

Facts and the Function of Truth

Huw Price

Full text of First Edition (with minor typographical corrections)

Preface

Some years ago I worked on the philosophy of probability. For a long time I thought about probabilities – what kinds of things they could be, and so on. Then it struck me that this might be the wrong approach. We didn't have to assume that there must be something that probabilistic judgements are characteristically *about*, and so concern ourselves with the nature of that distinctive subject matter. Instead we could interpret probabilistic judgement in terms of a characteristic way of talking and thinking about other things. At the time it seemed quite a revelation. Of course I soon realized that others had expressed the same view about probability, as about several other topics. But in the case of probability, at least, the view was not well developed. I went on to write a thesis (Price 1981) that argued its advantages over other views of probability, and defended it against certain objections raised against similar 'non-factualist' views in other areas.

By contemporary standards the unusual feature of this view of probability is that it takes a linguistic approach to metaphysical issues. Non-factualist theories were a common and characteristic product of the mid-century linguistic movement in philosophy; but these days that movement is widely felt to be discredited. Contemporary metaphysics regards itself as free of the peculiar concern with language that characterized that period. So as I eventually came to appreciate, a defence of such a treatment of probability (or indeed other topics) calls for some defence of the linguistic stance. More importantly, it calls for an elucidation of the cluster of notions whose proper application non-factualists are concerned to dispute – notions such as *assertion* and *statement of fact*. In this respect, as in others, the linguistic approach needs to be grounded in the philosophy of language.

I found that such a grounding had not in practice been obtained. This was not simply because it had usually not been sought. Contemporary philosophy of language has appeared to have reason to dismiss non-factualism – a reason that turns on the role of the notion of truth in Fregean linguistic theory. In effect, non-factualism and contemporary philosophy of language have wanted to draw the limits of fact-stating discourse at different places. This conflict has not had the attention that it deserves, I think. The linguistic approach to difficult areas of philosophy has more to be said for it than its reputation suggests. And the philosophy of language has much more to learn from the encounter than the history would indicate.

This book is in part an attempt at mediation. I try to show that with a proper understanding of what an account of statement of fact should amount to the conflict can be resolved. Neither side escapes unscathed. The resolution turns, in particular, on the rejection of a shared assumption about the appropriate goal of an elucidation of the notion of statement. But both sides benefit from the new perspective thus required. The major benefit is a particularly fruitful perspective on the nature and origins of truth.

I appreciate that the approach may seem misguided. Many philosophers will feel that the conflict described is long since decided in favour of the philosophy of language. I suspect that others, on both sides of the fence, will not have noticed that it ever existed. I can only urge such people to read the book – or at least to read chapter 1, which contains a more detailed guide to the nature and origins of the conflict, and the structure and benefits of the resolution I propose.

In order to make it a little easier to read the book, I have starred a number of sections of a more technical or incidental nature. These may be omitted without serious discontinuity. Chapter 1 mentions a further possibility, which involves beginning at chapter 6.

The book was written during my tenure of an Australian National Research Fellowship, in the School of Philosophy at the University of New South Wales. The Editor of *Mind* has kindly allowed me to reproduce parts of 'Truth and the nature of assertion' (Price 1987), which appear mainly in chapter 2. I am also indebted to Philip Pettit, whose assistance and encouragement long since transcended the

merely editorial; and to Peter Menzies and Phillip Staines, for their insight and patience in many conversations on this material.

1

Introduction

1.1 WHY READ THIS BOOK?

Few areas of contemporary philosophy have not been touched by the suggestion that their distinctive problems are in part a product of some sort of misconception about language. One view of this kind has been particularly widespread. It is the claim that in virtue of their grammatical form, certain uses of language are mistakenly construed as descriptive, or fact-stating. Perhaps the most familiar cases arise in ethics. Moral philosophers have often argued that despite the syntactical resemblance that ethical utterances bear to genuinely descriptive uses of language, such expressions should not be interpreted as statements of fact. There are actually several versions of this view, differing as to what they take it that ethical judgements really are. One at least of these is common currency among non-philosophers. Known to philosophers as ‘emotivism’, this version holds that moral claims are ‘value judgements’ – expressions of attitudes of approval, disapproval, or something of the kind. The Prime Minister might say for example that the Soviets have the best possible system of government, with the exception of all the rest. An emotivist would interpret the remark as an expression of the PM’s own political preferences, of her marked antipathy to the Soviet state, rather than as a factual claim in comparative political theory.

We shall need a name for the general philosophical strategy that emotivism here exemplifies. I shall call it ‘non-factualism’. ‘Anti-realism’ would be a little less ugly, and is perhaps more widely used in this context. However, it is widely used for other programmes as well. ‘Non-cognitivism’ suffers from the opposite problem, being largely confined to the ethical case. A non-factualist thus holds that there are certain utterances whose grammatical form is that of statements of fact, but whose linguistic role is actually quite different.

Further philosophical consequences will usually be held to flow from this linguistic insight. If nothing else, it allows the non-factualist to dismiss previous philosophical concerns about the nature of the ‘facts’ in question. An emotivist will say for example that the traditional philosophical problem of the nature of good and evil rests on a misconception about the language of morality. As I said, ethics is by no means the only branch of philosophy whose practitioners have been attracted to non-factualism. Aesthetic judgements, mathematical statements, theoretical sentences in science, Lockean secondary qualities, knowledge claims, meaning ascriptions, indicative and subjunctive conditionals, and probabilistic and modal judgements: all of these topics have been given non-factualist treatments, and readers may well be able to add to the list.

This book is not about any of these topics as such. Its initial concern lies rather at the next level of generality, with the foundations and hence the viability of non-factualism itself. In particular it is concerned with the nature of the distinction between factual and non-factual uses of language. I am going to argue that the view of this distinction normally taken for granted by non-factualists is radically mistaken. Non-factualists have taken the issue to concern the distribution in language of a property that we may call ‘statementhood’. Their claim is then that this property is less extensive than we ordinarily assume. Certain classes of utterance – moral judgements, or whatever – appear to possess it but actually lack it. Once the issue is couched in these terms the important questions become ‘What does the property of statementhood consist in?’ and ‘How do we determine that a particular kind of discourse possesses or lacks the property so construed?’ In other words, non-factualism then calls for an *analysis* of statement of fact – preferably, of course, an analysis that serves to support the view that there are fewer statements in language than we initially suppose.

The main task of the first part of the book will be to challenge this analytic perspective. I shall argue that of the available ways of cashing out the notion of statement of fact, none provides what the non-factualist evidently requires: a means to exclude the possibility that there are actually more statements than the non-factualist holds there to be. Indeed, we shall see that none of the available analyses can exclude the possibility that *all* linguistic utterances are fact-stating.

More in a moment on the structure of this argument. I want to make it clear first of all that the book is not against non-factualism as such. On the contrary, I am going to defend a version of non-factualism, albeit an unorthodox one, and argue that it vindicates the non-factualists' concern with language as a route to enlightenment on a range of philosophical topics. At the very least it vindicates the intuition that there is something unusual about the language of morality, probability, conditionals and a range of other topics with which non-factualists have been concerned.

In particular, therefore, the book is not aligned with the influential view that in virtue of some rather simplistic mistakes about language, non-factualism itself is misconceived. As I shall explain, this view stems particularly from the Fregean tradition in modern philosophy of language. Like non-factualism itself, this tradition is committed, by and large, to the existence in language of a distinction between fact-stating or assertoric discourse and discourse of other kinds. The conflict arises as to where this distinction should be drawn. A non-factualist wants to narrow the class of genuine statements. A Fregean, on the other hand, is inclined to do exactly the opposite – to construe the bounds of assertoric discourse as widely as possible. We shall see that this inclination is grounded particularly in the role that truth plays in Fregean linguistic theory. Its effect is that non-factualism, the epitome of philosophy in the linguistic style, comes ironically to be rejected by the philosophy of language itself.

In criticizing the usual presentation of non-factualism, how then do I fail to align myself with its Fregean critics? By rejecting the principle that non-factualists and Fregeans have in common, namely that there is a real property of statementhood, or some such, whose distribution in language can be a matter for dispute. As I said, I argue in part I that however we try to characterize this supposed property, we are left with no basis on which to show that it applies to some utterances but not to others. In particular, we cannot exclude the possibility that all utterances are statements: an admission that would be as unacceptable to (most) Fregeans as it would be to non-factualists.

I shall describe the conflict between non-factualism and Fregean semantics in more detail in the next two sections. Its effect, I think, has been to stunt the development of an adequate theory of the linguistic notions on which non-factualism depends. Both sides suffer as a result. For its part, non-factualism has not died out. It remains popular in traditional strongholds, such as ethics and aesthetics, and it continues to win converts elsewhere. As always, however, it lacks adequate conceptual foundations. On the other side, I think that contemporary philosophy of language has suffered from a failure to take seriously the non-factualist's concerns. The notions non-factualism appeals to – factuality, statement, truth, and so on – are central to an understanding of language. By focusing on some of the hard cases, non-factualism offers a particularly fruitful perspective on these notions. To dismiss non-factualism is to miss this perspective.

The book thus approaches the philosophy of language by way of linguistic philosophy. The needs of non-factualism, as viewed in the light of part I, lead in part II to a novel account of truth. One of the advantages of this account is its ability to deal with these difficult cases, but its interest and application is more general. Its novelty, by contemporary standards, lies in the fact that it is explanatory rather than analytic. By this I mean that the question it addresses is not so much 'What is truth?' as 'What use is the concept of truth, to creatures like us?' In other words, it seeks to explain our possession of the notion of truth in terms of its function in a linguistic community. It is a genealogical theory. It asks why we should ever have developed such a notion.

I shall have more to say later on about the distinction between analytic and explanatory theories. In case naturalists in the audience are already tempted to leave, however, I emphasize now that the proposed account will be thoroughly naturalistic. For any concept in common use there are (at least) two possible naturalistic accounts of how it got there. One approach is to analyse the concept concerned in naturalistic terms, and hence explain its use in terms of some general account of our reaction to our natural environment. But another – often overlooked, I think – is to find some non-referential function for the concept in the lives of natural creatures such as ourselves. The latter approach is most clearly appropriate in cases in which we have reason to deny referents to the concepts concerned – religious concepts, or those of certain former scientific theories, for example. It is not excluded in other cases, however, and the proposed account of truth will be of this kind.

In view of the wide importance of the notion of truth in contemporary philosophy, this may seem a case in which a solution outgrows the problem from which it stems. Some readers will inevitably feel that in grounding such an account of truth on the needs of non-factualism, I use a small tail to wag a very large dog. Others, more favourably disposed to the general approach of part II, will feel that the dog is quite capable of wagging itself.

To the former objection I would reply that if the argument of part I goes through then it is not simply non-factualists who need a new approach to truth and related notions. Their Fregean opponents should be equally concerned by the failure of the usual ways of construing assertoric discourse. That said, non-factualism provides a valuable guide to and constraint on alternative approaches. One of the advantages of the account of truth proposed in part II is that it does make very good sense of non-factualist intuitions. Moreover, non-factualist concerns are a powerful antidote to another non-analytic approach to truth which has had significant support in recent years. This approach reacts to the perceived failure of attempts to analyse truth by dismissing the problem, recommending instead that we change the subject and stop worrying about truth. In drawing our attention to the limits of truth, non-factualism does much to counter this unhealthy temptation to let a sleeping dog lie.

I have more sympathy with the latter objection – the view that the proposed account of truth might be motivated without reference to the needs of non-factualism. I would simply encourage the reader whose interest is more in the dog than the tail to feel free to begin at part II. Chapter 6, sections 1 to 5 of chapter 7 and chapter 9 deal largely with the general account of truth. They make little reference to non-factualism and the concerns of part I, and might be read independently. The remainder of chapter 7 and chapter 8 apply the general account to the non-factualist cases, and might be read next. Part I would then recommend itself to the reader who found the orientation implausible, but was curious to see where I had gone astray. This approach to the book aside, I have starred a number of sections to indicate that they may be omitted without serious discontinuity.

Who then might profitably read this book? First, I think, anyone with an interest in non-factualism in any of the areas in which it has been applied. I imagine that most such readers will take it that they have a reasonably clear conception of what distinguishes fact-stating discourse, and therefore as to why a non-factualist interpretation might be judged appropriate in any particular case. Part I aims to shake such confidence, and part II to show what might replace it. Secondly, I recommend the book to those with an active disinterest in non-factualism – those in particular who take its misconceptions to have been exposed by contemporary philosophy of language. Here too part I aims to undermine such prejudices. And thirdly and no doubt most numerous, I recommend it to anyone with an interest in truth. The account of truth here offered is novel and potentially important. The route will be unfamiliar to most readers, but I think the result will repay their efforts.

The remainder of this chapter is in four sections. The next, section 1.2, outlines the historical background to the conflict between non-factualism and contemporary philosophy of language. In section 1.3 I explain how the conflict can be thought of as a disagreement concerning the extension of the property of statementhood in language, and therefore how I propose to attack the common assumption that there is such a property. Section 1.4 is then a brief guide to the argument of part I,

and section 1.5 some remarks about orientation of the alternative programme to be developed in part II.

1.2 NON-FACTUALISM AND FREGEAN SEMANTICS

With the benefit of hindsight, at any rate, it is clear that non-factualism predates the linguistic movement of the mid-twentieth century. In ethics it traces its origins at least to Hume, who may be regarded as an early emotivist. Hume seems to have taken a similar view of causal judgements: namely that they do not refer to a realm of causal facts, but rather express our expectations, formed on the basis of observed regularities. But the great popularity of non-factualism in this century – indeed the recognition that Hume should be seen in these terms – seems largely associated with the linguistic turn. In those halcyon days after the war the non-factualist techniques were applied to a wider and wider range of traditional philosophical topics.

In the 1960s, however, non-factualism acquired a reputation as a cheap approach to subtle philosophical problems: too quick to claim new territory, and too slow to justify its existing claims. The reputation is not entirely undeserved. The non-factualists of the 1950s do seem to have paid too little attention to the foundations of their own position. But it also owes something to the fact that the approach had the misfortune to coincide with what, from the non-factualist's point of view, was a particularly inhospitable period in the philosophical study of language itself. For non-factualism is at odds with a view that at this period philosophers of language had increasing reason to uphold – the view that assertion or statement of fact is the central and primary use of language, rather than one use among others of equal significance.

At first sight, indeed, a reader might gain the impression that there is no place for the notion of a non-fact-stating use of language in contemporary philosophy of language. Writers will refer, for example, to the task of specifying the truth conditions of 'the sentences of a language', as if there were no such thing as a sentence without truth conditions – as an utterance whose point is not to state that certain conditions obtain. As it stands, however, this may be said to reflect a benign idealization, not inconsistent with a more pluralistic view of natural language. The real basis of the doctrine of the primacy of assertion lies deeper, I think, in the Fregean use of the notion of truth.

Once truth is part of a philosopher's repertoire, there are at least two routes to the primacy of assertion. One turns on the fact that a systematic approach to the theory and taxonomy of language will need to distinguish the 'genuine' uses of languages, to which the theory should be expected to apply, from various sorts of pseudo-linguistic noise – for example the kind of noises we make for dog rousing, baby settling and the like. A natural suggestion is that the genuine uses of language are the sentential utterances. (Thus Davidson 1984, p. 60: 'In defining sentencehood what we capture, roughly, is the idea of an independently meaningful expression.') In order to apply this suggestion we have to know what to count as a sentence. If the notion of truth is already at hand, it may suggest the following solution: a sentence is the bearer of a truth value. The significant utterances are those that are capable of being true or false: in other words (or so it seems), the assertions, or statements.

In practice, this route to the primacy of assertion is overshadowed by another. The notion of truth not only promises to delineate the raw material for a systematic theory of language; it also offers an approach to problem of meaning – an approach developed since the 1960s in the influential work of Davidson and his followers. (See particularly Davidson 1984, essays 1 to 5; and Evans and McDowell (eds) 1976.) Davidson proposes that a systematic theory of meaning for a language – a theory that relates meaning to syntactic structure, in such a way as to show how the meanings of complex sentences depend on those of their constituents – should be a systematic theory of the truth conditions of the sentences of the language concerned. He suggests also that this truth theory be modelled on Tarski's definition of truth for formal languages. Such a theory is bound to accord a central place to those linguistic expressions that can be said to have truth values (and hence truth conditions). At first sight, indeed, there seems to be room for nothing else.

It may seem then that whichever route it takes to the doctrine that assertion is the primary use of language, a Fregean theory of language should be embarrassed by the existence of utterances such as questions and commands. These would appear to be complete, independent, meaningful uses of language, and yet not to be assertions – to lack truth values, for example. However, there are at least two ways in which a Fregean theory might deal with these cases. One way, obviously, would be to claim that questions, commands and the like do have truth values after all. Questions and commands would thus be construed as special kinds of assertions – perhaps as reflexive assertions about their speakers. As we shall see in chapter 3, this option has been explored by some of the proponents of the truth-conditional approach to general semantics.

On the whole, however, advocates of Fregean general semantics have preferred a less radical solution, turning on the Fregean distinction between sense and force. Roughly, the idea is to confine the significance of the notion of a truth condition to the determination of a certain core factor in the meaning of an utterance (its *sense*); and to regard the complete meaning of an utterance as a product of this factor and another (its *force*). Thus, the meaning of the question ‘Has the dog been fed?’ will be explained as a product of the interrogative force (here indicated by the interrogative mood) and of a sense that the sentence shares with its assertoric transform, ‘The dog has been fed.’

