the schema as

> If I believe *that D*, then *that D* is true,

which has the logical form

> If I believe X, then X is true,

enabling it to be generalized in the normal way as

> All my directly action-guiding beliefs are true.

This is what we have good practical reason to want. But, in addition, bear in mind that directly action-guiding beliefs are derived from other beliefs by means of inferential rules that tend to preserve truth. And any one of our beliefs might be a premise in some such inference. It thus becomes pretty clear why we might like *all* our beliefs to be true.[15]

7. But doesn't the existence of empirical generalizations about truth—including the fact that true beliefs facilitate successful action—suggest that truth is, after all, a property with an underlying nature?

The idea behind this objection is that an empirical generalization is typically explained by the underlying natures of the properties that it relates to one another. Therefore one might expect that if there are any general, empirical facts about truth, their existence would depend, contrary to minimalism, on truth's having some underlying nature. And, as we have seen, it does appear to be a general fact about truth that, as Hartry Field puts it, 'agents tend to be more successful when the truth conditions of their beliefs are realized'.[16]

It is a mistake, however, to suppose that generalizations involving

---

[15] See the response to Question 11 and Barry Loewer (1993) for refinements of this idea.

[16] Both Field (1992) and Michael Devitt (1991) endorse the suspicion, already articulated in Question 11, that minimalism may well be undermined by the role of truth in explaining successful action.

truth may be explained by analysing that property. Given the function of our concept of truth, we can see that these generalizations are not focused on truth, not really about truth. Rather they belong to that class of special schematic generalizations that rely on the equivalence schema for their formulation. It was shown in section 5 that such facts must be explained in terms of yet more basic schemata and in terms of the equivalence axioms; nothing can be gained by attempting to identify truth itself with some deeper property.

As we have just seen, it is more or less possible to explain the generalization in question in terms of the truth of the instances of the schemata

   I. If S believes that if p then q, and if S wants that q, then S wants that p;

  II. If S wants that he does A, then S does A;

 III. If S believes that q follows from $p_1, p_2, \ldots,$ and p_n, then (probably) q does follow from $p_1, p_2, \ldots,$ and p_n.

And, by the same token, each of these schematic facts 'about truth' will be explained by reference to yet more basic schemata and to instances of the equivalence schema. No analysis of truth is called for.

8. Is truth a property?

Minimalism does not involve, in itself, any particular answer to this question. For it may be combined with a variety of different conceptions of property, some of which will yield the conclusion that the truth predicate does stand for a property, and some that it doesn't.

Consider, for example, the liberal conception according to which every term that functions logically as a predicate stands for a property. Now it is a vital feature of truth that we can argue as follows:

    x is true;
    x = the proposition *that p*;
  ∴ The proposition *that p* is true;
  ∴ p.

Therefore the truth predicate must indeed be rendered in logic as a predicate. Thus there is a perfectly legitimate, weak conception of property according to which minimalism implies that truth certainly is one. However, this does not preclude the construction of various stricter, more 'robust' conceptions of property, by adding further conditions. One might say, for example, that a predicate expresses a 'substantive' property if and only if there is no a priori obstacle to its being reducible to non-semantic terms. Minimalism, in so far as it maintains, on the basis of a priori considerations, that truth is not naturalistically reducible, will imply that it is not, in that sense, 'substantive'. Thus the only reasonable way to approach the question of whether, from a minimalistic point of view, truth is a property, is to distinguish various conceptions of property, and to answer on a case by case basis.

Paul Boghossian (1990) has argued that minimalism is incoherent ('unstable') on the grounds that it implies both that truth *is*, and that it *is not*, a property. For, according to Boghossian, the essence of minimalism—that which distinguishes it from traditional accounts like the correspondence and coherence theories—is its claim that the truth predicate does not really stand for a property. However, he continues, it is also constitutive of the doctrine that any grammatically appropriate sentence be substitutable into the disquotation schema

'p' is true iff p.

And this is tantamount to accepting the generality of the schema

'p' expresses a fact iff p,

which implies

'Fx' expresses a fact iff Fx,

which, quite plausibly, implies

'F' stands for a property of x iff Fx.