This solution enables a Fregean theory of meaning to accommodate non-assertoric utterances. Indeed, in principle it allows for the assertoric mode of utterance to be treated as simply one among many. Each mode calls for an account of its peculiar *pragmatic* significance. The assertoric mode might be said to be distinguished by the convention that in using it, speakers aim to utter true sentences; the imperative, by the convention that speakers utter sentences they require their audience to make true; and so on. In principle, the list could be very long indeed.

In practice, however, the theorist has a strong incentive to peg the boundaries of the class of assertions as widely as possible, at the expense of non-assertoric varieties of force – to expand the semantic component of a theory of meaning at the expense of the pragmatic. Faced with a new linguistic expression, the theorist has to decide whether its contribution to the meaning of a sentence consists in a modification of sense or an effect on force. The former answer has the advantage, for a Fregean, that it brings the case within the scope of the core component of the general theory of meaning. It avoids the need for a further addition to the supplementary theory of force. The latter answer, on the other hand, is a threat to the simplicity and comprehensiveness of the Fregean approach.

Most seriously, however, the latter answer would undermine the Fregean’s claim to provide a systematic account of meaning. As Davidson emphasizes, the great advantage of the truth-theoretic approach is that it promises to account for the finite-based ‘creativity’ of language use – the ability of speakers to use and interpret a potentially infinite number of novel utterances. It does so by linking syntactic structure to conditions for truth. The significance of sentence-forming components is explained in terms of their contribution to the truth conditions of sentences in which they occur. On this view then, the open-endedness of language simply mirrors the open-endedness of possible truth conditions. This elegant solution would be crippled by the admission of a dimension of linguistic creativity not matched to a range of truth conditions. If sentences expressing moral judgements do not have truth values, for example, then the infinity of novel utterances we can generate from moral sentences (using the ordinary logical connectives, say) cannot be matched to a corresponding infinity of truth conditions; and our capacity to understand such sentences cannot be explained in terms of such a correspondence.

So if creativity is to be confined to the semantic part of a theory of meaning, non-assertoric expressions can be tolerated only on the condition that they do not proliferate. Like questions and commands, non-assertoric expressions must not significantly combine with each other (or indeed with assertoric sentences) to generate further such expressions. However, this condition clearly excludes the cases of interest to non-factualists: expressions that behave like fact-stating sentences, at least to the extent of participating in the usual modes of supra-sentential complexity. Truth-conditional

general semantics thus has a powerful incentive to oppose non-factualism in general, and to dispute its claims in individual cases. The same applies, incidentally, if truth is replaced with a notion such as ‘assertibility’; the resulting theory will have the same reason to confine creativity to what it regards as the non-pragmatic component of a theory of meaning.

As a distinctive product of the linguistic movement in philosophy, non-factualism was thus overwhelmed by developments in the philosophy of language itself: by the growing interest in the project of a general theory of meaning, and the growing popularity of a Fregean approach to that project. This conflict has not had the attention it deserves, I think. Proponents of truth-conditional general semantics, and indeed of other Fregean approaches to a general theory of meaning, have shown a degree of professional blindness to the attractions of non-factualist treatments of particular topics; while on their side, non-factualists often seem to have been unaware of the problem of reconciling their theories with a workable approach to general semantics.

There are signs of renewed interest in non-factualism, of a new respectability. At least one source of this interest lies within the philosophy of language itself: in recent interest in Wittgenstein's later views on meaning, and in particular in the suggestion (notably by Kripke 1982) that the notion of meaning itself might require a non-factualist treatment. However, it seems to me that without proper attention to its foundations, and to its place in a general semantic programme, the approach will prove no more durable at the end of the century than it was in the middle. This book, then, is in part an investigation of these foundations, in the light of the conflict just described.

1.3 THE NATURE OF STATEMENT: CONFLICTING CONSTRAINTS AND A SCEPTICAL PROGRAMME

As I noted earlier, some Fregeans think that all discourse is assertoric (and hence, in effect, that all language is fact-stating); but most prefer to distinguish assertions from utterances such as questions and commands. The more common view is thus that the fact-stating/non-fact-stating distinction roughly corresponds to the English indicative/non-indicative distinction.

A major advantage of this view has been thought to be that it gives us an account of grammatical mood. Mood is a striking feature of natural languages, whose existence we might expect a theory of language to be able to explain. It has been widely assumed that an adequate explanation would come in two parts. The first would identify some property characteristic of just those speech acts for which we do employ the indicative mood; while the second would consist in the claim that, by convention, the indicative mood is used to ‘mark’ the performance of a linguistic act with this particular property. This view of mood – the *marker view*, as I shall call it – thus takes the indicative mood to be the conventional sign of a certain property, or factor, in the meaning of an utterance. The following remarks are representative; they come from a recent text on mood by the linguist F. R. Palmer.

It is undoubtedly the case that most, perhaps all, languages have a clear way of indicating that the speaker is making a statement that he believes to be true. This is what may be called the Declarative – the grammatical form that is typically used for such statements ... In languages that have systems of mood, the indicative ... typically indicates a declarative. (Palmer 1986, pp. 26–7)

Where a language has an indicative and an imperative mood, these are the formal grammatical markers associated with the notions of statement and [command], though there is no exact one-to-one correspondence. (1986, p. 24)

As Palmer notes, the marker view needs to be qualified, in practice, to cope with various additional grammatical devices. In English, for example, stress and inflexion enable us to do many things with indicatives normally done with non-indicatives, and *vice versa*. However, a proponent of the marker view will take this to mean that in practice the marks are just more complicated than the initial

account recognized. For present purposes we may ignore such refinements. As long as we don't criticize the marker view on the grounds that it misses these complexities, we won't go wrong in criticizing the simpler version.

In drawing the bounds of assertoric discourse at the limits of the indicative mood, most Fregeans thus align themselves with the marker view; and marker theorists will often characterize their position in terms of the Fregean force categories. The views thus have a common interest in the elucidation of these categories; in particular, in an account of assertion in terms of which the status of a given utterance can be established. For both views want to defend the claim that the non-indicatives are not assertions; and to be able to show, *pace* the non-factualist, that all indicatives are assertions.

The marker view will be the second front for the argument of the next four chapters. Our main target will be non-factualism, as usually presented. However, I want to show that the untenability of this position can be traced to its dependence on the analytic presupposition – on the view that there is a real property of statementhood, about whose distribution in language we are liable to be misled. I think that expressed in other terms, non-factualism contains important insights about language. So I shall need to exclude an alternative response to the argument – that of rejecting non-factualism, while retaining its usual analytic footings. The marker view makes this response particularly appealing. In other words, we have to show that the failure of orthodox non-factualism is not attributable to the success of the marker view.

Fortunately, the same style of argument works on both fronts. In effect, indeed, we shall be able to encompass both targets in a single advance, arguing first that the available analyses do not enable the non-factualist to hold a line short of the marker view itself; and secondly that they do not enable the marker theorist to hold the line at the indicatives (and so to exclude the possibility that all discourse is factual discourse).

I emphasize that the conclusion we are aiming for is not that all discourse is factual discourse; merely that neither non-factualists nor marker theorists can exclude this possibility. The argument will thus be a sceptical critique of the analytic assumption these views have in common – a demonstration that in neither case can analysis serve its intended purpose, in justifying a particular division of language into fact-stating and non-fact-stating parts.

1.4 GUIDE TO PART I

The structure of the next four chapters is as follows. In view of the importance of the notion of truth in general semantics and to the proposal of part II, chapter 2 treats non-factualists and marker theorists in tandem, examining the suggestion that assertions be distinguished in virtue of their connection with the properties of truth and falsity. Both sides have found this idea appealing. I argue, however, that without a prior understanding of the nature of assertion, we cannot restrict the notion of truth in such a way that not every utterance has the appropriate connection with these properties.

With truth out of the way, it is then convenient to separate our targets. Chapter 3 concentrates on the marker view, arguing by way of *reductio* that if there were an underlying property of the kind the view requires – a property marked by the indicative mood – we could not exclude its possession by non-indicatives. This shows, in effect, that the line cannot be held at the limits of the indicative mood. If all indicatives are fact-stating, then there is no way to exclude the possibility that non-indicatives are also fact-stating. The remaining task is then to show that the line cannot be held within the indicative mood – that of the various possible accounts of the nature of statement, none suffices to show that some indicatives are not statements of fact.

In chapters 4 and 5, with this end in mind, we take the side of the non-factualist's opponent in particular cases. Chapter 4 argues that in two important and arguably representative cases (probability and moral judgement) such an opponent forces the non-factualist to rely on a psychological

characterization of statement of fact – on the claim, essentially, that statement is the expression of belief. Chapter 5 then examines some possible bases for the required distinction between beliefs and other propositional attitudes. We find that the attempt to make good this distinction leads the non-factualist once more to appeal to truth – in effect, to the proposal rejected in chapter 2. The chapter concludes with a look at some recent arguments that seem to favour non-factualism about conditionals – confirming that here, too, non-factualism lacks the basis it requires.

1.5 THE EXPLANATORY TURN

One option at this point would be to blame the problem for the failure of the analytic solution – to conclude that there is simply nothing of philosophical interest to be said about a distinction in language of the kind non-factualism requires. Such ‘quietism’ has an evident appeal: to dismiss a problem is to save oneself the trouble of finding a solution. It is the relief the lapsed theist might feel in being able to put theological problems aside. But of course, this advantage is like the claimed psychological benefits of religious faith itself. It is a motive for wanting a particular conclusion to be true, but not a reason for believing it so. In the present case there seem to be two good reasons to remain at least open-minded. Quietism cannot explain the significance of mood, or the continuing attraction of non-factualist approaches to particular philosophical topics.

Part II therefore explores a different response. This concentrates on the assumption that non-factualism needs to be couched as a claim concerning the distribution of a certain property in language. I think that the prevalence of this assumption among non-factualists is particularly a product of their native tendency to set aside linguistic appearances. Non-factualists want to say that appearances are deceptive, that assertions cannot readily be distinguished in virtue of their syntactical form. This is needed to explain how more traditional philosophical theories can have been led astray. And there must be something in the picture for these theories to have been wrong about – in other words, it seems, some underlying property that syntax can mislead us about. Ask what property this is, and one is committed to what I have called the analytic programme.

The solution, I think, is to recognize that there is room for deceptiveness at the syntactical level itself. An expression that behaves as an assertion in most respects and circumstances might none the less behave differently in minor respects or unusual circumstances. If the differences are then overshadowed by the likenesses, it is easy to see how interpreters might be led astray. With the deceptiveness of linguistic appearances explained in these terms, non-factualism doesn’t need to appeal to linguistic categories underlying usage. So the non-factualist no longer owes us an analysis of statementhood. Non-factualism simply doesn’t need to appeal to any such property. True, it ought to explain the syntactical anomalies to which it does now appeal – the features of usage that distinguish moral judgements, probabilistic claims, or whatever. As we shall see, however, these anomalies can be explained without being taken to reflect any underlying semantic distinction.

What are these anomalies in usage, and what is the standard pattern from which they depart? We shall see that both involve the notions of truth and falsity. Dealing first with the ordinary use of these notions, I shall propose that we explain it in functional and evolutionary terms. The main idea will be that in virtue of the normativity of truth and falsity, speakers are encouraged to resolve disagreements. This has long term survival advantages. On the whole, it improves the behavioural commitments with which language users meet the world. As I said earlier, this is not intended to be an analysis of truth. It is a genealogical theory, an explanation as to why a language community might come to possess such a notion. It is comparable for example to a biological view of the origins of morality – the view say that human societies tend to evolve such a value system in virtue of the communal advantages that flow from its regulative effects.

Returning to non-factualism, we shall see that this account of truth leaves room for certain borderline cases – cases in which the use of truth and falsity departs from the standard pattern. These variations turn out to be explicable in similar terms to the standard pattern itself. Moreover, they appear to

correlate well with the cases in which non-factualism has seemed attractive. The explanatory perspective thus provides a sympathetic reconstruction of non-factualist intuitions – while vindicating, in a sense, the popular conception that truth is at the basis of the fact-stating/non-fact-stating distinction. It also finds a place for the non-indicatives.

It will be necessary to defend this account against the charge that it merely offers another analysis of the nature of statement. An obvious source of such a charge will be the fact that the account professes to explain variations in a pattern of usage across language. Hence it needs apparently to appeal to some more basic differences between utterances, or possible uses of language. Could not these differences themselves be made the basis of a characterization of fact-stating discourse? The reply to this will turn particularly on holistic considerations – the crucial point being that the relevant underlying difference need not show up in the individual case.

We began with the conflict between non-factualism and the Fregean programme for general semantics over the scope of the fact-stating/non-fact-stating distinction. In the final chapter I shall indicate how this conflict is resolved by the proposed treatment of truth. I shall locate the proposed account in relation to some recent discussions of realism and objectivity. And I shall finish with some comments on the significance of the explanatory perspective in other areas – particularly its relevance to the problem of the status of folk psychology.

PART I

A Sceptic's Guide to the Matter of Fact

The Place of Truth

Truth has often seemed to be the key to an analysis of fact-stating discourse. Marker theorists and non-factualists have both been tempted by the idea that an utterance is genuinely assertoric or fact-stating just in case it stands in a particular relation to the properties of truth and falsity – for example, just in case it bears a truth value, or is capable of doing so. In this chapter I want to show that truth cannot be invoked in this way.

The argument will turn on the demands placed on truth by the requirement that it provide the base property for an *analysis* of the notion of statement of fact. I cannot emphasize too strongly that I am not suggesting that truth has no role to play in an adequate account of fact-stating discourse. On the contrary, truth will be central to the theory to be developed in part II. But that theory will not be an analysis of statement of fact. Instead it will offer a non-reductive explanation of the characteristic marks in usage of the distinction between fact-stating and non-fact-stating discourse – including, perhaps most characteristically, the use of the notions of truth and falsity.

The present chapter goes as follows. I begin by clarifying the demands on the notion of truth that flow from its attempted use by marker theorists and non-factualists to elucidate the notion of assertion. With reference to familiar theories of truth, I then illustrate some of the ways in which a theory of truth can fail to meet these demands (fail to provide a *substantial* theory of truth, as we shall say). I offer some general arguments against the possibility of a substantial theory of truth, and conclude by examining a specific proposal: the approach suggested by David Wiggins in a paper called ‘What would be a substantial theory of truth?’ (1980). Wiggins’s approach – I think the most explicit attempt in the literature to tailor a theory of truth to these purposes – has several features of interest, from our present point of view. For one thing it is grounded in the truth-theoretic approach to general semantics, whose constitutional aversion to non-factualism we noted in chapter 1. Wiggins is aware of this tension, observing that the notion of truth he derives from truth-theoretic considerations – a notion applicable to the indicatives as a whole – leaves little scope for the more restricted notion the non-factualist requires. In showing that such a notion of truth cannot in fact be counted substantial, we shall defend non-factualism against this objection and clear the ground for its promised reconciliation with general semantics. For another thing, Wiggins’s theory sets useful targets for any alternative account of truth. Wiggins identifies properties or ‘marks’ that are plausible characteristics of any satisfactory account of truth. We shall see that his own derivations of these marks are illicit, given the intended role of a substantial theory of truth. But the marks themselves will set a standard for our own explanatory theory of truth. We shall want to be able to explain why truth has these particular characteristics.

2.1 TWO WAYS OF APPEALING TO TRUTH

In attempting to characterize assertion in terms of truth, marker theorists and non-factualists must both aim for more than a merely descriptive account of a distinction found in usage. Neither side can simply observe that in practice we apply the notions of truth and falsity selectively – that only certain kinds of utterance are treated as capable of possessing these properties. In the marker theorist’s case, this is because the marker theory itself is an attempt to explain a distinction we observe in usage. It may be illuminating to describe this distinction further – to note that in practice the use of the indicative mood correlates with that of ‘true’, ‘false’ and related terms. But to gloss the problem in this way is not to provide a solution. The ordinary use of truth is as much in need of explanation as the indicative mood itself.

In the non-factualist's case a merely descriptive characterization would be inadequate for a different reason. The non-factualist's interest in usage is not so much to explain it as to escape it. Non-factualism urges us to draw distinctions that are missed by ordinary usage. Its characteristic claim is that we are misled by usage into mistaking the real nature of a certain kind of discourse (typically with philosophical consequences). True, the non-factualist might say that we are misled only by insufficient attention to usage – that if we concentrate on the details, we do find evidence of the distinctions in question. But evidence of what? Even in this case the non-factualist assumes that the usage distinctions in question reflect the deeper distinctions on which the philosophical moral depends. The point is made by Michael Dummett. In his (1958) paper on truth Dummett considers the suggestion that an appeal to the use of 'true' and 'false' can settle non-factualist claims, for example about the status of ethical statements. 'At one time', he says,

it was usual to say that we do not call ethical statements 'true' or 'false', and from this many consequences for ethics were held to flow. But the question is not whether these words are in practice applied to ethical statements, but whether, if they were so applied, the point of doing so would be the same as the point of applying them to statements of other kinds, and, if not, in what ways it would be different. (1958, p. 144)

Non-factualists and marker theorists should therefore not be construed, in trying to characterize assertion in terms of truth, as merely labelling a distinction they claim to find in ordinary usage. Their appeal is not to the ordinary use of 'true', 'false' and related terms, but to the properties of truth and falsity. Their common claim is that an assertion (or statement of fact, or whatever) is an utterance that stands to these properties in some distinctive way.

At this point the possible views begin to diverge, in two senses. First, the non-factualist and the marker theorist clearly apply this characterization in different ways. The marker theorist takes standing in the specified way to truth and falsity to be the property conventionally marked by the indicative mood (and as we shall see, offers parallel accounts of the other moods); whereas the non-factualist claims that not all indicatives stand in this way to truth and falsity. Secondly, there are a number of views of the precise role of the properties of truth and falsity in a characterization of this general kind. Is it that assertions *have* these properties, that they are *capable* of having them, that they are governed by the convention that we *try to ensure* that they have one of them (i.e. truth), or some yet further connection?

These various elaborations have this much in common, however: they all depend on truth and falsity, and hence on the availability, at least in principle, of some account of the nature of these properties. The point is a trivial one. Unless the property of truth is subject to some theoretical constraints there is nothing to exclude its application not merely to the *utterances* the non-factualist or marker theorist wants to regard as non-fact-stating, but to any object whatsoever. Without a theory of truth, it is not simply that we cannot exclude the possibility that imperatives or ethical judgements have truth values (or stand in some other important relation to truth); we cannot exclude the corresponding possibility for trees, rocks and clusters of galaxies. (It is no good appealing to usage. We have seen that non-factualists and marker theorists both need more than that.)

Of course, not just any theory of truth will do. Again, one only needs to consider some trivial possibilities. Let us say that if a theory of truth can be invoked to characterize fact-stating discourse in the way that either non-factualists or marker theorists require, then it is a *substantial* theory of truth. (I think the term was first used in this way by Bernard Williams 1973, p. 202). I want to show that there can be no such theory.

2.2 TWO WAYS FOR TRUTH TO BE INSUBSTANTIAL

One of the most explicit versions of a truth-based formulation of the marker view is that of Stenius, who says that 'we may ... formulate the rule for the indicative in the following way: *Produce a*

sentence in the indicative mood only if its sentence-radical is true (1967, pp. 267–8). By ‘sentence - radical’ Stenius, following Wittgenstein, means a sentence considered as devoid of assertoric force. Thus:

Consider the following sentences:

- (1) You live here now.
- (2) Live (you) here now!
- (3) Do you live here now?

These three sentences have something in common, which ... I call the *sentence-radical*; what is different in each of them I call the *modal element*. The sentence-radical signifies the *descriptive content* of the sentence, the modal element signifies its *mood*. (1967, p. 254)

Stenius suggests the following rules for the use of the imperative and interrogative moods: ‘React to a sentence in the imperative mood by making the sentence-radical true’, and ‘Answer the question by “yes” or “no”, according as its sentence-radical is true or false.’

Let us focus on the possible justification for a theory of this kind. If we want to appeal to a semantic property to explain the significance of the indicative mood, we need to be sure that we can tell whether an utterance has that property, without appealing to its grammatical mood. A theory such as Stenius’s cannot simply take for granted that certain sentence-forming expressions are mood indicators, and others are modifiers of descriptive content. Rather it needs to justify the particular dividing line it claims to find in language.

This point is liable to be overlooked, if – in the grip of the orthodox view – we fail to notice that there are alternatives. But consider negation, for example. Why not think of ‘It is not the case that ...’ as a mood indicator? It will be governed, of course, by this rule: Produce a sentence in the negative mood only if its sentence radical is *false*. Or on the other side, why not think of the imperative as a content modifier (so that imperative utterances have assertoric force)? We cannot rule out these possibilities on the grounds that negatives but not imperatives are indicatives, if what we want to explain is the significance of the indicative construction itself.

These considerations place the following constraint on a substantial theory of truth. It must not take for granted that only indicative utterances have truth values, or stand in the requisite relation to the properties of truth and falsity. Yet at the same time it must exclude the extension of these properties to non-indicatives. (The non-factualist will want the bounds to be narrower still.) We shall see that there is a tendency to avoid the latter mistake at the cost of the former.

The second major constraint on a substantial theory of truth is merely an instance of a general requirement for acceptable analysis: an account of a base property cannot itself depend on the very notion that property is intended to elucidate. A substantial theory of truth cannot rely on the notion of assertion, for example. In particular, it cannot rely on the otherwise appealing intuition that a sentence is true just in case it would be correct to assert it. For that would reduce the supposedly distinctive claim that assertions are governed by the convention that speakers aim to speak the truth, to the truism that assertions are governed by the convention that speakers aim to make them correctly. Any convention-governed activity is governed by that convention.

Stenius is bothered by this sort of point, and suggests that

there are two concepts of truth: (i) the concept of truth ... which refers to the sentence-radical – I shall call this kind of truth *descriptive truth*; (ii) a concept of truth referring to a sentence in the indicative – I shall call this kind of truth *modal truth*.

When stating that somebody ‘speaks the truth’ one does not mean that he is producing a true sentence-radical – for one does not say that somebody is speaking the truth if he presents a true sentence-radical in the imperative or interrogative mood The expression ‘speaking the truth’ thus refers to the modal concept of truth, and means that one is following the rule for the indicative correctly. This moral rule is often motivated by the argument that by not following it one is undermining the communicative function of language, which morals must safeguard. (1967, p. 268)

As it happens, of course, to use the indicative correctly is to utter a sentence radical which is true in the descriptive sense; and so the two kinds of truth seem co-extensive. On Stenius’s view this seems to be a kind of accident, stemming from the rule which happens to govern the indicative mood. But there is a danger here: the more the modal notion of truth is detached from the actual use of the indicative mood, the more it needs to be explained why the modal notion should not extend to other moods. Why, for example, should we not say that someone ‘acts the truth’ when he or she obeys a command, thus following the rule for the imperative correctly? The modal notion of truth thus floats free of the indicative mood, while we are still in need of a substantial theory of the descriptive notion of truth.

In summary then, there are two main ways in which a theory of truth can fail to be substantial. It can ‘get the grammar wrong’, either illicitly assuming that only indicatives stand appropriately to truth or failing to exclude the possibility that non-indicatives so stand. Or it can appeal to the sorts of notions it is supposed to elucidate (or, not quite such a serious sin, to notions that would do a better job by themselves). We shall see that the first disability is endemic to the proposal to analyse assertion in terms of truth, though it is liable to be disguised by the second.

It is easy to find both kinds of insubstantiality in popular accounts of truth. This is not an objection to the accounts concerned, of course. It simply means that they cannot be applied, as in general they were not intended to be applied, in the special role of a substantial theory.

Thus traditional correspondence and coherence accounts of truth are likely to embody the second form of insubstantiality. To explain truth in terms of ‘correspondence to the facts’ is to take for granted the very notion we are supposing that the non-factualist wants to elucidate (and so restrict) by means of a substantial theory of truth. For example, it makes the issue as to whether there are genuine moral truths – whether moral claims can properly be said to be true or false – depend on that of whether there are moral facts. If we could already settle the latter issue, we wouldn’t need to characterize fact-stating discourse in terms of truth. Similarly, it is unlikely to be helpful to try to explain truth in terms of coherence with a system of beliefs. In the moral case, that leads us back to the question whether moral attitudes are genuine beliefs. But given the notion of belief, we would be able to use it to characterize assertion (in terms of expression of belief), and the excursion through truth would be redundant.

The first kind of insubstantiality is likely to be displayed by any account of truth that trades on the equivalence property of truth: on the fact that for any (‘appropriate’) sentence P, the sentences P and ‘P is true’ are in some sense equivalent in meaning. The most familiar such accounts are the ‘redundancy’ and ‘disquotational’ theories of truth (Ramsey 1978, chapter 2; Quine, 1970, chapter 1), for which this property characterizes truth. A less familiar case is Strawson’s (1949) ‘performative’ account of truth. This turns on the application we make of the equivalence property – on the fact that by saying ‘That’s true’ we can endorse an utterance made by another speaker. The novelty of the approach lies in taking this endorsing or confirmatory speech function to be the primary characteristic of truth; or as the view would think more accurate, of the use of the term ‘true’. (This view is also known, rather aptly, as the ‘amen’ theory of truth.)

The insubstantiality of such theories stems from their reliance on the assumption that the equivalence property involves a restriction of truth to indicative sentences. It is not in fact the case that for *any*

sentence P, P and 'P is true' are equivalent in meaning. Manifestly, 'Feed the dog' does not mean the same as "'Feed the dog' is true.' Hence the restriction to 'appropriate' sentences. But which are the appropriate sentences? Why, those that grammar permits us to substitute for P in the construction 'It is true that P'; in other words (for our purposes), the indicatives. So the equivalence property distinguishes the indicatives only by fiat. A theory of truth which rests on the equivalence property either fails to pick out the indicatives (or any sub-class thereof) at all, or does so only because the restriction of truth to such a class of utterances is assumed from the start. But if the latter, then we can again invoke the sceptical challenge: Why this restriction? Such a notion of truth does not therefore provide the base property that is supposed to be characteristic of assertion and marked by the indicative mood.

The fact that Strawson's theory is insubstantial in this way is noted by Bernard Williams, in the paper in which he introduces the idea of a substantial theory of truth. Williams says that he doesn't see 'how on such a theory it could be more than an accident of language that "is true" signified agreement with *assertions* rather than agreement with anything else' (1973, p. 203). In practice we cannot endorse an apt question or an appropriate command by calling it true. Of course, we can do it in less direct ways, but so we could in all the cases in which we can just say 'That's true.' The amen theory does not explain why we should be allowed the conversational convenience in some cases but not in others.

The above examples suggest that a substantial theory of truth would be an unusual theory, by existing standards. The common approaches seem to be insubstantial in one or both of the ways we identified. Could there be such a theory? I think we can show that there could not.

2.3 TWO ARGUMENTS THAT TRUTH CANNOT BE SUBSTANTIAL

The first argument generalizes the sceptical challenge. Marker theorists and non-factualists both claim to find a division in language between fact-stating and non-fact-stating uses. A substantial theory of truth would permit a particular account of this distinction. It would allow the theorist to say that the fact-stating uses are distinctive in standing in a certain (yet to be specified) relationship to the properties of truth and falsity. Given such a proposal, the sceptic challenges the theorist's warrant for assigning any given utterance to one side of the division or other.

For simplicity let us focus on the claim that marker theorists and non-factualists will have in common, that non-indicatives do not stand in the designated way to truth. How could such a claim be justified? Consider an example. On what grounds do we say that sentences such as those in this paragraph do not have the property supposedly marked by (or confined to some sub-class of) the indicatives? In other words, how do we know that these sentences are not fact-stating?

Ordinarily we might appeal to their mood, but what is now at issue is our right to the kind of theory that would allow such an appeal. Perhaps then we should appeal to the intuition that in uttering such sentences we do not express beliefs, make any claims about the world, and the like. However, to depend on claims of this sort is to render our theory of truth insubstantial in the second of the two ways we distinguished earlier. A substantial theory of truth cannot rely on alternative characterizations of fact-stating discourse.

The last alternative seems to be to appeal to ordinary usage – to the fact that the sentences concerned are not ordinarily ascribed truth values. We have noted, however, that in distinguishing fact-stating from non-fact-stating discourse, marker theorists and non-factualists both want more than a merely descriptive characterization of ordinary usage. Among other things a substantial theory of truth is supposed to provide an analytically grounded explanation of usage. But it seems that there can be no evidence for such a theory, other than the merely linguistic phenomena it is intended to explain.

It might be objected that this is a common predicament for scientific theory. It seems that the grounds for acceptance of a theory often consist simply in the fact that it offers the best available explanation of a range of phenomena. Such cases lack a crucial feature of the present linguistic case, however. In trying to account for distinctions in usage in terms of the fact-stating/non-fact-stating distinction, marker theorists and non-factualists both seem to have regarded the relevant features of usage as responsive to an underlying distinction. The marker theory, in particular, seems to involve that the choice of grammar be guided by at least an implicit recognition of underlying linguistic categories. The sceptical argument challenges the possibility of such recognition (given, in the present case, that these categories are supposed to rest on a substantial theory of truth). There is no parallel to this recognitional requirement in the scientific cases, and hence no easy answer to the sceptic by analogy with those cases.

The role of implicit recognition of the underlying categories is well illustrated by Stenius's version of a truth-based marker theory. We saw earlier that if assertions are to be characterized in terms of the convention that they aim at truth, truth must be considered a property of the members of a class of unasserted sentences: Stenius's sentence radicals. Making an assertion is then to be thought of as uttering a sentence radical, while indicating by means of its (indicative) mood that one takes it to be true. There are parallel conventions for the other moods. However, such conventions require that speakers be able to distinguish the sentence radicals in advance of the choice of mood. The sceptic argues that this is impossible (at least in terms of a substantial theory of truth). Our only guide to the limits of the class of sentence radicals is the indicative mood itself. So the convention that supposedly governs that mood turns out to presuppose an understanding of the significance of mood.

The second and perhaps more familiar argument against the possibility of a substantial theory of truth also turns on the fact that such a theory makes truth conceptually prior to assertion. If fact-stating constructions are characterized as sentences that bear a certain relationship to the properties of truth and falsity then the distinction between truth and falsity comes into question. For example, couldn't we imagine a group of speakers who followed the convention that they aimed to utter *false* sentences. This convention seems no less useful than the ordinary one for purposes of communication. If one of these speakers says 'I've made the tea', then knowing the convention we know what he means. He means that he has not made the tea. However, what he *said* was supposed to be false. (We were imagining that he follows the convention of trying to utter false sentences.) But the truth value of a sentence depends on what it means. If he meant that he had not made the tea, then if we agree we should say that the sentence he uttered was true, not false. 'For a sentence is true if things are as we say they are by means of it; and if by "*p*" we mean " $\sim p$ ", and things are as we mean them, then "*p*" is on the new interpretation true and not false.' The sentence comes from the *Tractatus* (4.062). Dummett (1973, p. 318) refers to it in support of the contention that assertion cannot be analysed in terms of truth. Indeed, in his (1958) paper on truth Dummett criticizes 'Frege's theory of truth and falsity as the references of sentences' on the grounds that it 'leaves ... quite out of account' that

it is part of the concept of *truth* that we aim at making true *statements* Frege indeed tried to bring it in afterwards, in his theory of assertion – but too late; for the sense of the sentence is not given in advance of our going in for the activity of asserting, since otherwise there could be people who expressed the same thoughts but went in instead for denying them. (1958, p. 143, my italics)

The two arguments of this section show the absurdity of Stenius's 'descriptive' notion of truth – of the idea that sentences could be considered to have truth values (or indeed meanings) in advance of an understanding of their conventional linguistic role. On the contrary: a sentence can be taken to have a significant truth value only in virtue of the understanding that its conventional use is such as to admit a standard of right and wrong, of correctness and incorrectness. (This is a minimum requirement; presumably not just any standard will do.) The true sentences are those whose utterance would be correct, according to this standard. In other words (now taking for granted what we are trying to explain), a true sentence is true in virtue of the fact that it would be correct to assert it – which reduces

the claim that in making assertions we aim to speak the truth, to the truism that in making assertions, we aim to do so correctly.

Dummett criticizes Stenius in similar terms, but seems to think that the point concerns only the distinction between truth and falsity:

Stenius insists that the ‘sentence-radical’ ... itself has a ‘*directed*’ sense, independently of the assertion which can be expressed by attaching the assertion sign to it. Admittedly, the ‘sentence-radical’ has ‘true/false poles’, in that its sense consists in the distinction between conditions under which it is true and conditions under which it is false: but we can determine the ‘direction’ of this sense ... only by considering its use when assertoric force is attached to it, because it is only by reference to its role in assertions that we can identify one set of conditions as those in which it is *true* and the other as those in which it is *false*. (1973, p. 327)

However, the sense of a sentence radical seems equally dependent on its role in assertion. Dummett’s own views on truth and sense suggest as much: if ‘the notion of truth has first to be explained’ in terms of ‘a distinction between correct and incorrect assertions’ (1978, p. 20), and ‘sense consists in a distinction between conditions under which [a sentence radical] is true and conditions under which it is false’ (1973, p. 327), then a grasp of sense already depends on a conception of potential assertoric use.

2.4 TRUTH THEORIES: INSUBSTANTIAL THEORIES OF TRUTH

Let us turn now to the promised example: David Wiggins’s (1980) attempt to construct a substantial theory of truth. ‘What would be a substantial theory of truth?’ is Wiggins’s contribution to a *Festschrift* for P. F. Strawson. He begins, appropriately, with some remarks about Strawson’s views on truth. In particular, he claims (1980, pp. 190–1) to find in Strawson’s (1950) article on truth the view that ‘of proper concern to the theory of truth’ is ‘the task of elucidating the nature of a certain type of communication’: the ‘empirically informative’ or ‘fact-stating’ use. He thus counts Strawson as an early advocate of the need for what Williams was to call a substantial theory of truth: ‘I do not doubt that what [Williams] felt he needed was the very same thing that Strawson had envisaged in 1950’ (Wiggins 1980, p. 192).

Strawson does refer to the task just mentioned. However, when he asks ‘why should the problem of Truth ... be seen as this problem of elucidating the fact-stating type of discourse?’ his ‘answer is that it shouldn’t be The problem about the use of “true” is to see how this word fits into [the fact-stating] frame of discourse. The surest route to the wrong answer is to confuse this problem with the question: What type of discourse is this?’ (1950, pp. 142–3). At that time, at any rate (there is no hint of it in his reply to Wiggins in van Straaten 1980, pp. 294–6), Strawson’s view thus seems to have been that the problem of truth arises only *within* the fact-stating form of discourse, and hence depends on, or assumes, an understanding of that form of discourse.