Therefore, since, for example, the proposition *that snow is white* is true, it follows that 'true' stands for a property.

There is, however, no reason for a minimalist to fall into this contradiction. The second leg of Boghossian's argument invokes the weak, logical notion of property; and we can accept that truth is a property of this sort. In that case we will say that what distinguishes the minimalist conception from traditional accounts is its character-

istic view of the meaning and function of the truth predicate, together with what that implies about the nonexistence of any reductive theory of truth. Alternatively, we can operate throughout with some more substantive conception of property. And in that case we might agree with the characterization of minimalism as implying that truth is not a property. However, relative to this notion of property the minimalist should not accept the general schema

'*F*' stands for a property of *x* iff *Fx*,

and can well deny that this follows from the generality of the equivalence schema.

Crispin Wright (1992, 1997) has also claimed that the minimalist view of truth is 'unstable'—but on quite different grounds from those of Boghossian. Wright argues that since truth is a 'distinctive norm'—i.e. since (a) it is desirable that our statements be true, and (b) truth is distinct from justification—then one must recognize that truth is a 'real' property.

But we must consider whether these claims really conflict with minimalism. And to that end we must review what might be meant by 'a *real* property'. There are various possibilities.

(1) Were it to mean 'the sort of property that any logically normal predicate stands for', then, as we have seen, the minimalist has no disagreement that truth is indeed this sort of property.

(2) Perhaps, then, Wright has in mind what I have called a 'substantive' property: namely, the sort of property for which there might well be a constitution theory of the form

x is true = *x* is *F*.

But it surely does not follow, from the normative character of truth, that truth is 'real', in that sense. On the contrary, we saw in section 7 that no reductive theory of truth is likely to be correct because no such account will explain the equivalence axioms.

(3) Another possibility is that we count the following infinitely disjunctive property as 'real':

    x = the proposition *that snow is white* and snow is white;
  or *x* = the proposition *that God exists* and God exists;
  or . . ..

But again this is unlikely to be what Wright is supposing since, apart from terminology, it is more or less exactly the minimalist account of truth.

(4) Of course one might introduce a special sense of 'real property' for which it is *stipulated* that any 'distinctively normative' property (as defined above) is to count as a 'real' property. But again the minimalist could have no quarrel—except concerning the *point* of such a stipulation.

(5) The only other possibility I can think of is that Wright's notion of 'real' property resembles my notion of 'substantive' property— except for not requiring that there be any prospect of reduction. It may be, in other words, that his conception of 'property with a real nature' appeals to an intuitive picture whereby it is possible for a property to have a 'real nature' even though we can see a priori that no such nature is likely to be specifiable. If this is indeed the right interpretation of Wright's thesis then there are two replies. First, in the absence of any explicit articulation of the criterion for being such a property, and in the absence of any argument for Wright's contention that a distinctively normative property would have to satisfy the criterion, we have absolutely no reason to accept that contention. Second, there is intuitive evidence for the opposite conclusion. For suppose that a concept of 'truth' (perhaps not identical to our own) is introduced by means of the stipulation that it will apply to the proposition *that snow is white* if and only if snow is white, to the proposition *that $E = mc^2$* if and only if $E = mc^2$, and so on. Then it would seem to be consistent with our intuitive conception of 'real nature' and of 'property constitution' that the 'truth' of the proposition that snow is white simply consists in snow being white, that the 'truth' of the proposition that $E = mc^2$ simply consists in E being equal to mc^2, etc.— which will imply that 'truth' *as such* has no real nature. And this despite the fact that it would nonetheless follow from the equivalence schema that 'truth' is a 'distinctive norm' (in Wright's sense). Consequently, it does not follow from the fact that our actual concept of truth is 'distinctively normative' that it stands for a property with a real nature.

Conclusion

The minimalist picture of truth has three principal components: first, an account of the utility of truth (namely, to enable the explicit formulation of schematic generalizations); second, an account of the concept of truth (namely, that 'true' is implicitly defined by the equivalence schema); and third, an account of the nature of truth (namely, that truth has no underlying nature, and that the explanatorily basic facts about it are instances of the equivalence schema).