Wiggins also claims Dummett as an early advocate of Williams’s project, quoting from Dummett’s (1958) paper on truth. But we have seen that in that paper and elsewhere Dummett rejects the idea that assertion can be explained in terms of truth. He again emphasizes the point in a postscript to the (1958) paper, first published in 1972:

What has to be added to a truth definition for the sentences of a language, if the notion of truth is to be explained, is a description of the linguistic activity of making assertions; and this is a task of enormous complexity. What we can say is that any such account of what assertion is must introduce a distinction between correct and incorrect assertions, and that it is in terms of that distinction that the notion of truth has first to be explained. (Dummett 1978, p. 20)

We shall see in chapter 3 that Dummett prefers a quite different account of the distinctive features of assertion.

History aside, Wiggins's approach is to consider what must be true of truth if the predicate 'is true' is to figure in a truth theory adequate for radical interpretation. More precisely, the idea is to look at theories in the style of a Tarski truth theory for a language L, yielding equivalences of the form

(2.1) S is ϕ if and only if P.

Such a theory is to be regarded as specifying the meaning of the L-sentence S. There are many such theories, but some of them make better sense than others of the speakers of L. The constraints on ϕ imposed by the anthropological constraint of 'making sense' are the marks of the notion of truth.

I am going to argue that the distinction that a substantial theory of truth is supposed to help us with – between assertoric or fact-stating uses of language, and all the rest – is already built into the truth-theoretic approach to meaning, as Wiggins employs it. Wiggins gets out only what Davidson and others have already built in. The project of radical interpretation, even when constrained by the structural insights we have from Tarski, does not in itself yield either the Fregean distinction between sense and force, or (what would come to the same thing) a substantial theory of truth. These have to come from somewhere else.

Wiggins's strategy does not consist simply in a Tarskian definition of truth. Tarski's approach takes for granted the notion of translation. His 'condition of adequacy' is that a truth theory should yield sentences of the form

(2.2) S is true if and only if P

where P is a translation in the metalanguage of the object language sentence to which the description S refers. Whereas for Davidson and the truth theorists who follow him, the priorities are in a sense reversed. The T-sentences are now expected to yield interpretations of the sentences of the object language, and the theory is constrained, roughly speaking, by the requirement that it make good anthropological sense of the speakers of the object language. Moreover, it is not crucial to this programme (as it was to Tarski) that the predicate the theory invokes be the predicate 'is true'. John McDowell has repeatedly emphasized this point. McDowell describes what he calls 'the best version of the truth-conditional proposal ... along these lines':

We may reasonably set ourselves the ideal of constructing, as a component of a complete theory of meaning for a language, a sub-theory which is to serve to specify the contents of (for instance, and surely centrally) assertions which could be made by uttering the language's indicative sentences [A] direct assault on that task would be to look for a sub-theory which generates, on the basis of structure in the object-language sentences, a theorem, for every appropriate sentence, of this form: 's can be used to assert that p'. Now there is a truistic connection between the content of an assertion and a familiar notion of truth ...; the connection guarantees, as the merest platitude, that a correct specification of what can be asserted, by the assertoric utterance of a sentence, cannot but be a specification of a condition under which the sentence is true. A radical proposal at this point would be as follows: as long as the ends of the theorems (think of them as having the form 's... P') are so related that, whatever the theorems actually say, we can use them as if they said something of the form 's can be used to assert that p', it does not actually matter if we write, between those ends, something else which yields a truth in the same circumstances; our platitude guarantees that 'is true if and only if' fits that bill, and this gives a more tractable target than that of the direct assault. (1981, pp. 228–9)

McDowell here takes for granted the association between assertoric force and the indicative mood. In speaking of the indicative sentences of an arbitrary language, he clearly has in mind the assertoric sentences. The idea is thus that the truth theory (or rather, what McDowell says may as well be regarded as a truth theory) yields an account of the descriptive content, or sense, of what are effectively the sentence radicals of the language in question. For a complete theory of meaning, we need to add to this content-specifying sub-theory a theory of force – an account of the various different sorts of things that speakers of the language do with sentence contents (including, of course, asserting them).

As we noted in chapter 1, the extent of this additional theory of force depends on how many varieties of force there are reckoned to be in the object language in question. This suggests that other things being equal it would be better to treat differences of meaning as differences of sense than as differences of force; for this way they are incorporated into the theory of sense, and don't have to be dealt with later. Every effort should be made to handle a difference of meaning *within* the sort of sub-theory that McDowell describes, before consigning it to the assortment of leftovers which will have to be dealt with when that sub-theory is complete. Of course, if we had some criterion for whether a difference of meaning was a matter of sense or a matter of force, there would be no question of *trying* to accommodate it in the sub-theory; it would be in or out, according to what the criterion told us. But at the moment we have no criterion. A substantial theory of truth is supposed to give us one.

Moreover, McDowell's care in distinguishing a theory whose primary role is to state the contents of sentences of an object language from a theory of truth, removes what may have seemed a barrier to accommodating certain sorts of differences of meaning within a theory of this general form. The standard 'filling' for a T-sentence – 'is true if and only if – suggests that what goes into the blanks must at the left hand end designate the sort of object language sentence to which the predicate 'true' applies; and at the right hand end be the sort of metalanguage sentence that can form the antecedent or consequent of a material conditional. But McDowell's insistence that the standard filling is not obligatory should alert us to the possibility of a filling that does the same work, without imposing such a grammatical restriction; and that hence enables the same process of meaning specification to extend to what are traditionally regarded as matters of force.

McDowell suggests the filling

(2.3) ... can be used to assert that

Of course, he has already restricted the theory to assertions – to 'the language's indicative sentences'. Except in that it actually hints at this restriction, 2.3 is no different from

(2.4) ... can be used to say that

However, 2.3 and 2.4 might still seem to restrict the kinds of sentences we can use to fill the blanks. On the right-hand side, in particular, we are still effectively restricted to indicative sentences; for English grammar won't allow anything else to fill the gap in this sort of 'that' clause.

At this point, however, we can appeal to Davidson's own (1968) 'paratactic' analysis of this kind of grammatical construction. On this analysis the logical form of the sentence 'Ramsey said that truth is redundant' is

(2.5) Truth is redundant. Ramsey said that.

This allows that the clause which appears after 'that' in the original sentence is *used*, rather than *mentioned*, and hence makes sense of the intuition that in translating this remark about Ramsey into another language, we should not leave this clause in English. (We are not *quoting* Ramsey.)

The present relevance of this proposal is that it frees us from grammatical restrictions on the nature of ‘that ...’ clauses.

(2.6) Make truth redundant. Ramsey said that.

makes as much sense as 2.5. So indeed does

(2.7) Is truth redundant? Ramsey said that.

Of course, in this case we would be inclined to say ‘Ramsey *asked* that’, but this is incidental. The important point is that just as in 2.5, the imperative and interrogative sentences in 2.6 and 2.7 are used rather than mentioned. We are not reporting Ramsey’s exact words. We are *samesaying*, as Davidson (1968, p. 140) puts it, in our own words.

Applying the same analysis to 2.4, the ‘filler’ we derived from McDowell, we find that the theorems of resulting sub-theory can be regarded as having the form

(2.8) S can be used to say this: P,

and that there is now nothing in the form of these theorems to prevent the sub-theory from extending to the non-indicative sentences of the object language. If the object language contains imperatives, for example, there is now nothing to prevent the sub-theory giving us their meaning in the same way it does for indicatives: by *using* a metalanguage sentence, and telling us that this sentence says the same thing as the given sentence of the object language. Note that the fact that the metalanguage sentence is still used rather than mentioned means that the theory is still distinguished from a translation manual in the way that truth theorists have been keen to stress. (See, for example, Evans and McDowell 1976, p. x.)

It now appears that the preferred form for what was supposed to be a theory of the *content* of the sentences of the object language, makes no distinction between differences in content (or sense) and differences in force. The tasks initially consigned to the separate theory of force are in fact taken care of by the so-called sub-theory. The significance of the force indicators of the object language – its imperative and interrogative moods, for example – is explained in the same way as that of sense-modifying expressions: by the use of corresponding constructions in the metalanguage. The supplementary theory, originally thought of as a theory of force, survives, at most, in a force-neutral account of the activity of *saying*; a mere gloss to the main theory, whose main task would be to distinguish the full-blooded use of sentences from other sorts of appearances in writing and in speech.

This need not mean that there is no longer a place for the notion of force; only that the distinction between force and sense need not be reflected in the radical interpreter’s methodology, as usually supposed. As long as we draw such distinctions in the metalanguage, the process of interpretation will draw corresponding distinctions in the object language. But the process makes no intrinsic demands on such distinctions – which suggests that we can’t find an elucidation of the fact-stating/non-fact-stating distinction (or, what is supposed to come to the same thing, a substantial theory of truth) on the requirements of radical interpretation.

If this conclusion seems surprising, I think the impression stems from a distorted view of the aim of radical interpretation. Radical interpretation is most naturally seen as the programme of systematic translation of the sayings of others into the home language. ‘Sayings’ here means sayings of all kinds – as many kinds as we need, to understand what our subjects are saying. From this point of view the introduction of a Fregean sense/force distinction looks rather odd – a case of changing horses in mid-stream. So from the radical interpreter’s natural perspective there is an odd discontinuity in a truth-theoretic theory of meaning. We have seen however that truth-theoretic general semantics shows a

constitutional tendency to concentrate on assertoric uses of language. This emphasis obscures the discontinuity, because it obscures the fact that a truth-theoretic theory of meaning leaves us some way short of the interpreter's goal. *Saying* comes to be equated with *asserting*, and the radical interpreter's programme comes mistakenly to be identified with the task of specifying the contents of assertoric utterances.

I want to mention two objections to the above conclusion – two considerations that might be thought to show that a radical interpreter does have methodological grounds to distinguish indicatives from non-indicatives. (I discuss these and other objections at greater length in Price 1987.) The first stems from a popular view of the empirical grounds of a theory of radical interpretation: the view that 'the evidential base ... will consist of facts about the circumstances under which speakers hold sentences of their language to be true' (Davidson 1974, p. 320). It might be claimed that the radical interpreter's theory will need to embody a sub-theory of the kind McDowell describes, in order properly to reflect its empirical base.

The objection is easily turned, however. The available 'facts about circumstances under which speakers hold sentences of their language to be true' cannot exceed the available facts about circumstances under which speakers are prepared to utter these sentences, with serious intent. Adding the truth predicate doesn't help to distinguish the literal utterances from all the rest; and doesn't, for example, help us to decide whether reluctance to make an utterance is due to politeness or disbelief. The subject who is too polite to say, 'You linguists all look the same to me', will be too polite to assent to the suggestion that we linguists all look the same to him.

On the other hand, to insist on framing the evidence in terms of 'holding true' is to restrict one's gaze, quite unnecessarily, to those sentences that can be held true. It is not clear that such a restriction is possible – it depends on being able to distinguish (what will be interpreted as) the indicative sentences at a very early stage. But even if it is possible, there can be nothing to be gained by it – far better to collect facts about the circumstances in which sentences of all kinds are seriously uttered. The facts about 'holding true' can then be distilled, if necessary, when we've distinguished the indicatives – and the residue will do us nothing but good. (The interpreter's standard question will thus be, not 'Would you assent to this: P?', but 'Would it be appropriate to say this: P?')

Secondly, the truth theorist might object that the justification for restricting one's initial meaning-specifying theory to the indicatives lies in the advantages of expressing its theorems in the predicative form exemplified in 2.2. This point is best coupled with another. The main thing that recommends a truth-theoretic approach to a theory of meaning is its ability to deliver plausible recursive clauses for at least some of the sentence-forming operations of natural language; while at the same time keeping faith with the intuition that an adequate theory of meaning will give us the meanings of expressions of the object language by using, rather than mentioning, expressions of the metalanguage. It might be thought that these virtues justify what otherwise seems a needless complexity: the restriction of the meaning-specifying theory to indicative sentences, and the consequent introduction of a separate theory of force.

This argument depends on the assumption that it is impossible to express the theorems of our extended sub-theory in the predicative form of 2.1. In other words, it turns on the impossibility of a predicate ϕ such that for *any* sentence P, 'P is ϕ ' can be used to say that P – i.e. such that

(2.9) 'P is ϕ ' can be used to say this: P.

However, in a sense we have already encountered such a predicate, in Strawson's amen theory of truth. We saw that the amen theory does not explain why 'true' should not be used to endorse any sort of utterance whatsoever. It is thus insubstantial in the first of the ways we distinguished in section 2.2.

We noted that the same can be said of any theory of truth that relies on the equivalence property – on the fact that for any (appropriate) sentence P, P and ‘P is true’ are in some sense equivalent in meaning. However, it is precisely this feature of truth that is exploited by the present objection. For imagine that we could generalize truth to the non-indicatives, its significance being given by the generalized equivalence principle:

(2.10) For all sentences P, P is true if and only if P.

The theorems of our general meaning-specifying theory could then be written in the predicative form of 2.2, and the advantages of such a form could not be the grounds for a discontinuity in a programme of radical interpretation. Our generalized truth theory would do what an orthodox truth theory does, but do it more generally. Certainly, we could carve the orthodox theory out of the general one; but only if we already knew what to cut away.

So the objector needs a reason why we shouldn’t generalize the equivalence principle, and hence apply ‘true’ to non-indicatives. As we saw, however, such a reason will call for an independent characterization of the indicatives, in order to account for the required restriction. This defeats the avowed aim of a substantial theory of truth. There is surely some reason why we don’t use ‘true’ in this way. But unless we know more about truth than the equivalence property tells us, it is bound to be a reason to which a substantial theory of truth is not allowed to appeal.

In a footnote to the passage quoted above, McDowell remarks that ‘it is a philosophical issue whether there are respectable purposes for which a stronger notion of truth is required’ – a stronger notion, in other words, than that guaranteed by the ‘platitude’ that ‘... is true if and only if ...’ is a suitable filling for the theorems of the meaning-specifying sub-theory. McDowell notes that non-factualists will want a stronger notion, but says that he is ‘inclined to suppose that this is a matter not so much of an alternative notion of truth as of a characteristically philosophical misconception of the only notion of truth we really have: one which the platitude in fact suffices to determine’ (1981, p. 229, n. 9). We find, however, that the platitude does not account for the ordinary use of truth. The truth-theoretic approach to truth is insubstantial in the first of the ways we earlier distinguished. It picks out the indicatives (or any smaller class) only to the extent that it takes for granted that such utterances are the only bearers of truth. So McDowell’s platitude does not suffice to determine the fact that truth is not applied to the non-indicatives; and the non-factualist is not alone in wanting something stronger.

2.5 THE MARKS OF TRUTH

Wiggins claims to derive certain ‘marks of a theory of truth’ from the anthropological constraints on a Davidsonian programme of radical interpretation. However, we have seen that that programme does not yield the kind of distinction that a substantial theory of truth is supposed to underpin. If we already have the distinction, interpretation will reflect it; but it needs to come from somewhere else. What then of Wiggins’s marks of truth? Are they properties of truth that, like the equivalence property, could equally extend to non-indicatives? Or do they depend on some prior conception of the nature of fact-stating discourse – on the sort of conception that a substantial theory of truth is supposed to provide? The answer, I think, is the latter. Wiggins allows himself the notions of assertion and belief.

Thus ‘the first mark of assertibility (and so of truth)’ consists in ‘assertibility’s being *the primary dimension of assessment for sentences*, or that property which sentences have *normally* to be construed as aiming to enjoy’. If there were no such norm, Wiggins says, ‘interpretation would either be impossible ... or unjustifiable’ (1980, p. 205). Wiggins’s point seems to be that unless we take utterance to be constrained by such a norm, we shall be unable to take seriously anything that our speakers say. But only the implicit assumption that by ‘utterance’ we mean ‘assertion’ can justify Wiggins’s way of stating the constraint. Certainly, we need to assume that by and large our subjects

are speaking seriously – that they mean what they say – but only in the case of assertions does this amount to the assumption that they are aiming to speak the truth.

As for Wiggins's remaining marks of truth, they make essential use of the claim that 'language ... *par excellence* ... is the vehicle for the communication and expression of belief (1980, p. 205). Thus: 'The argument for [the third mark] ... is similar to the argument for the second mark in turning upon the notion of belief (pp. 208–9); '... the fourth mark [is] simply... a condensation of the first, second and third ...' (p. 211); and as for the fifth mark, 'it is a norm of rational belief that one is *committed* to conform one's beliefs to this requirement; and the interpretation of beliefs as beliefs has to see belief as *answerable* to it' (p. 212).

As a result, Wiggins's theory of truth is insubstantial in the second of the two senses of section 2.2. Given the notions of belief and expression of belief on which Wiggins here relies, we could quite well characterize the fact-stating use of language in terms of expression of belief. The basic distinction would be that between those utterances whose standard function is to express a belief and those that do something else. Far from looking for a substantial theory of truth, we ought then to expect our account of truth to flow from the same source as our theory of assertion. The obvious approach would be to regard truth as primarily a property of belief, and thence a derivative property of expressions of belief. The correctness of an assertion would thus be a matter of the correctness of the belief it would normally express.

The argument of this chapter should alert us to a potential hazard for an account of this kind. We have seen that assertion cannot be characterized in terms of truth. On the contrary, an account of truth and falsity as linguistic properties seems bound to depend on an elucidation of the fact-stating type of discourse. In offering such an elucidation we must be careful not to depend on a linguistic notion of truth. In particular, therefore, we cannot rely on a characterization of belief that itself depends on a notion of truth as a property of linguistic entities. This notion should arise from, not underpin, our psychological theory. If we want to analyse the fact-stating/non-fact-stating distinction in terms of a distinction between beliefs and other sorts of psychological state, we must be prepared temporarily to forego the linguistic notion of truth. This point will be very much in evidence in chapter 5, where we shall be looking in more detail at attempts to find a psychological basis for the notion of statement of fact.

3

The Insignificance of Mood

Marker theorists and non-factualists agree that there is a non-trivial distinction between fact-stating and non-fact-stating uses of natural language. They differ, however, as to where this distinction should be drawn. We have seen that if the marker theorist tries to settle the issue by appealing to truth, the attempt backfires. A notion of truth that secures the indicatives tends also to encompass the non-indicatives. We shall see now that this is the inevitable predicament of the marker view. No analysis of fact-stating discourse that includes the indicatives can exclude the non-indicatives. This will enable us to put the marker theory and the non-indicatives to one side. We shall be able to concentrate, in other words, on whether there is a distinction to be drawn within the indicatives.

Our goal is thus a generalized sceptical challenge to the marker view – an argument that non-indicatives cannot be shown not to be assertions. Since this would be entailed by a demonstration that non-indicatives are assertions, our course runs parallel to that of the ‘universal factualist’, who holds that *all* discourse is fact-stating. However, we’ll see that the path leads not to universal factualism but to our own less ambitious destination. It does not establish that non-indicatives are assertions; but it does provide a plausible *reduction* of any purported demonstration that non-indicatives are not assertions.

It will be helpful to have in mind some of the characteristics that have been thought to distinguish the non-indicatives, other than their standing to the properties of truth and falsity. So let us again refer to the views of Dummett, whom we have seen to be a critic of the attempt to analyse assertion in terms of truth. In his discussion of assertion in *Frege: Philosophy of Language*, Dummett mentions four main alternative characteristics. He says that assertions are the expressions of mental states with inner consequences (so that we can make sense of an internal assertion – a judgement – but not for example an internal command); are not typically associated with definite consequences (as are questions and commands); can be used to deceive; and can be used (perhaps stripped of assertoric force) as components of complex sentences. (Dummett 1973, chapter 10).

Not all of these features could plausibly be offered by way of analysis of assertoric discourse (or are, I think, so intended by Dummett). The last property is clearly grammar-dependent; and hence, as we saw in chapter 2, no use to the marker theorist, whose aim is to explain the grammar. At best this criterion amounts to a further characterization – a rather obvious one – of the distinction in usage that calls for explanation.

The suggestion that only assertions can be used to deceive also seems an unlikely basis for an analysis of assertion. If true, it surely rests on some more fundamental property of assertion. However, it proves a useful focus for our present concerns. In asking how non-indicatives can be used to deceive, we shall be led to a notion of the informational content that such an utterance could be considered to convey. From there it will be a short step to the claim that a non-indicative can be construed as an assertion with such a content – or rather, cannot be shown not to be such an assertion.

The attempt to make sense of non-indicative deception will lead us in particular to the proposal that non-indicatives are equivalent to explicit performatives (these being themselves assertoric). This proposal – advocated in the interests of general semantics by David Lewis (1972) – has been criticized on a number of grounds. I shall defend it against the better known objections, but argue that it is nevertheless untenable. Once recognized, however, the main problem is easily solved. I shall argue that the improved account meets our purposes (if not the universal factualist’s); and confirm that in so far as they need to do so, the assertoric paraphrases suggested by the account do indeed have the features Dummett takes to distinguish the non-indicatives.

3.1 NON-INDICATIVE DECEPTION

The universal factualist has two possible responses to the claim that assertions are distinctive in allowing deception: to argue that some assertions (and in particular non-indicative assertions) cannot be used to deceive; or to try to show that non-indicatives can be used to deceive. Let us begin with the latter possibility.

In support of the claim that assertions can be characterized in this way, Dummett argues that non-indicatives show no parallel to Moore's paradox. He concedes that 'the asking of a question is in a sense the expression of a desire to know the answer: so one may ask a question to deceive someone into thinking that one wants to know the answer.' But he counters that this

is not ... part of the convention which is learned when one learns to use and answer questions: it merely results from the fact that a desire to know the answer provides the usual motive for asking a question. This is shown by the fact ... that there is nothing self-defeating in asking a question and at the same time letting it be seen that one has no interest in the answer (if, e.g., one is an official whose duty involves the posing of certain questions); whereas the making of an assertion while letting it be seen that one does not believe it to be true is intrinsically self-defeating. (1973, p. 356)

At best, however, this shows us that a question does not express an interest in knowing its answer. It does not show that there is not some other attitude whose disavowal will frustrate an interrogative speech act. And on the face of it, any linguistic act can be defeated by an appropriately chosen attitude report. One simply needs to add that the utterance was not intended seriously, or sincerely. This takes advantage of the most general linguistic convention of all, namely that one means what one says. There may be various ways of signifying that a previous utterance should not be taken as governed by this convention, but there is always at least one way: thanks to Quine, one need only add that the utterance should be counted a mention rather than a use. True, not all effects of an utterance can be cancelled in this way. Some effects are insensitive to a speaker's intentions: for example, effects on voice-operated machinery. Others are immediate, so that the disavowal after the fact is bound to come too late. (Think of crying 'Wolf!' – or for that matter, of the effects of 'Think of a wolf!') Clearly, however, such cases do not distinguish the non-indicatives.

Does the possibility of insincerity allow non-indicatives to be used to deceive? It might be objected that pretending to sincerity is not the same as lying. The liar pretends to sincerity, but not all pretenders to sincerity are liars. A person who utters an insincere non-indicative misleads his audience about his own sincerity, but does not *lie*. Whereas a person who lies misleads her audience in two ways: about her own sincerity, and (if she is believed, at least) about the facts. To lie is to try to mislead in the second sense. In the case of commands, for example, the claimed contrast would thus be that a person who obeys a command is in no way at fault for doing so, even if it turns out that the command was insincere. Whereas a person who unknowingly accepts an insincere assertion may end up, not at fault for accepting it – that was the appropriate thing to do – but at fault in possessing a false belief. It is a virtue of this difference, it seems, that lying misleads about more than a speaker's sincerity.

However, it is easy to see how obedience to insincere commands might be inappropriate. In general obedience will involve some disadvantage, or behavioural cost; but this will be regarded as preferable to the disadvantages of disobedience – in other words, to being punished for failure to comply. However, the relevant social conventions might well recognize the possibility of insincere commands, and administer punishment accordingly. Disobedience to an insincere command would thus go unpunished. Hence we would have a worldly basis for the state of being misled by an insincere command, in the consequent performance of an unnecessary and otherwise undesirable action. Moreover, we could match such a state to a false belief. We could say that someone wrongly believed that he was *required* to perform some action. This suggests that we might interpret a command as an

assertion with content of this belief – for example, that we equate ‘Fill in this form’ with ‘You are required to fill in this form.’

To forestall an objection, let me clarify the role of the doubtful (and certainly contingent) proposition that insincere commands do not establish real obligations – that a person is not at fault for failing to comply with what turns out to be an insincere imperative. The suggestion does not require that actual commands conform to this convention. The point is rather that since the giving of commands clearly could be constrained in this way, we can make sense of a worldly basis for imperative deception. In terms of that basis we can then explain, if necessary, the fact that some or all commands do not permit this sort of deception. I shall come back to this point. But briefly: we can say that some or all actual commands are governed by the convention that they be self-justifying – the convention that their own utterance be sufficient to establish the requirement that makes their suggested paraphrase true.

There would be other objections to any serious proposal to paraphrase imperatives as reports of obligations. As yet, for example, such a report does not capture the first-person character of a command. Perhaps the paraphrase could be refined to cope with such objections, but for present purposes there is a more serious concern. Even if reports of requirements can be made to do duty for imperatives, we shall be left with the other non-indicatives. Insincere commands may mislead us about our obligations, but what about insincere questions, requests, wishes, and so on? A case by case argument is no use to us. We need a general guarantee that some such transformation will always be available. I want to show that the marker theory gives us such a guarantee. This will deliver the promised *reductio*. It will show that from the marker theorist’s point of view, universal factualism remains a viable alternative.

3.2 HOW TO CONSTRUE THINGS IN WORDS

The general strategy relies on some familiar points about linguistic communication. Utterances have various effects on the social environment of speakers and audiences. Among these effects are some that depend on linguistic understanding – on the fact that audiences are able to interpret what speakers say. Effects of this kind are convention governed. They rely on the meaning conventions in force in the community in question. Interpretation of an utterance can be thought of, in part, as a process of acquiring beliefs about these convention-governed effects. These beliefs are not often explicit; but it would seem that any of them could be elicited by an appropriate enquiry, at least from a subject with the necessary conceptual resources. Learning to understand a language seems in large part a matter of learning to form such beliefs in the right circumstances.

I take this much to be uncontroversial. The difficulties arise when we try to fill in the details. In particular, the interpretation of a single utterance will in general give rise to a string of beliefs. The proper characterization of some of these beliefs is far from obvious. Interpretation presumably involves the application of the interpreter’s knowledge of meaning, and hence the proper form of interpretative beliefs will depend on that of a theory of meaning for the language in question. A theory of meaning will bring with it an account of the corresponding interpretative beliefs.

The marker theory, for example, claims that mood distinctions mark significant meaning categories. If so, then the process of interpretation should recognize the factor of meaning marked by mood. Hence on this theory we might think of the main elements of the interpretative string as the following four beliefs: that someone has spoken (rather than simply made a noise); that a certain speech act (a command, an assertion, or whatever) has been performed; that the speech act concerned has a particular content; and that one is oneself thereby affected in some way (for example in being thereby obliged to perform some action).

The present significance of these points lies in the fact that beliefs can be acquired not only directly, in virtue of perceived features of one’s environment; but also indirectly, as a result of an assertion by somebody else. So if the interpretation and hence the effects of an utterance always depend on chains

of belief, we might achieve the same final effects in a different way: that is, by simply *telling* a person what they would otherwise come to believe at some point in the chain of interpretation. From the marker theorist's point of view this is what happens when we replace a command with an assertion of the existence of a corresponding requirement. In both cases we normally cause our audience to believe that there is such a requirement, but the initial pathways are different. In one case the belief concerned is inferred from the observation that we have issued a command; in the other it flows in the normal way from the observation that we have made an assertion – from the fact that we have *said* that there exists the requirement in question. But note that not only do the pathways converge, in giving rise to the same belief; but also the effects of initial insincerity may be the same, in making this belief false.

The marker theory suggests a simple and elegant way to adapt this construction to the general case. We want to break into a chain of interpretative beliefs, so as to assert what a hearer would otherwise infer. It doesn't seem to matter where we break in, as long as we can characterize the information the new assertion needs to convey. Given the marker theory, the obvious target is the point at which the mood itself is first deciphered. At this point the interpreter comes to believe that he is faced with a question, a request, or whatever. Why not replace the original utterance with an assertion to this effect?

The suggestion is thus that we exploit the marker theorist's account of the significance of mood. That account entails that the process of interpretation of an utterance requires, *inter alia*, a belief that a speech act of a certain kind has taken place – the kind being the meaning category that the marker theorist associates with the mood of the utterance in question. The marker theorist takes this belief to determine the further effects of the utterance concerned (or those, at any rate, that depend on its mood). The suggestion is that we imagine the same belief to be transmitted in another way: not inferred from mood, but reported in content. Non-indicatives are thus to be paraphrased in terms of explicit performatives (these being themselves regarded as assertoric). 'Fill in the form' will now be read not as 'You are required to fill in the form', but simply as 'You are commanded to fill in the form'; or more explicitly, as '*I* command you to fill in the form.'

In this way, the marker theory seems to provide the universal factualist with the material for an assertoric interpretation of the non-indicatives. It allows the factualist to propose that non-indicatives are merely 'disguised' reports of the occurrence of the very speech acts the marker theory claims to identify. To admit such readings would be to abandon the claim that the indicative mood is the mark of assertoric force. But how is the marker theorist to exclude them, given the intuitive appeal of the interpretative model from which they stem?

3.3 PERFORMATIVES: TOO EXPLICIT FOR THE FACTUALIST

The marker theorist might well follow J. L. Austin (1962, pp. 5–6, for example) in denying that explicit performatives are fact-stating – as of course the above suggestion requires, if it is to give us *assertoric* as opposed to merely *indicative*, substitutes for non-indicative utterances. In our terms, Austin was a non-factualist: he took it as a discovery about language that it contains this large class of non-assertoric indicative sentences. On Austin's view, the expression 'I command ...' is not used to describe one's own behaviour – to *assert* that one commands – but simply to command; similarly with 'I request ...', 'I promise ...', and many others.

Some of Austin's critics, however, though agreeing that explicit performatives are used for these purposes, have argued that such utterances are also assertions. One such critic is David Lewis, whose interest in explicit performatives is much the same as ours: faced with the task of incorporating non-indicatives into a general theory of meaning, Lewis recommends that they be treated as disguised performatives, and that these be regarded as assertoric, or declarative. In Lewis's words, the result is that

the distinction between declarative and non-declarative sentences becomes a purely syntactic, surface distinction. The only distinction among meanings is the distinction between those sentential meanings that can only be represented by declarative sentences and those that can be represented either by suitable declarative sentences (performatives) or by non-declarative paraphrases thereof. (1972, p. 208)

Lewis's proposal has been criticized on several grounds. We shall see that although the critics have mostly missed the mark, there is a valid objection. However, the problem stems from the unnecessarily Austinian character of the performative approach. Cured of its fixation with speech acts, the proposal turns out to be just what we want.

Lewis's critics have concentrated on the claim that explicit performatives are declarative; and in particular on the consequence that an utterance of an explicit performative must, as McGinn puts it,

be tantamount to the performance of *two* speech acts, where the (felicitous) performance of one (saying or asserting) is sufficient for the performance of the other (commanding, etc.) So in uttering a performative one both performs a certain speech act and declares that one does. (1977, p. 305)

McGinn's own objection to this consequence of Lewis's account is that

no such duality is apparent in respect of straight nonindicatives 'I command you to shut the door' mentions me and commanding, but 'Shut the door' manifestly does not. So they are simply not equivalent in meaning. (1977, p. 305)

In advance of a semantic analysis of non-indicatives, however, I think this is premature. The significance of my utterance of 'Shut the door' clearly depends in some way on the fact that it was I who said it (you may decline to obey, for example, for precisely that reason). Is it obvious that this significance is not the result of an implicit reference to myself? Similarly, one of the effects of this utterance is surely that you come to believe that I have ordered you to shut the door. Is it clear that this should not be explained by saying that I have implicitly stated that I have so commanded you? (Compare: 'Bring me pig meat' mentions pigs and meat. 'Bring me pork' manifestly does not. So they are simply not equivalent in meaning.)

The best attempt to make an objection out of the duality of the explicit performative seems to be that of Stephen Schiffer. Schiffer takes over from Austin the principle that 'ordinarily one uses the explicit performative formula to make explicit the precise act one is performing in the issuing of one's utterance' (1972, p. 107). He then points out that if we assume that performatives have assertoric force, this principle leads to a vicious regress. To make explicit the fact that 'I order you to shut the door' performs an assertoric speech act, we would have to say 'I assert that I order you to shut the door.' And here the same problem occurs again: to make explicit the assertoric force of this new assertion, we would need a yet longer explicit performative.

Schiffer's solution is to say that although performatives are constative, they are ordinarily used without assertoric force. This will not do for Lewis, who wants to claim, in effect, that we give commands by asserting that we are doing so. If 'I order you to shut the door' did not have assertoric force, you could justify your inactivity with the remark that I had not told you that there was a command to obey – that you had understood that I was merely entertaining the thought that I was thereby ordering you to shut the door. But as Kent Bach (1975) points out, a more plausible solution to Schiffer's regress is simply to reject Austin's view that performatives make explicit the precise nature of the speech act they perform. Instead, we can say that they make explicit one speech act in performing another.

The real difficulty for Lewis's account seems to me to stem not from the dual role it accords to the explicit performative, but from the fact that the roles are not independent. Lewis wants a semantic analysis of non-indicatives. In effect, his account is meant to be a guide to the interpretation of non-indicative utterances. It tells us, first, that they have the internal structure of a particular kind of assertion: namely an explicit performative. The structure thus looks like this:

(3.1) I perform the speech act A.

So the account tells us that if we receive a non-indicative linguistic 'package', its outer layer is to be treated as an assertion that we are being offered another package, another speech act. To understand the original, we have to unpack the package it claims to contain (unless we do so, after all, there will be nothing to distinguish one non-indicative from another). But what is this new package? Why, the very one we started with. Offered a command, for example, we are told to treat it as the assertion that we are being offered a command. Obligingly we refer to our account of commands, to understand what it is that the assertion says we have been offered. But it is a command, and so the account tells us that it is equivalent to an explicit performative – an assertion that we have been offered a command. The chain of interpretation thus has a loop in it.

This is not Schiffer's objection. Schiffer tries to build a regress outwards, working on the assertoric part of an explicit performative. This objection goes inwards, working on the speech act that the assertoric part reports. The problem stems from the attempt to analyse the speech acts performed by the non-indicatives in terms that refer to these speech acts themselves – that is, as reports of the performance of these speech acts. As Charles Hamblin puts it,

the so-called reduction ... [is] a logician's blind alley; ... it is not really a reduction at all. If someone did not know what an order was, and needed to be told what meaning to attach to a given one *O*, it would be no help at all to tell him that it was a locution that, uttered by me to you, meant

I order you to carry out order *O*,

since he wouldn't understand this any better; *O* remains embedded in it, unanalysed and un glossed. (1987, p. 135)

The marker theorist can thus object that the universal factualist has not offered a coherent assertoric analysis of the non-indicatives. An analysis of meaning is required to account for linguistic understanding. The explicit performative account cannot do this, because the performative refers to the very linguistic act whose analysis is in question. The factualist must do more than show that for any speech act conventionally performed by non-indicative means, we can find an indicative utterance that could be taken to perform the given speech act. Speech acts being conventional, any indicative utterance (indeed, any public action) can be made to do such additional duty. But the trick requires a prior conception of the speech act we propose to add to the conventional significance of the chosen indicative. So it cannot provide an alternative analysis of the speech act concerned. It is as if we had tried to show that pocket knives can be thought of as potato peelers, by producing an object that physically combines the two sorts of implement. What we need to show is rather that an arbitrary pocket knife can already be thought of as a potato peeler (by showing, of course, that it can be grasped as such).