These ideas are supported as follows. The thesis about utility is justified by a couple of considerations. First, there appear to be no other convenient ways of expressing the generalizations that can be captured with the concept of truth; the alternative would be to supplement our language with the relatively cumbersome apparatus of sentential quantification. And second, there appear to be no other advantages of having the concept.

The thesis about the meaning of 'true' is based on two assumptions, both of which are themselves justifiable. First, the facts in virtue of which we mean what we do by 'true' are those that best account for our use of the term. And second, our use of the term is best explained by our acceptance of the equivalence schema. The first of these assumptions comes out of a general theory of meaning sketched in Chapter 6 and elaborated in my *Meaning* (1998). The second is justified by the difficulty of finding any uses of the truth predicate that cannot be explained in terms of the equivalence schema. Or to put it another way: it is justified by showing that the generalizing function of the truth predicate is explained by our acceptance of that schema.

The thesis that truth has no underlying nature derives from the foregoing. Given the function of truth, we may infer that the general facts which we need the concept of truth to articulate are not really *about truth*; therefore their explanation would not be facilitated by an account of truth's underlying nature. Rather, the facts about truth that will enter into the explanation of those generalizations will be instances of the equivalence schema. These instances are conceptually basic and a priori; hence very likely to be unsusceptible to reductive explanation. It is reasonable to conclude, therefore, that truth is not constituted by some more fundamental property.

This is not the place to expand further on the philosophical

ramifications of minimalism—but they are evidently considerable. For the concept of truth is deployed throughout philosophy, often in ways that are in tension with the minimalist picture. As we have seen, one example is the formulation of emotivism as the theory that ethical pronouncements have no truth value. Another is the view that the realism/anti-realism debate concerns the nature of truth. Another is the doctrine that meaning should be explained in terms of truth. Another is the idea that truth is precluded by vagueness and other forms of indeterminacy. In each of these cases, progress towards clearer formulations, and a better sense of where the problems really lie, is achieved by appreciating that truth is metaphysically trivial—nothing more than a device of generalization.

Bibliography

AUSTIN, J. L. (1950). 'Truth', *Proceedings of the Aristotelian Society*, supp. vol. 24: 111–28.

—— (1961). 'Unfair to Facts', *Philosophical Papers*. Oxford: Oxford University Press.

AYER, A. J. (1935). 'The Criterion of Truth', *Analysis*, 3.

—— (1936). *Language, Truth and Logic*. London: Dover.

—— (1963). 'Truth', *The Concept of a Person and Other Essays*. London: Macmillan.

BALDWIN, T. (1989). 'Can There Be a Substantive Theory of Truth?', *Recherche sur la Philosophie et le Language*, 10. Grenoble: Université des Sciences Sociales de Grenoble.

BARWISE, J., and PERRY, J. (1983). *Situations and Attitudes*. Cambridge, Mass.: MIT Press.

BLACK, M. (1948). 'The Semantic Definition of Truth', *Analysis*, 8: 49–63.

BLACKBURN, S. (1984). *Spreading the Word*. Oxford: Oxford University Press.

BLACKBURN, T. (1988). 'The Elusiveness of Reference', in P. French, T. Uehling, and H. Wettstein, eds., *Midwest Studies in Philosophy*, 12. Minneapolis, Minn.: University of Minnesota Press.

BLANSHARD, B. (1939). *The Nature of Thought*, vol. ii. London: Allen & Unwin.

BLOCK, N. (1986). 'Advertisement for a Semantics for Psychology', in P. French, T. Uehling, and H. Wettstein, eds., *Midwest Studies in Philosophy*, 10. Minneapolis, Minn.: University of Minnesota Press.

BOGHOSSIAN, P. (1990). 'The Status of Content', *Philosophical Review*, 99: 157–84.

BRADLEY, F. H. (1914). *Essays on Truth and Reality*. Oxford: Oxford University Press.

BRANDOM, R. (1984). 'Reference Explained Away', *Journal of Philosophy*, 81: 469–92.

—— (1988). 'Pragmatism, Phenomenalism and Truth Talk', in P. French, T. Uehling, and H. Wettstein, eds., *Midwest Studies in Philosophy*, 12. Minneapolis, Minn.: University of Minnesota Press.