3.4 SPEECH ACTS AND SPEECH EFFECTS

The problem for the performative approach to an assertoric treatment of non-indicatives is thus that if performatives are thought of as reporting speech acts, the resulting analysis is viciously self-referential. The key to a solution, I think, is to appreciate that what such an assertion reports does not have to be a speech act at all. It is again crucial that we are dealing with conventional acts. We need

the fact that if a social state of affairs is actually the conventional effect of an utterance, then it could instead be taken to be the conventional effect of almost any action (or indeed of other sorts of event: there are societies in which conventional obligations are taken to be dependent on random physical processes, for example).

Thus if convention enables a speaker to impose a requirement by saying ‘Please leave the room’, it could quite well enable the same speaker to impose the same requirement by other means – including, I think, a simple mental decision that the person concerned shall be required to leave the room. It might be objected that imposing the requirement in this way would enable people to become subject to requirements about which they had no way of knowing. This is true, I think, but hardly problematic. There are other sources of obligations to which individuals may unknowingly be subject. (The possibility is reflected in the principle that ignorance of the law is no excuse.) In any case, it could easily be avoided by treating such requirements as conditional on publicity.

We are liable to be misled here by the very view we are attempting to undermine. If we begin with the idea that a command is a type of speech act – different from, but on a par with, the act of assertion – then of course we read ‘I command ...’, if as a report at all, as a report of the occurrence of such a speech act. But if we are not already committed to a non-assertoric view of imperatives, we don’t have the same reason to think of commands as a category of speech acts. And hence we are free to think of ‘I command ...’ as a report of a conventionally generated state of affairs: that very state of affairs that on the orthodox view results from the corresponding imperative speech act.

Looking at things from the orthodox point of view, we can thus separate the *effect* of a speech act – in the case of a command, for example, the requirement or obligation it establishes – from the speech act which normally produces that effect. Not only does the existence of the effect not depend on the occurrence of such a speech act; but we are free to characterize the effect in terms that do not refer to such a speech act. And hence we can *report* the effect, without explicitly (due to our description of the effect) or implicitly (due to its implied causal history) reporting another speech act. So in place of a speech act that produces a conventional effect – and in the process, so to speak, alerts an audience to this effect – we can always imagine an assertion that there has been such an effect, understood as produced by the speaker by some other conventional means. We thus replace a speech act with a report, not of the occurrence of the speech act itself, but of the social state that speech act would normally produce – roughly, by a report of the relevant change that would have been made to an audience’s social environment by a performance of the original speech act.

We thus aim to intersect the chain of interpretation where recognition of the nature of a speech act gives way to response. We want to report just those effects of an utterance that a speaker is normally able to produce at will, without depending on audience compliance (except with respect to the relevant meaning conventions). Even at this level, of course, there are cases in which a speaker will fail to produce an intended effect. Lack of authority will vitiate a command, for example. (Austin 1962 describes many ways in which an intended speech act may be ‘unhappy’, as he puts it.) But these will simply be the cases in which the relevant reports are false.

I think this proposal gives us the indicative paraphrases we have been looking for. We have shown, in effect, that for each category of non-assertoric speech acts the marker theorist claims to identify, we could construct a class of indicative utterances that would serve the same linguistic function. These would be indicatives in the ordinary sense. At worst we would have to coin some new descriptive terms to refer to the typical effects of the non-indicative speech acts we were concerned to replace. The marker theorist cannot object to this procedure. Pressed to explain our terminology we can appeal to the marker theorist’s own account of the social significance of the force distinctions supposedly marked by the grammatical moods. In distinguishing questions, commands and other non-assertoric speech acts, the marker theorist incurs an obligation to elucidate these categories. Such an elucidation is bound to appeal to supposedly distinctive features of the social role of each of these forms of communication, and in particular to their distinctive effects on conventional relations between hearers

and speakers. These effects (or some of them) provide the subject matter for the factualist's assertoric paraphrases.

In summary then, the argument shows that if there were non-assertoric speech acts of the kind the marker view takes to be associated with the non-indicative moods, the function of these speech acts could quite well be performed by utterances that the marker view would clearly take to be assertoric. On what grounds then can the marker theorist maintain that non-indicatives are not simply 'disguised' assertions?

I emphasize again that the argument does not show that non-indicatives are assertions. It merely undermines the marker theorist's grounds for claiming that non-indicatives are not assertions. So it supports the universal factualist only to extent that it counts against a contrary view. Indeed, I think its effect is to call into question the enterprise that universal factualism and the marker view have in common, of seeking to discern the 'true' logical form of natural utterances. For it suggests that the categories both views employ are sufficiently flexible to prevent any authoritative judgement on how they should be applied to natural language.

3.5 EXPLAINING THE APPEARANCES

The marker theorist's best hope of a defence would be to find some significant characteristic of non-indicative speech acts not shared by the claimed indicative paraphrases. It is a slim hope, however. For one thing, the argument exploits the marker theory itself, in conjunction with a simple and plausible model of linguistic interpretation. For another, the factualist's indicative paraphrases seem to account in one of three ways for the evidence that appears initially to favour the marker view. The supposed characteristics of the non-indicatives divide into three classes. Some supposedly distinctive properties of non-indicatives turn out to be shared by their indicative transforms. Some properties supposedly lacked by the non-indicatives turn out to be applicable after all, in ways determined by the indicative transforms. And some properties, as we have seen, turn out to be distinctive but merely grammatical – the kind of thing that the marker theory claims to explain.

All three possibilities are exhibited by the range of characteristics that Dummett takes to distinguish the non-indicatives. As we saw, Dummett says that assertions are not typically associated with definite consequences; are the expressions of mental states with inner consequences (so that we can make sense of an internal assertion – a judgement – but not for example an internal command); can be used to deceive; and can be used as components of complex sentences. Taking these in order, the impression that non-indicatives have definite consequences is confirmed and explained by the subject matter of their claimed indicative para- phrases. These are assertions about the social environment of a speaker and a recipient – in general, about features of this environment of immediate relevance to the recipient's actions. The indicative substitute for a command, for example, informs its recipient of a social obligation (or of a state of affairs of which such an obligation is normally an immediate consequence). So it is hardly surprising that such utterances have consequences which are unusually definite, by the standards of assertions in general.

The apparent lack of 'internal' correlates of non-indicative speech acts falls into the second of the above categories: the indicative transforms provide such internal correlates. Thus if we think of commands as assertions of the form 'You are required to do A', then their internalizations will take the form 'X is required to do A', where X will normally be a term that refers to someone other than the speaker. It is true that once this reference is made explicit, we can consider the case in which it is a self-reference. Externally or internally, judgements of the form 'I am required to do A' are perfectly respectable. Indeed, they have just the form in which the recipient of the indicative substitute for a command will report his or her newly acquired belief. Whether they have another use, as the transforms of self-commands, depends on extraneous factors: on whether such requirements can be self-imposed; and on whether, if so, there is ever a need to report them.

Similarly, the indicative transforms provide a sense in which commands can be said to express mental states with separate internal consequences. If commands express beliefs about requirements, for example, then such beliefs will have other effects: believing that you are required to complete the form, I take steps to ensure that if you do not do so (having been properly informed of the requirement), appropriate measures will be taken. It does not count against this view that the belief may be the result of my own decision to impose such a requirement, or that the requirement may be conditional on your actually receiving news of it.

The supposed impossibility of non-indicative deception also falls into the second category. We have already seen that one can be misled by an insincere command just as by an insincere report of a corresponding requirement. We noted that the assertoric model can cope with the possibility that disobedience might be regarded as punishable, irrespective of the commander's intentions. The basis of such requirements is bound to be a matter of social convention. One relevant convention might be that in the absence of explicit indications to the contrary, any utterance of a command or its indicative substitute shall count as establishing the requirement that makes that substitute true. In limited ways this sort of convention governs many assertions. Think, for example, of 'This is a stick-up', uttered in a bank. Deceit is not always easy.

This leaves the suggestion that non-indicatives do not occur as sub-sentential components. As noted earlier, this is a grammatical restriction, and not therefore a feature of non-indicatives to which the marker view is entitled to appeal. As such, moreover, it collapses in the face of indicative substitutes for non-indicative speech acts. For example: 'If you are required to complete the form, you are required to do so in black ink.' The fact that we do not say 'If complete the form then complete it in black ink' is on this view no more than a simple consequence of the rules of non-indicative grammar. This does not mean that it is insignificant, but that its significance is just that of broader question as to the reason for the existence of the non-indicative moods.

We shall return briefly to this question in chapter 8. By then we shall have an alternative to the marker theorist's claim that the moods are the conventional marks of underlying semantic categories. We have argued here that such a claim cannot be justified. Indeed, it follows from the argument that the convention the marker theory calls for would be unworkable. *Qua* speakers, just as *qua* theorists, we lack the mood-independent access to the claimed underlying categories that the convention would require.

In the main, however, the rest of the book will be concerned with the indicative mood and its internal structure. The aim of this and the previous chapter has been partly to sanction this concern, by showing that the issues the non-factualist raises within the indicative mood are not to be swept aside in the marker theory's grand account of the moods as a whole. More importantly, however, it has been to show that there is no refuge in the marker view, in the face of a sceptical challenge to non-factualism. Such a challenge will be the major task of the next two chapters. We may now appreciate that what the sceptic challenges the non-factualist to exclude is not merely the marker view. That outcome, if not as philosophically illuminating as non-factualism is thought to be, would at least be familiar. Given that the marker theory is itself vulnerable to sceptical challenge, however, the sceptic's alternative is not so tame. The threat to non-factualism is rather the bleak possibility that there is no way to exclude universal factualism – no analysis of statement of fact that entitles us to draw non-trivial limits to fact-stating discourse. We have seen that for non-factualists, as for marker theorists, to admit this possibility is to admit defeat. Appealing as it may seem to general semantics, universal factualism cannot support the distinctive claims of either doctrine. It cannot explain mood in terms of the limits of statementhood, and it cannot attribute philosophical errors to the misattribution of statementhood.

4

A Retreat to the Head

In this chapter and the next our goal is to embarrass the non-factualist with an abundance of facts. More precisely, we want to show that the non-factualist's best arguments are inconclusive, not excluding factualist interpretations of the topics in question. So again, our intentions are mainly sceptical: we want to show not that there are more facts than the non-factualist supposes, but that the non-factualist cannot establish that there are not such facts.

Non-factualism is more a species of philosophical theory than a theory in its own right. Since we cannot examine every member of the species, I propose to concentrate on three of its major manifestations: the non-factualist treatments of ethical judgements, probability, and (in chapter 5) indicative conditionals. Among the advantages of this selection is the fact that on the face of it, they have little to do with each other. This will add weight to the claim that the striking parallels which emerge, between probability and ethics in particular, reflect a characteristic pattern of non-factualist argument.

Broadly, the plan of these two chapters is as follows. In this chapter we want to show that non-factualists can be forced to depend on a psychological distinction – in effect, on the claim that genuine statements of fact are the expressions of genuine beliefs. In chapter 5 we shall then challenge their entitlement to this distinction. We shall argue that beliefs cannot be distinguished from the relevant non-beliefs without appealing – illicitly, so far as the non-factualist is concerned – either to truth or to existence of appropriate facts in the world.

First, however, a point of clarification. Such a case against non-factualism may seem superfluous, given the argument of chapter 3. We criticized the marker view on the grounds that by the standards of that view, any convention-governed speech act can be construed as an assertion. The marker view requires that a successful speech act convey, in some sense, the information marked by mood. Consequently, the marker theorist cannot exclude the possibility that the relevant feature is *reported*, rather than merely *marked*. This argument may seem to count as much against non-factualism as against the marker theory. If no utterance can be shown not to be an assertion, then in particular ethical judgements, conditionals, and the like cannot be shown to be non-fact-stating.

However, I think non-factualism fares better than the marker theory. For one thing, it is not clear that the non-factualist need require that the fact-stating/non-fact-stating distinction be marked in speech, or accessible in principle to ordinary speakers. If not, then the distinction need not be manifest in the chain of interpretation, in the way that the marker theory evidently requires. Since the *reductio* argument of chapter 3 depends on this feature of the marker view – the sceptic proposes an alternative chain of interpretation with the same final effect – non-factualism may well be immune.

For another thing, the contents of the assertoric paraphrases on which the *reductio* relies are themselves closely based on speech act categories. The idea was to map force distinctions onto a class of content distinctions (roughly, distinctions within the set of assertions about certain conventionally generated states of affairs). This means that the taxonomy of non-assertoric speech acts that the non-factualist claims to find within the indicative mood would survive the adoption of this sort of universal factualism – albeit on the basis of content distinctions rather than force distinctions. So the non-factualist's case could simply be rephrased. The question would be, not whether a moral judgement should be regarded as an assertion at all, but whether it should be regarded, at face value, as an assertion about a realm of moral facts. The alternative would depend on the version of moral non-factualism on offer. It might for example be an assertion to the effect, roughly, that the speaker has expressed an evaluation of a certain kind.

The non-factualist could thus adapt to the possibility of universal factualism, by presenting the issue in terms of the *contents* of disputed utterances: granted that moral judgements, like everything else, are statements of fact, are they really about *moral* facts, or rather about some utterance-related conventional state? The non-factualist would still claim to show that certain classes of utterances are not assertions with the contents they seem at first sight to have. The negative claim will remain unchanged, though the positive claim will now be that the utterances concerned are assertions with some other content, rather than that they are not assertions at all. However, the negative thesis is the basis of the non-factualist's 'resolutions' of philosophical problems in various areas. So the approach seems essentially intact.

Moreover, our present interest, here and in chapter 5, is in criticizing the non-factualist's own case for the negative claim. The issue of what would follow from accepting this thesis will thus be largely irrelevant. It is true that in principle we might want to argue against the thesis by *reductio*, and hence be interested in its consequences. But in fact we shall concentrate our attack on the presuppositions of the non-factualist's typical grounds for accepting the thesis. The marker view's problems are thus of no immediate relevance. They stand in the background – a warning that we would be unwise to take the downfall of the non-factualists as a vindication of the marker theory.

4.1 PROBABILITIES AND VALUES: THE NON-FACTUALIST INTERPRETATION

I want to exhibit some similarities between the usual grounds for non-factualism about ethical judgements and about probability. In particular, I want to show to show that for essentially parallel reasons, non-factualists in both areas can be forced to appeal to psychology. It will be helpful to begin with a characterization of what is at issue in each case. These characterizations are far from comprehensive, and perhaps exaggerate the initial formal similarity between probabilistic and ethical non-factualism. They will do for our purposes, however; and are not, I think, an artificial source of the striking parallels we find later on.

Let us begin with the probabilistic case. In English a probabilistic expression often has the syntactical form of a sentential operator: operators such as 'It is probable that ...', 'There is a 60 per cent chance that ...', and 'Probably, ...' itself, for example. Let us call a sentence formed by attaching such an operator to a suitable sub-sentence a *simple probability ascription*, or *SPA*. I have argued elsewhere (Price1983b) that SPAs correspond to the so-called *single case* applications of probability, much discussed in the philosophical literature. For the present, however, the relevant question is whether a SPA should be construed as a statement of fact. Does a probabilistic operator modify the *content* or the *force* of a sentence to which it is attached?

Most philosophical accounts of probability take for granted the factualist answer. In effect, they are attempts to characterize the distinctive content of a SPA – to tell us what SPAs are characteristically *about*. There are several candidates: class ratios (finite or infinite, actual or hypothetical), logical relations, propensities, objective chances or probabilities, degrees of rational belief, and even a speaker's own degrees of confidence. As the list suggests, there is a wide range of opinion on what should be expected of an account of probability (and especially on the extent to which one should aim to reduce that notion to others).

The non-factualist, however, says that all these views are misconceived. SPAs have no common content. Literally, there is no more something that all sentences of the form 'Probably P' are about, than there is something that all sentences of the form 'Is it the case that P?' are about. 'Probably P' is about what P is about. It differs in that whereas an utterance of P is normally an assertion, 'Probably P' is normally a different sort of speech act – what may be labelled a *qualified* or *partial* assertion.

As in many cases, the linguistic distinction on which the non-factualist here relies is conveniently elucidated in psychological terms. It assumes a familiar model of assertion, belief and relevant action: an assertion that Q is typically an expression of a (full) belief that Q; and such a belief is typically

displayed in a person's acting as if Q whenever, in the light of her desires, she thinks it makes a difference to the outcome of her actions whether it is the case that Q. The factualist subsumes SPAs under such a model, regarding them as assertions that express (full) beliefs with some characteristic content. The non-factualist, on the other hand, takes SPAs to express 'partial' beliefs, or degrees of confidence. 'Probably P' is thus the characteristic expression not of the full belief that it is probable that P, but of a strong partial belief that P. A partial assertion (hence the term) is thus the expression of a partial belief. Moreover, because full beliefs are the limiting cases of partial beliefs, an ordinary or 'full' assertion is a special sort of partial assertion.

For the moment, not wishing to prejudge the question of the priority of psychological over metaphysical presentations of non-factualism, we should treat this version as merely expository. In particular, we want to hold open the possibility that the belief/non-belief distinction might need to be founded on the availability of relevant domains of facts in the world.

Turning now to the ethical case, we can characterize the issue in similar terms. For present purposes we can confine our attention to a particularly simple sort of ethical judgement: those of such forms as 'It is a good thing that ...', 'It is disgraceful that ...', and so on, in which the evaluative expression can be regarded as a sentential operator. Following the pattern of the probabilistic case, I shall call this a *simple ethical ascription*, or *SEA*.

The competing views are then as follows: the factualist holds that SEAs are assertions, and hence that ethical sentential operators are content modifiers; the non-factualist that the role of the sentential operator is rather to signal a non-assertoric force. A psychological elucidation again suggests itself. The factualist says that SEAs express ethical beliefs – beliefs about moral facts or states of affairs. Whereas the non-factualist takes SEAs to express some sort of distinctive moral attitude – approval or disapproval, say – towards the non-moral state of affairs referred to within the scope of the operator concerned.

Not all the arguments for non-factualism about SPAs run parallel to corresponding arguments about SEAs. For example, a probabilistic non-factualist would do well to help himself to some of the objections his factualist opponents raise against one another. Thus, he might well endorse the standard criticism of (most) frequency accounts of probability, concerning single case probabilities (the so called 'problem of the reference class' – see Mellor 1971, chapter 3, for example).

Similarly, the non-factualist will do well to concur with several serious objections to naïve subjectivism – the view that SPAs report a speaker's own degrees of confidence. One such objection is the phenomenological one: in considering the probability of snow this afternoon, I don't consider my own degree of belief in snow this afternoon. How could I, since the purpose of deliberation is to lead me to some such degree of belief? Another objection points to ordinary disagreements about SPAs. Clearly these are not disagreements about what each of the participants currently believes. The argument gets its point from the fact that the participants both know what they each believe, and therefore know that they disagree. (In this case, clearly, there is a parallel argument against ethical subjectivism.)

However, arguments of this sort merely encourage the factualist to retreat to a stronger position. The strongest is the simplest: renouncing the attempt to provide a *reductive* interpretation of SPAs, it settles on the claim that SPAs are about, simply, probabilities. Probabilities are objective features of world, and hence a respectable subject matter for statements of fact.

It is easy to imagine a corresponding factualist position in the ethical case. And in the face of doctrines as bare as these, the non-factualist's options are rather limited. There is little for an objection to latch on to – and little, similarly, to distinguish the probabilistic and ethical cases. I think that in each case the non-factualist has two main avenues to explore. The first develops a metaphysical argument, while the second appeals to the functional role of ethical or probabilistic

judgement in our lives. I want to show that the probabilistic and ethical cases do indeed run parallel here, and that in both cases both avenues lead to the head.

4.2 ODD FACTS: THE ARGUMENT FROM QUEERNESS

In the ethical case, more perhaps than in any other, non-factualists have been prepared to appeal to the apparent peculiarity of facts of the disputed kind: 'If there were objective values, then they would be entities or qualities of a very strange sort, utterly different from anything else in the universe' (Mackie 1977, p. 38). Mackie calls this 'the argument from queerness'. It is often linked to an epistemological point. Non-factualists argue that if there were such facts, then 'if we were aware of them, it would have to be by some special faculty of moral perception or intuition, utterly different from our ordinary ways of knowing everything else' (Mackie 1977, p. 38).

As we shall see, the non-factualist's metaphysical case does not rest here. And just as well, for there is nothing here to trouble a committed factualist. With a touch of irony, the factualist can object that queerness is in the eye of the beholder. Moral facts seem queer to a non-factualist because he already has a view of the world that excludes them. It is true that they differ from say physical facts; but this is no more than a truism. Facts about different sorts of things are bound to differ – for how else could they be about different sorts of things?

This leaves the epistemological point. But here the factualist will point out that her opponent agrees that we do not make ethical judgements (or whatever sort of judgement is at issue) at random. On any account, there is some process that leads us to make some such judgements and not others. Why not regard this process, whatever it is, as the so-called moral intuition? Once again (the factualist will say) the non-factualist's mistake is to start with some preconceived notion – this time of what a belief-forming process should be like – and then to dismiss the moral case on the grounds that it does not conform to this model.

The non-factualist may feel that these replies do not do justice to the argument from queerness and its epistemological shadow. But I think the feeling stems from the fact that the non-factualist has a stronger metaphysical argument in mind. The trouble with moral facts, the non-factualist suspects, is not that they are odd but that they are unnecessary. Indeed, we are entitled to regard them as odd because they are unnecessary. The non-factualist thinks that we can explain our *talk* of values on a much sparser metaphysical basis.

Similarly in the probabilistic case. Here the argument from queerness consists in an appeal to a conception of the world in which there seems no place for objective probabilities. The non-factualist tells us that the world is the realm of the actual, the domain of physical things. Everything is determinate, every point either occupied or unoccupied, every particle in its place, and every proposition either true or false. Where, then, is the room for probability?

As it stands, this is no more convincing than its ethical counterpart. True, it presents non-factualism as a corollary of an attractive Humean metaphysical minimalism. But the blessed poor notwithstanding, economy is not a virtue in itself; it needs to keep company with sufficiency. So a proper appeal to metaphysical minimalism is at best one side of a more complex argument. The non-factualist has to show that probabilistic facts are unnecessary – that we can account for our talk of probability without construing it as talk about a distinctive feature of the world. Only thus will non-factualism be shown to have the conjunctive virtue of economical sufficiency.

4.3 PROJECTIVISM

Like the metaphysical position it serves, the argument for sufficiency is attributable to Hume. It turns on the claim that in addition to belief itself, there are other kinds of attitude one may hold to things in the world. There may be partial beliefs, and perhaps valuations and expectations of various kinds.

However, we are inclined to treat these attitudes as if they were genuine beliefs. We thus mistake for a difference in the world what is really a difference in the head. Such a mistake may have pragmatic advantages – for example, it may enable us to improve our attitudes by reasoning and argument ‘about the facts’. But metaphysically speaking, it is a mistake all the same: the ‘facts’ we seem to find in the world about probabilities (or whatever) are actually the images or ‘projections’ of our own attitudes. We project these attitudes onto the world, and hence come to see it as containing corresponding realms of facts.

Perhaps the leading contemporary proponent of this Humean approach is Simon Blackburn, who calls the view ‘projectivism’. He emphasizes its economical advantages:

The projective theory intends to ask no more from the world than what we know is there – the ordinary features of things on the basis of which we make decisions about them, like or dislike them, fear them and avoid them, desire them and seek them out. It asks no more than this: a natural world, and patterns of reaction to it. By contrast a theory assimilating moral understanding to perception demands more of the world. Perception is a causal process: we perceive those features of things which are responsible for our experiences. It is uneconomical to postulate both a feature of things (the values they have) *and* a mechanism (intuition) by which we are happily aware of it. (1984, p. 182)

Projectivism thus claims much the same advantage as buying furniture in kit form. By doing some of the work ourselves, we reduce our initial commitment – financial or ontological, as the case may be. But in both cases it is possible to overdo things. If a tree is a furniture kit then kit builders become cabinet-makers. And unless the projectivist can distinguish the real beliefs from the ‘patterns of reaction’, then projectivism becomes a species of idealism – the view that all ‘facts’ are projections of our reactions to the world. This outcome obliterates the distinction at the very core of the non-factualist’s thesis. In the process, moreover, it overturns the projectivist’s claim to metaphysical economy. The idealist’s world can be as cluttered as anybody else’s: idealists differ from realists not so much as to what exists, as to what its existence amounts to.

To avoid the slide to idealism, the projectivist thus needs a core of genuine facts: a secure foundation, on which the explanation of higher-level appearances can be constructed. Blackburn talks of asking ‘no more from the world than we know is there’. But how do we know what is really ‘there’, and what we put there? This issue has an important corollary: it means that projectivism itself is no guide to the facts. To arrest the slide to idealism, the projectivist needs to distinguish the genuine facts from the realm of ‘apparent’ facts for which the projective account is appropriate. So the possibility of a projective account of an area of discourse cannot in itself provide the motive for a non-factualist construal of that topic. To be sure, the projective account pays off a debt that the non-factualist incurs in urging such an account: it meets the obligation to explain the appearances. But the net effect is to balance the books. There is no profit in projective explanation.

The projectivist might object that this ignores the cost of the raw materials. Since Hume, projectivists have been motivated by metaphysical economy – by the suggestion that if we take ontological costs into account, then projectivism comes out ahead. But we can now see that the projectivist’s deal is too good to be true. If we can choose our raw materials, and less is better, then the best choice is none at all: that is, idealism. But idealism defeats non-factualism, so the projectivist cannot really accept that raw materials are a matter of choice. In other words, the economy projectivism offers us is ours already – or so the honest projectivist must believe. Like a good accountant, projectivism may show that we are better off than we thought, but it cannot take the credit for making us so.

What then would it take to arrest the slide to idealism? Have we perhaps underestimated the role of other metaphysical considerations, such as the argument from queerness? We shall see below that Blackburn himself offers another metaphysical argument for a non-factualist interpretation of moral

judgement. These arguments bear on the wrong side of the issue, however. They lever the projectivist cart into motion, by rejecting factualist accounts of moral usage. But the non-factualist's problem is to stop the cart, before it bears him to idealism. He needs an argument to the effect that certain topics *are* factual – and the better projectivism is made out to be at explaining factual appearances, the less likely it is that there can be any such argument.

The one live possibility seems to be to appeal to psychology. If the non-factualist could distinguish the genuine beliefs – in the head, as it were – from the various attitudes from which we project, then the barrier to idealism would be readily apparent. Projectivism would be safely contained, its task being to explain the fact that some non-beliefs are treated like beliefs. We shall return to the question of the availability of such a psychological distinction in chapter 5. But even if the non-factualist is entitled to such a distinction, the fact remains that projectivism provides no grounds for non-factualism. Projectivism is powerful medicine. Unless the need that it serves is grounded elsewhere, however, its effect is to make us idealists – which for the non-factualist counts as an overdose.

Let us turn then to the second major strand in the contemporary case for ethical and probabilistic non-factualism. This line of argument has a more overt link with psychology than the supposedly metaphysical arguments we have just discussed. Indeed, it might be thought to provide the psychological basis the projectivist seems to require. It focuses on the distinctive functional roles of probabilistic and moral judgements in our mental and behavioural lives – their significance in thought, speech and action. The notion of functional role might seem to provide the grounds for a distinction between beliefs and other propositional attitudes. I want to show that in appealing to aspects of the use of probabilistic and ethical judgements, non-factualists have taken for granted that the required psychological distinction can be characterized in these terms.

4.4 FUNCTIONAL ROLE (I): PROBABILITY

Probability provides an important mode of qualification of ordinary empirical judgements. Empirical evidence is rarely conclusive. Ordinary judgements are rarely felt to be certainly true, to be justified beyond a shadow of a doubt. An interpretation of probabilistic sentential operators needs to be sensitive to the fact that such operators are used to register this feature of judgement, to indicate the *degree* to which a judgement is felt to be justified. True, we often make judgements with no such qualification, even if we are less than certain that they are true. But we are quick to forgo this convenience if the circumstances call for precision – if there is a lot at stake, say. Often then we resort to SPAs.

The fact that probability plays this role raises two sorts of problem for the factualist. For one thing it threatens a vicious regress: if probabilistic qualification is a more or less universal feature of accurate factual judgements, and probabilistic judgements are themselves factual, then accuracy seems to demand that in place of 'It will rain' we say 'It will probably rain'; that in place of that we should say 'It is probable that it will probably rain'; and so on. The regress is vicious because an interest in accuracy – an interest in asserting only what we are *justified* in asserting – will prevent us from resting content with any particular stage.

The factualist has a couple of options here. For a start, she could say that probabilistic judgements are a special case, and can be justified with certainty. The main problem with this is that, notoriously, the statistical rules of inference that are held to justify probabilistic conclusions are themselves probabilistic. A given range of frequency data will make it *likely*, but certainly not certain, that a particular probability has a specified value. (This is the source of concern about the circularity of statistical inference; see for example Braithwaite 1953, chapters 5 and 6.)

A much more promising suggestion relies on what we may call the *limiting case principle*: the principle that as the probability concerned becomes sufficiently high, a probabilistic judgement becomes 'equivalent' in some useful sense to the corresponding unqualified judgement. Given this

principle, and the fact that in practice our first level probability judgements are usually themselves justified to a sufficiently high probability to allow us to apply it, it will follow that in practice we rarely introduce second (or higher) level probabilities.

The limiting case principle is so natural that it is easy to overlook its importance, and to fail to notice that for the factualist its justification is far from obvious – indeed, its justification is the second and more fundamental difficulty that arises, for the factualist, from the role of probability as a qualifier of ordinary empirical judgements. Clearly there is some important connection between the meanings of ‘It will rain’ and ‘It will probably rain’. Yet for the factualist these statements have different contents. The problem is to explain the connection between these contents, in such a way as to account for the limiting case principle, and hence make sense of the use of probabilistic qualifiers of ordinary judgements.

The non-factualist might explain these things as follows: probabilistic judgements express degrees of confidence, or degrees of partial belief. So ‘It will rain’ and ‘It will probably rain’ simply express different degrees of belief in the same proposition. And a series of judgements of increasingly strong probability that P, with limit 1, is thus associated with a series of increasingly strong partial beliefs that P, whose limit is the full belief that P. A judgement of high probability is practically ‘equivalent’ to the corresponding unqualified judgement in the sense that, and in so far as, a high degree of confidence in a given matter is practically equivalent to certainty. This explanation turns on the notion of the degree of a belief. Hence it is useless to the factualist, who is committed to explaining the use of probabilistic operators in terms not of differences of degree but of differences in content. The factualist seems to require an (increasing) similarity of content between the belief(s) that it is (increasingly) probable that P and the belief that P.

This problem stems from the role of probability in judgement. A related problem arises from its role in decision making. Factualists and non-factualists agree that probabilistic judgements have certain typical behavioural manifestations, at least in certain straightforward cases, such as the choice of betting odds. For present purposes we can assume that both sides accept that Bayesian decision theory gives an approximate even if very idealized description of this process. The non-factualist again has a ready explanation: this behaviour is simply the characteristic (non-linguistic) behavioural display of partial belief. What Bayesian decision theory describes, in effect, is the functional role of a partial belief, in terms of its content and degree. (The more familiar story about the functional role of a full belief falls out as a special case.)

For the factualist, however, there is a problem: Why should the full belief that it is probable that P have the functional role that the non-factualist accords to a partial belief? Why should beliefs about probability be reflected in betting behaviour, for example? After all, the factualist will expect another sort of manifestation of the belief that it is probable that P: the sort of behaviour standardly associated with a full belief, given in this case a content of the form ‘It is probable that P’.

On the face of it, the factualist’s problem is thus to explain why probabilistic judgements should be accompanied by ‘corresponding’ degrees of confidence: for example, to explain why the belief that it is probably going to rain should be associated with a high degree of confidence, or strong partial belief, that it is going to rain. I shall call this *the confidence problem*. Note that a solution to the confidence problem would not only allow the factualist to explain the practical manifestations of probabilistic beliefs; it would also justify the limiting case principle, and hence explain probability’s role as a qualifier of empirical judgements.

Why should someone who believes that it is probable that P be confident that P? Factualists who notice the difficulty are inclined to appeal to rationality at this point. (See, for example, Braithwaite 1966, a paper entitled ‘Why is it reasonable to base a betting rate upon an estimate of chance?’) There are two difficulties with this move. First, the notion of rationality it depends on is problematic, to say the least. Its natural interpretations either themselves invoke probability, or cognate notions –

‘rational’ means ‘most likely to be successful’, for example – or they invoke some notion such as hypothetical limiting frequency, already invoked by certain of the more reductive factualist accounts of probability. Either way, the confidence problem remains: why should a belief about probability, or about hypothetical long run frequency, have a bearing on degrees of confidence, and hence on action in the single case. It is doubtful whether an account of the relevant notion of rationality can escape circularity of one of these two kinds. Notoriously, deductive inferences from probabilistic premisses yield probabilistic conclusions; while the non-deductive inference to degrees of confidence is what needs justification.

Secondly, it is far from clear that the appeal to rationality makes the connection strong enough to do justice to our intuitions. Faced with subjects who regularly failed to match their degrees of confidence, and hence their actions, to their claimed estimates of ‘probability’, I think that we would be inclined to say not that they were irrational, but that they did not mean what we do by ‘probability’ and related terms.

These factors have led some writers (e.g., Mellor 1971, p. 67) to the view that the inference from a full belief about a probability to the corresponding partial belief – let us call it *the downward inference* – is in some sense analytic. The simplest such account, taking to heart the intuition on which the earlier suggestion was based, is what may be called the *rationalist* proposal: probabilities simply are degrees of rational belief, so that ‘It is probable that P’ simply means ‘It is reasonable to be confident that P.’ There is then no question of believing that it is probable that P, and yet doubting that it is reasonable to be confident that P.

However, it is important to realize that there is still a question as to why people who believe that it is reasonable to be confident that P are (on the whole) *actually* confident that P. What is it about the content of this belief (about what it is rational to believe) that leads people to adopt the corresponding degree of confidence? Of course, this is simply the original confidence problem, dressed up in new clothes – clothes it acquires in the rationalist’s paraphrase of a SPA.