BRANDOM, R. (1994). *Making it Explicit*. Cambridge, Mass.: Harvard University Press.

BROAD, C. D. (1914). 'Mr. Bradley on Truth and Reality', *Mind*, 23: 349–70.

BURGE, T. (1979). 'Individualism and the Mental', in P. French, T. Uehling, and H. Wettstein, eds., *Midwest Studies in Philosophy*, 4. Minneapolis, Minn.: University of Minnesota Press.

CARNAP, R. (1942). *Introduction to Semantics*. Cambridge, Mass.: Harvard University Press.

—— (1947). *Meaning and Necessity*. Chicago, Ill.: University of Chicago Press.

CARTWRIGHT, R. (1962). 'Propositions', repr. (1987) in his *Philosophical Essays*. Cambridge, Mass.: MIT Press.

—— (1987). 'A Neglected Theory of Truth', *Philosophical Essays*. Cambridge, Mass.: MIT Press.

COHEN, J. (1950). 'Mr. Strawson's Analysis of Truth', *Analysis*, 10: 136–44.

COUSIN, D. R. (1959/60). 'Carnap's Theories of Truth', *Mind*, 59: 1–22.

DAVIDSON, D. (1967). 'Truth and Meaning', *Synthese*, 17: 304–23. Repr. in Davidson (1984).

—— (1969). 'True to the Facts', *Journal of Philosophy*, 66: 748–64. Repr. in Davidson (1984).

—— (1984). *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.

—— (1990). 'The Structure and Content of Truth', *Journal of Philosophy*, 87: 279–328.

—— (1996). 'The Folly of Trying to Define Truth', *Journal of Philosophy*, 93: 263–78.

DEVITT, M. (1984). *Realism and Truth*. Princeton, NJ: Princeton University Press.

—— (1991). 'Deflationist Truth: A Critical Notice of Paul Horwich's *Truth*', *Mind and Language*, 6: 273–83.

—— and STERELNY, K. (1989). *Language and Reality*. Cambridge, Mass.: MIT Press.

DEWEY, J. (1916). *Essays in Experimental Logic*. Chicago, Ill.: Dover.

—— (1938). *Logic: The Theory of Inquiry*. New York: Irvington.

DRETSKE, F.I. (1981). *Knowledge and the Flow of Information*. Cambridge, Mass.: MIT Press.

DUHEM, P. (1954). *The Aim and Structure of Physical Theory*. Princeton, NJ: Princeton University Press. First published in 1906.

DUMMETT, M. (1959). 'Truth', *Proceedings of the Aristotelian Society*, 59: 141–62; repr. in Dummett (1978).

—— (1975). 'What is a Theory of Meaning? Part I', in S. Guttenplan, ed., *Mind and Language*). Oxford: Oxford University Press.

—— (1976). 'What is a Theory of Meaning? Part II', in G. Evans and J.

McDowell, eds., *Truth and Meaning: Essays in Semantics*. Oxford: Oxford University Press.

—— (1977). *Elements of Intuitionism*. Oxford: Oxford University Press.

—— (1978). *Truth and Other Enigmas*. Cambridge, Mass.: Harvard University Press.

EVANS, G. (1982). *The Varieties of Reference*. Oxford: Oxford University Press.

—— (1985). *Collected Papers*. Oxford: Oxford University Press.

EZORSKY, G. (1963). 'Truth in Context', *Journal of Philosophy*, 60: 113–35.

FIELD, H. (1972). 'Tarski's Theory of Truth', *Journal of Philosophy*, 69: 347–75.

—— (1977). 'Logic, Meaning and Conceptual Role', *Journal of Philosophy*, 74: 379–409.

—— (1986). 'The Deflationary Conception of Truth', in G. MacDonald and C. Wright, eds., *Fact, Science and Morality*. Oxford: Blackwell.

—— (1992). 'Critical Notice: Paul Horwich's *Truth*', *Philosophy of Science*, 59: 321–30.

—— (1994). 'Deflationist Views of Meaning and Content', *Mind*, 94: 249–85.

FINE, A. (1984). 'The Natural Ontological Attitude', in J. Leplin, ed., *Scientific Realism*. Berkeley, Calif.: University of California Press.

—— (1984). 'And Not Anti-Realism Either', *Nous*, 18: 51–65.

FINE, K. (1975). 'Vagueness, Truth and Logic', *Synthese*, 30: 265–300.

FODOR, J. (1984). 'Semantics, Wisconsin Style', *Synthese*, 59: 231–50.

—— (1987). *Psychosemantics*. Cambridge, Mass.: MIT Press.

FORBES, G. (1986). 'Truth, Correspondence and Redundancy', in G. MacDonald and C. Wright, eds., *Fact, Science and Morality*. Oxford: Blackwell.

FREGE, G. (1891). 'On Function and Concept', in *Translations from the Philosophical Writings of G. Frege*, ed. and trans. M. Black and P. Geach. Oxford: Blackwell, 1960.

—— (1892). 'On Sense and Reference', in *Translations from the Philosophical Writings of G. Frege*, ed. and trans. M. Black and P. Geach. Oxford: Blackwell, 1960.

—— (1918). 'The Thought', trans. A. Quinton and M. Quinton, *Mind*, 65 (1956): 289–311.

FRIEDMAN, M. (1979). 'Truth and Confirmation', *Journal of Philosophy*, 76: 361–82.

GEACH, P. T. (1960). 'Ascriptivism', *Philosophical Review*, 69: 221–5.

—— (1965). 'Assertion', *Philosophical Review*, 74: 449–65.

GROVER, D., Camp, J., and Belnap, N. (1975). 'A Prosentential Theory of Truth', *Philosophical Studies*, 27: 73–125.

150   *Bibliography*

GUPTA, A. (1982). 'Truth and Paradox', *Journal of Philosophical Logic*, 11: 1–60; repr. in R. L. Martin, ed., *Recent Essays on Truth and the Liar Paradox*. Oxford: Oxford University Press, 1984.

—— (1993*a*). 'A Critique of Deflationism', *Philosophical Topics*, 21: 57–81.

—— (1993*b*). 'Minimalism', in James Tomberlin, ed., *Philosophical Perspectives*, 7: 359–69.

—— and BELNAP, N. (1993). *The Revision Theory of Truth*. Cambridge, Mass.: MIT Press.

HARMAN, G. (1973). *Thought*. Princeton, NJ: Princeton University Press.

—— (1974). 'Meaning and Semantics', in M. K. Munitz and P. Unger, eds., *Semantics and Philosophy*. New York: New York University Press.

—— (1982). 'Conceptual Role Semantics', *Notre Dame Journal of Formal Logic*, 23: 242–56.

—— (1987). '(Nonsolipsistic) Conceptual Role Semantics', in E. LePore, ed., *New Directions in Semantics*. London: Academic Press.

HEMPEL, C. (1935). 'On the Logical Positivist's Theory of Truth', *Analysis*, 2: 49–59.

HINTIKKA, J. (1962). *Knowledge and Belief*. Ithaca, NY: Cornell University Press.

HORNSBY, J. (1997). 'Truth: The Identity Theory', *Proceedings of the Aristotellan Society*, 97: 1–24.

HORWICH, P. G. (1982*a*). 'Three Forms of Realism', *Synthese*, 51: 181–201.

—— (1982*b*). *Probability and Evidence*. Cambridge: Cambridge University Press.

—— (1987). *Asymmetries in Time*. Cambridge, Mass.: MIT Press.

—— (1990). *Truth*, 1st edn. Oxford: Blackwell.

—— (1991). 'On the Nature and Norms of Theoretical Commitment', *Philosophy of Science*, 58: 1–14.

—— (1993). 'Gibbard's Theory of Norms', *Philosophy and Public Affairs*, 22: 67–78.

—— (1994). 'The Essence of Expressivism', *Analysis*, 45: 19–20.

—— (1995). 'Meaning, Use and Truth', *Mind*, 104: 355–68.

—— (1996). 'Realism and Truth', in J. Tomberlin, ed., *Philosophical Perspectives*, 10.

—— (1997*a*). 'The Composition of Meanings', *Philosophical Review*, 106: 503–31; repr. in Horwich (1998).

—— (1997*b*). 'The Nature of Vagueness', *Philosophy and Phenomenological Research*, 57: 929–36.

—— (1997*c*). 'Deflationary Truth and the Problem of Aboutness', in E. Villanueva, ed., *Philosophical Issues*, 8: 95–106. Atascadero, Calif.: Ridgeview Publishing Company.

—— (1997*d*). 'Realism Minus Truth', *Philosophy and Phenomenological Research*, 56.

Bibliography

151

—— (1997*e*). 'Implicit Definition, Analytic Truth, and Apriori Knowledge', *Nous*, 31: 423–40.

—— (1998). *Meaning*. Oxford: Oxford University Press.

—— (1999). 'Davidson on Deflationism', in U. Zeglen, ed., *Discussions with Donald Davidson: On Truth, Meaning and Knowledge*. London: Routledge.

JACKSON, F., OPPY, G., and SMITH, M. (1994). 'Minimalism and Truth Aptness', *Mind* 103: 287–302.

JACOB, P. (1997). *What Minds Can Do*. Cambridge: Cambridge University Press.

JAMES, W. (1909). *The Meaning of Truth*. Cambridge, Mass.: Harvard University Press.

KAPLAN, D. (1969). 'Quantifying In', in D. Davidson and J. Hintikka, eds., *Words and Objections*. Dordrecht: Reidel.

KATZ, J. J. (1972). *Semantic Theory*. New York: Harper & Row.

KAUFMAN, F. (1948). 'Rudolf Carnap's Analysis of Truth', *Philosophy and Phenomenological Research*, 9: 294–9.

KINCADE, J. (1958). 'On the Performatory Theory of Truth', *Mind*, 67: 394–9.

KIRKHAM, R.L. (1992). *Theories of Truth*. Cambridge, Mass.: MIT Press.

KRIPKE, S. (1972). 'Naming and Necessity', in D. Davidson and G. Harman, eds., *Semantics of Natural Language*. Dordrecht: Reidel.

—— (1975). 'Outline of a Theory of Truth', *Journal of Philosophy*, 72: 690–716.

KUHN, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago, Ill.: University of Chicago Press.

LEEDS, S. (1978). 'Theories of Reference and Truth', *Erkenntnis*, 13: 111–29.

LEMMON, E. J. (1966). 'Sentences, Statements, and Propositions', in B. A. O. Williams and A. C. Montefiore, eds., *British Analytic Philosophy*. London: Routledge & Kegan Paul.

LEWIS, D. (1972). 'General Semantics', in D. Davidson and G. Harman, eds., *Semantics of Natural Language*. Dordrecht: Reidel.

LEVISON, A. B. (1924). 'Logic, Language and Consistency in Tarski's Theory of Truth', *Philosophy and Phenomenological Research*, 25: 384–92.

LOAR, B. (1981). *Mind and Meaning*. Cambridge: Cambridge University Press.

—— (1987). 'Truth Beyond All Verification', in B. Taylor, ed., *Michael Dummett*. Dordrecht: Nijhoff.

LOEWER, B. (1993). 'The Value of Truth', in E. Villanueva, ed., *Philosophical Issues*, 4. Atascadero, Calif.: Ridgeview Publishing Company.

MAYO, B. (1959). 'Truth as Appraisal', *Mind*, 68: 80–6.

McGEE, V. (1991). *Truth, Vagueness, and Paradox*. Indianapolis, Ind.: Hackett.

MILLIKAN, R. (1984). *Language, Thought and Other Biological Categories.* Cambridge, Mass.: MIT Press.

MOORE, G. E. (1899). 'The Nature of Judgment', *Mind*, 8: 176–93.

—— (1910/11). Lectures, in *Some Main Problems of Philosophy.* London: Allen & Unwin, 1953.

—— (1927). 'Facts and Propositions', *Proceedings of the Aristotelian Society*, supp. vol. 7: 171–206.

PAP, A. (1952). 'A Note on the "Semantic" and the "Absolute" Concept of Truth', *Philosophical Studies*, 3: 1–8.

—— (1954). 'Propositions, Sentences, and the Semantic Definition of Truth', *Theoria*, 20: 23–35.

PAPINEAU, D. (1987). *Reality and Representation.* Oxford: Blackwell.

—— (1991). 'Truth and Teleology', in D. Knowles, ed., *Explanation and its Limits.* Cambridge: Cambridge University Press.

PEACOCKE, C. (1986). *Thoughts.* Oxford: Blackwell.

—— (1992). *A Study of Concepts.* Cambridge, Mass.: MIT Press.

PEIRCE, C. S. (1905). 'What Pragmatism Is', *The Monist*, 15: 161–81.

—— (1955). 'How to Make Our Ideas Clear', in *Philosophical Writings of Peirce*, ed. J. Buchler. New York: Dover.

—— (1932–3). *Collected Papers*, vols. ii–iv. Cambridge, Mass.: Harvard University Press.

PITCHER, G., ed. (1964). *Truth.* Englewood Cliffs, NJ: Prentice-Hall.

POPPER, K. (1955). 'A Note on Tarski's Definition of Truth', *Mind*, 64: 388–91.

—— (1962). *Conjectures and Refutations.* New York: Basic Books.

PUTNAM, H. (1975). 'The Meaning of "Meaning"', in K. Gunderson, ed., *Language, Mind and Knowledge*: *Minnesota Studies in the Philosophy of Science*, 7. Minneapolis, Minn.: University of Minnesota Press.

—— (1978). *Meaning and the Moral Sciences.* London: Routledge & Kegan Paul.

—— (1981). *Reason, Truth and History.* Cambridge: Cambridge University Press.

—— (1983). 'Vagueness and Alternative Logic', *Philosophical Papers*, vol. iii: *Realism and Reason.* Cambridge: Cambridge University Press.

—— (1988). *Representation and Reality.* Cambridge, Mass.: MIT Press.

QUINE, W. V. (1953). *From a Logical Point of View.* Cambridge, Mass.: Harvard University Press.

—— (1960). *Word and Object.* Cambridge, Mass.: MIT Press.

—— (1970). *Philosophy of Logic.* Englewood Cliffs, NJ: Prentice-Hall.

—— (1977). 'Intensions Revisited', in *Midwest Studies in Philosophy*, 2.

—— (1990). *Pursuit of Truth.* Cambridge, Mass.: Harvard University Press.

RAMSEY, F. (1927). 'Facts and Propositions', *Proceedings of the Aristotelian Society*, supp. vol. 7: 153–70.

—— (1991). 'The Nature of Truth', *On Truth: Original Manuscript*

Materials (1927–1929) from the Ramsey Collection at the University of Pittsburgh, edited by N. Rescher and U. Majer, *Episteme*, 16: 6–16.

RICHARD, M. (1997). 'Deflating Truth', in E. Villanueva, ed., *Philosophical Issues*, 8. Atascadero, Calif.: Ridgeview Publishing Company.

RORTY, R. (1982). *Consequences of Pragmatism*. Minneapolis, Minn.: Minnesota University Press.

RUSSELL, B. (1903). *The Principles of Mathematics*. New York: Norton.

—— (1904). 'Meinong's Theory of Complexes and Assumptions', *Mind*, 13: 204–52. Repr. in D. Lackey, ed., *Essays in Analysis*. New York: Braziller, 1973.

—— (1905). 'On Denoting', *Mind*, 14.

—— (1906). 'On the Nature of Truth and Falsehood', repr. in his *Philosophical Essays* (1966). New York: Simon & Schuster.

—— (1919). 'On Propositions: What They Are and How They Mean', *Proceedings of the Aristotelian Society*, supp. vol. 2: 1–43.

RYLE, G. (1929). 'Are There Propositions?', *Proceedings of the Aristotelian Society*, 30: 91–126.

SCHIFFER, S. (1994). 'A Paradox of Meaning', *Nous*, 28: 279–34.

SCHLICK, M. (1935). 'Facts and Propositions', *Analysis*, 2: 65–70.

SEARLE, J. R. (1958). 'Proper Names', *Mind*, 63: 166–73.

—— (1962). 'Meaning of Speech Acts', *Philosophical Review*, 71: 423–32.

—— (1978). 'Literal Meaning', *Erkenntnis*, 13: 207–24.

SELLARS, W. (1954). 'Some Reflections on Language Games', *Philosophy of Science*, 21: 204–28.

—— (1962). 'Truth and "Correspondence"', *Journal of Philosophy*, 59: 29–56.

SMITH, M. (1994*a*). 'Why Expressivists about Value Should Love Minimalism about Truth', *Analysis*, 54: 1–12.

—— (1994*b*). 'Minimalism, Truth-Aptitude, and Belief', *Analysis*, 54: 21–6.

SOAMES, S. (1984). 'What is a Theory of Truth', *Journal of Philosophy*, 81: 411–29.

—— (1997). 'The Truth about Deflationism', in E. Villanueva, ed., *Philosophical Issues*, 8. Atascadero, Calif.: Ridgeview Publishing Company.

STALNAKER, R. (1984). *Inquiry*. Cambridge, Mass.: MIT Press.

STAMPE, D. W. (1977). 'Toward a Causal Theory of Linguistic Representation', *Midwest Studies in Philosophy*, 2: 42–63.

STERELNY, K. and DEVITT, M. (1989). *Language and Reality*. Cambridge, Mass.: MIT Press.

STRAWSON, P. (1950). 'Truth', *Proceedings of the Aristotelian Society*, supp. vol. 24: 129–56.

—— (1964). 'A Problem About Truth—A Reply to Warnock', in Pitcher (1964).

TARSKI, A. (1943/4). 'The Semantic Conception of Truth', *Philosophy and Phenomenological Research*, 4: 341–75.

—— (1958). 'The Concept of Truth in Formalized Languages', in his *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*. Oxford: Oxford University Press.

THOMSON, J. F. (1948). 'A Note On Truth', *Analysis*, 8: 67–72.

USHENKO, A. (1944). 'A Note on the Semantic Conception of Truth', *Philosophy and Phenomenological Research*, 4: 104–7.

VAN FRAASSEN, B. (1980). *The Scientific Image*. Oxford: Oxford University Press.

VAN INWAGEN, P. (1988). 'On Always Being Wrong', in P. French, T. Uehling, and H. Wettstein, eds., *Midwest Studies in Philosophy*, 12. Minneapolis, Minn.: University of Minnesota Press.

WALSH, W. H. (1952). 'A Note on Truth', *Mind*, 61: 72–4.

WARNOCK, G. (1964) 'A Problem about Truth', in Pitcher (1964).

WHITE, A. R. (1957). 'Truth as Appraisal', *Mind*, 66: 318–30.

WILLIAMS, B. (1996). 'Truth in Ethics', *Ratio*, 8: special issue, *Truth in Ethics*, ed. Brad Hooker, 227–42.

WILLIAMS, M. J. (1986). 'Do We (Epistemologists) Need A Theory of Truth?', *Philosophical Topics*, 14: 223–42.

WILLIAMSON, T. (1994). *Vagueness*. London: Routledge.

WITTGENSTEIN, L. (1922). *Tractatus Logico-Philosophicus*. London: Routledge & Kegan Paul.

—— (1953). *Philosophical Investigations*, trans. G. E. M. Anscombe. Oxford: Blackwell.

WRIGHT, C. (1987). 'Further Reflections on the Sorites Paradox', *Philosophical Topics*, 15: 227–90.

—— (1988). 'Realism, Anti-realism, Irrealism, Quasi-realism', in P. French, T. Uehling, and H. Wettstein, eds., *Midwest Studies in Philosophy*, 12. Minneapolis, Minn.: University of Minnesota Press.

—— (1992). *Truth and Objectivity*. Cambridge, Mass.: Harvard University Press.

—— (1996). 'Responses to Commentators', *Philosophy and Phenomenological Research*, 56: 863–8, 911–41.

ZIFF, P. (1962). *Semantic Analysis*. Ithaca, NY: Cornell University Press.

Index