At this point I think the rationalist has two possible moves. One is to attempt to explicate the notion of rationality, so as to display a decision procedure that justifies the (newly construed) downward inference, and in which ordinary self-interested speakers can plausibly be held to indulge. But this faces the very problems that, as we saw, plague attempts to show that the belief that it is probable that P makes it reasonable to be confident that P.

As in that case, the objection is perhaps not conclusive. But it does encourage another approach: to say that the use of the downward rule is constitutive of a grasp of the meaning of the relevant rationality ascriptions. A person will thus be said not to know what ‘It is reasonable to be confident that P’ means, unless in general when they profess to believe it, they are confident that P. Let us call this *contextual* rationalism: it takes the meaning of the term ‘rational’ to be bestowed by its inferential context.

The rationalist’s appeal to rationality in accounting for the meaning of SPAs may now seem redundant. Why not simply be a contextualist about probability itself – saying that the habit of making the downward inference is (partially, at any rate) constitutive of a grasp of the meaning of SPAs themselves – without making a detour via ascriptions of rationality? Either way, however, it can be seen that the confidence problem, concerning the typical *use* of the factualist’s full beliefs about probabilities, provides a strong motive for a particular account of their content: it favours the contextual theory. In effect, the factualist has been forced to retreat from ambitious reductive accounts of the meaning of SPAs, to the more modest and hence less exposed contextual position.

At this point, it may seem that the factualist has really conceded everything that was originally distinctive about her position. The factualist claims to take SPAs to express full beliefs about probabilities; and in response to the confidence problem, has taken pains to establish that in virtue of

the content of beliefs about probabilities, such beliefs are typically accompanied by the corresponding partial beliefs. In effect, then, the factualist has acknowledged that the use of SPAs – indeed, any manifestation of this sort of single case belief about probabilities – is also an indication of a speaker's possession of a corresponding partial belief. In consequence, the factualist seems vulnerable to a charge of theoretical redundancy. Beliefs about probability now seem idle cogs. Except in the factualist's account of probabilistic assertion, all the functional work is done by the corresponding partial beliefs. So why not abandon that account of assertion? Why not acknowledge that SPAs are products of partial beliefs?

The factualist thus appears to face a dilemma. If there are full beliefs about probabilities, then the question arises as to their connection with degrees of confidence. If the connection is too loose, then beliefs about probabilities will not have the behavioural significance that everyone is agreed that they should have, if there are such things. If it is too strong, then full beliefs about probabilities come to seem redundant – dispensable in favour of their constant companions, the corresponding degrees of confidence.

However, the factualist has one more card to play. The non-factualist's attack has rested on the distinction between the full belief that it is probable that P, on the one hand, and the attitude of strong partial belief that P, on the other. But why shouldn't the factualist simply reject this distinction, claiming that these two mental attitudes are one and the same – that a belief that it is probable that P just *is* a strong degree of confidence that P? Factualism would thereby take on the explanatory advantages of non-factualism, without an embarrassing duplication of psychological attitudes.

In other words, this 'radical' factualism challenges the psychological distinction on which the non-factualist has been relying. The non-factualist charges that factualist interpretations of probability are unable to make adequate sense of the functional role of probabilistic judgement – the role of what the non-factualist calls partial belief. The radical factualist responds by helping herself to the non-factualist's functional account; but insisting that it be regarded as a characterization of the distinctive content of probabilistic beliefs. The prospects for non-factualism again turn on its ability to distinguish genuine beliefs from other psychological attitudes.

4.5 FUNCTIONAL ROLE (II): VALUES

The confidence problem has an analogue in the ethical case. Indeed, it is perhaps the most venerable argument for a non-factualist treatment of moral judgements, tracing its origins at least to Hume. It relies on the principle that, as Hume puts it, 'reason alone can never be a motive to any action of the will; ... it can never oppose passion in the direction of the will' (*Treatise*, Book II, Part III, Section III; p. 413 in Hume 1978). In more modern terms, Hume's view is thus that beliefs by themselves are incapable of motivating action. Beliefs need to be supplemented by some desire or 'pro-attitude'. (Since precision won't matter, and I have a con-attitude to 'pro-attitude', I'll continue to talk of 'desires'.)

The non-factualist wants to show that SEAs do not express beliefs. Given Hume's principle, it is enough to show that the attitudes expressed in moral judgements are capable of motivating action; for this puts them on the desire or non-belief side of Hume's distinction. And the motivating character of moral judgement (or indeed of evaluative judgement in general) has usually seemed self-evident.

The non-factualist thus offers the following argument: moral commitments – the attitudes expressed in moral judgements – have a distinctive conceptual role. Briefly, they motivate actions. Hume's principle tells us that no belief, of itself, could fulfil this role. And hence the factualist, who takes SEAs to express beliefs, has to resort to the claim that such beliefs are 'accompanied by' attitudes that can play this role. The factualist faces the moral equivalent of the confidence problem – we might call it *the approval problem* – that of explaining, for example, why the belief that it is a good thing that it is raining should be associated with an attitude of approval towards the fact that it is raining. On the

face of it, the non-factualist will suggest, it would seem that ‘if moral commitments express beliefs that certain truth conditions are met, then they could apparently co-exist with any kind of attitude to things meeting the truth conditions’ (Blackburn 1984, p. 188).

Hume’s principle is clearly crucial. Someone who rejects it can agree that moral judgements express mental states on the active or motivating side, and yet insist that moral commitments simply are beliefs about moral facts. This will be the ethical equivalent of the radical factualism we have just encountered in the probabilistic case. In both cases the non-factualist has attempted to burden the factualist with a cumbersome and implausible duplication in the mental states associated with utterances whose status is in question (SPAs and SEAs, respectively). The radical factualist responds by accepting the model but denying that it involves any duplication: rather, she suggests, the non-factualist has confused the existence of two ways of describing a single mental state for the existence of two distinct mental states.

We shall return to Hume’s principle in chapter 5. Before that, however, I want to show the relevance of these issues of functional role to another prominent metaphysical argument for ethical non-factualism. I want to show that the argument cuts no ice with the radical factualist; and also, incidentally, that it has an analogue in the probabilistic case.

4.6 THE SUPERVENIENCE ARGUMENT

Moral properties appear to supervene on natural or non-moral properties: it seems uncontroversial that if two things are taken to be the same in all non-moral respects, then they should not be distinguished in moral terms. The problem is to explain this evident constraint on moral judgement. Simon Blackburn has argued that only a non-factualist has an adequate explanation. I want to show that the radical factualist does just as well.

Blackburn first presents the argument in a paper entitled ‘Moral realism’ (1971). He returns to it, changing the emphasis to some extent, in (1984) and (1985). Blackburn says, for example, that the explanation of supervenience is

especially hard for the realist [our factualist]. For he has the conception of an actual [moral] state of affairs, which might or might not distribute in a particular way across the [non-moral] states. Supervenience then becomes a mysterious fact, and one which he will have no explanation of (or no right to rely upon). (1984, p. 185)

As far as realism can show us, it could be true that the moral floats quite free of the natural. (1984, p. 221)

Blackburn wants to show that the non-factualist (or anti-realist) can explain the supervenience of the moral on the non-moral. In his first version of the argument he suggests that a

natural way of explaining the supervenience of moral properties would, for an anti-realist, be something like this. There can be no question that we often choose, admire, commend, desire, objects because of their naturalistic properties. Now it is not possible to hold an attitude to a thing because of its possessing certain properties and, at the same time, not hold that attitude to a thing which is believed to have the same properties. The non-existence of the attitude in the second case shows that it is not because of the shared properties that I hold it in the first case. (1971, p. 114)

Blackburn acknowledges that the factualist ‘need not refrain from talking about moral attitudes’: on the contrary, ‘it is just that [the factualist] thinks this is a less clear alternative to talking about moral

beliefs' (1971, p. 115). But he thinks that 'if we rephrase the preceding paragraph in terms of moral belief, it is obviously insufficient to explain supervenience' (p. 115). For the claim would be that

it is not possible to hold a moral belief about a thing, believe a second to be exactly alike in all naturalistic respects, yet at the same time not hold the belief about the other thing. But this doesn't explain supervenience at all: it merely shows the realist putting conditions upon what can be *believed* to be the truth, not upon what *is* the truth. Our belief, he is saying, has to be consistent across naturalistic similarities – but this is no explanation of why, on his theory, the truth has to be. (p. 115)

Blackburn's claim thus seems to be that the factualist has trouble justifying moral consistency – justifying the policy of applying the same moral predicates in the same non-moral circumstances. The non-factualist is better placed not because he can justify such a policy, but because he doesn't need to. The need for justification stems from the factualist's conception of truth – in particular, from the principle that the truth is independent of what it is possible for us to believe. Both sides have a more or less causal explanation of the fact that we follow the consistency principle: if moral attitudes are the effects of non-moral attitudes, consistency in practice is a consequence of the relevant causal laws. For the factualist, however, *de facto* fidelity is not enough. Consistency needs to be earned.

So the non-factualist escapes the supervenience problem – as construed in terms of consistency, at least – by rejecting the framework within which it arises. The problem arises because we talk of independent moral truths; and the non-factualist is free to reject that sort of talk.

Not all non-factualists will want to exercise this freedom, however. In particular the projectivist, far from rejecting factual talk, wants to account for it in non-factual terms. Blackburn repeatedly stresses this aspect of projectivism, regarding its explanatory potential as a major virtue. In particular, he thinks that the projectivist can explain our entitlement to doubt our own opinions, and to distinguish between our believing that something is the case and its actually being the case (Blackburn 1984, p. 219, for example). If so, then the projectivist is equally vulnerable to the charge that the causal explanation of consistency simply puts 'conditions on what can be believed to be the truth, not upon what is the truth'.

Whether the projectivist has an answer to this charge will no doubt depend on the details of the constructive account of truth, in terms of which he proposes to explain the factual appearances. But if the projectivist can secure supervenience by this means then what is to stop a factualist from doing the same? In the (1971) version of the argument, Blackburn takes it for granted that factualism arrives with its own account of truth. His target is the traditional realist, whose notion of truth is anchored to a conception of a 'real' external world. For present purposes, however – as to a large extent in Blackburn (1984) – the projectivist's more threatening opponent is the idealist. As noted earlier, idealism should here be counted a version of factualism; for it rejects the non-factualist's attempt to divide indicative discourse into fact-stating and non-fact-stating parts. And a factualist of this sort will have no qualms about helping herself to the projectivist's account of moral supervenience. Thus we find once again that the non-factualist's main problem is to stop the projectivist bandwagon – and that the better projectivism starts, the harder it is to stop.

Perhaps for this reason, the consistency issue gets less attention in Blackburn's more recent presentations of the supervenience argument. Of course, no version of the argument can hope to exclude idealism. As an argument for non-factualism, it simply bears on the wrong side of the issue. But at least the damage might be controlled: we might hope to present the argument in such a way that it favours projectivism about ethics without equally favouring projectivism about everything. Alternatively (though this would be less to Blackburn's own tastes) we might hope for a version of the argument that would favour a non-projectivist non-factualism about ethical discourse. It turns out, however, that the prospects of either course depend on our ability to refute the radical factualist.

Blackburn now offers the projectivist the following explanation of moral supervenience:

When we announce the [moral] commitments we are projecting, we are neither reacting to a given distribution of [moral] properties, nor speculating about one. So the supervenience can be explained in terms of the constraints upon proper projection. Our purpose in projecting value predicates may demand that we respect supervenience. If we allowed ourselves a system (shmoralizing) which was like ordinary evaluative practice, but subject to no such constraint, then it would allow us to treat naturally identical cases in morally different ways. This could be good shmoralizing. But that would unfit shmoralizing from being any kind of guide to practical decision-making (a thing could be properly deemed sbetter than another although it shared with it all the features relevant to choice or desirability). (1984, p. 186)

This might be taken as a pragmatic justification of the consistency principle – as if Blackburn were suggesting that in view of the role of moral judgements in influencing our practical decisions, it would be unwise, though not impossible, not to make one's moral judgements in accordance with that principle. But there is a more interesting and I think more accurate interpretation. We have noted that moral beliefs, indeed evaluative beliefs in general, appear to play a special role in guiding choice. A (positive) value judgement about a state of affairs is a *prima facie* reason for choosing that state of affairs over others not so valued. As Blackburn says, 'It seems to be a conceptual truth that to regard something as good is to feel a pull towards promoting or choosing it, or towards wanting other people to feel the pull towards promoting or choosing it' (1984, p. 188). To value A higher than B is to be disposed, other things being equal, to choose A in preference to B.

In order to explain supervenience, we need to add to this the plausible claim that in so far as they affect us, the outcomes of actions are completely determined by their natural or non-evaluative consequences. The value judgement guides our choice of destination, but the success or failure of the expedition depends on what we actually find when we get there. (I am simplifying here. In practice, the route from valuation to natural consequences of choice may be multi-staged: as, for example, when we judge that it is better to take the high road than the low road, because the scenery is more attractive on the high road.) This means that to imagine a case of moral difference without natural difference would be to imagine on the one hand that choice A would be preferable to choice B; but on the other that since the two choices have the same natural consequences, it would make no difference which we choose. This conceptual conflict seems a plausible basis for our belief in moral supervenience – indeed, for the supervenience of the evaluative on the non-evaluative in general.

If this explanation is on the right lines, then the supervenience of the evaluative on the non-evaluative depends on the principle that outcomes are judged, in the end, by their natural consequences. This principle might conceivably be challenged. On theological grounds, for example, someone might say that the outcome of an action can have an intrinsic value that is independent of its natural consequences (at least in this world). But this sort of challenge in fact works in favour of the above explanation. To admit intrinsic values of this kind would surely be to deny supervenience – to allow values to 'float free' of the natural, in Blackburn's phrase. However, the above explanation of supervenience turns on nothing more than the action-guiding role of evaluative attitudes. We already know that the factualist has a problem explaining this – it is the approval problem, the evaluative analogue of the confidence problem. But we saw that if the factualist is prepared to claim that evaluative attitudes simply are beliefs, the problem seems to be dealt with. And in the present context, the same identification will entitle the factualist to this explanation of evaluative supervenience. So it seems that the supervenience problem is not an objection to the radical factualism we have encountered in this chapter: a factualism which accepts the non-factualist's positive account of the mental states characteristically expressed in evaluative judgements, but insists that the mental states so described are themselves the beliefs whose existence the projectivist denies.

In his most recent discussion of the supervenience argument, Blackburn says that the explanation of supervenience depends

crucially upon the role of moralising being to guide desires and choices amongst the natural features of the world. [But] if, as a realist ought to say, its role is to describe further, moral aspects of [reality], there is no explanation at all of why it is constitutive of competence as a moralist to obey the [supervenience] constraint. (1985, pp. 56- 7)

Our radical factualist agrees that the explanation of supervenience lies in the fact that moralizing has the former role; agrees also that it has the latter role; but denies that the two roles are incompatible.

*4.7 SUPERVENIENT PROBABILITY

Blackburn presents the supervenience argument as a problem for the moral realist. In view however of the parallels we have already noted between the ethical and probabilistic cases, it is natural to wonder whether there is an analogous argument in the probabilistic case. Do probabilities supervene on non-probabilistic characteristics of things? I want to show that there is a parallel here, that again the best explanation turns on the action-guiding role of the relevant sort of attitude, but that here also the radical factualist has an answer.

Suppose then that we accept that there is a certain probability of a specified outcome to a given event – let us say an $n\%$ chance that a given experiment will result in the emission of an electron. Do (or should) we allow that, with all non-probabilistic facts held fixed, there could have been some different chance of this result? The answer depends in part on our view of chance. If we think of the chance in terms of the actual frequency of such outcomes in similar experiments, or some such, the appropriate answer seems to be ‘yes’. If we are talking about actual class ratios in finite classes, we have no difficulty accepting that those ratios could have been different, even if other aspects of the world had been the same. The possibility is illustrated by the man who broke the bank at wherever it was: because his fortunes depended on indeterminate events, he could equally well have lost his shirt.

Partly for this reason, however, actual frequencies have never provided a plausible approach to single case probabilities. It is more natural to suppose that the chance that an experiment will have a certain result is a further fact, over and above the facts about frequencies in actual similar cases. Could then this further fact vary independently of the other properties of the experiment – including, now, the actual relative frequencies themselves?

I think that we are inclined to deny that chances could vary in this way. The reason is nicely analogous to the above explanation of the supervenience of the moral on the non-moral. Suppose for example that two experimental set-ups are identical in (what we take to be) relevant non-probabilistic respects; and, most importantly, that they yield the same relative frequencies of corresponding outcomes. Let us try to imagine that the two set-ups in fact have different chances of some outcome – electron emission, say. In set-up A there is a high chance of emission; in set-up B, a low chance. The problem with such a supposition stems from the role of probabilistic beliefs in decision making. Roughly speaking, the higher the probability of an outcome, the more reasonable it is to bet on it, at any given odds. This means that on the given assumptions, we shall be inclined to prefer a bet on electron emission in set-up A to the same bet in set-up B. However, since there is no difference in the relative frequency of such an outcome in the two cases, a gambler who consistently bets on electron emission will do no better in case A than in case B. So there is no reason to prefer one bet to the other. This contradiction seems to show that chances do supervene – in the same sense as in the ethical case – on non-chance factors.

As it stands, however, the argument is vulnerable to a serious objection: namely that the fact that the two cases show the same relative frequency of electron emission does not entail that the same betting policy will have the same outcome – let alone that a *single* bet would have the same outcome in both cases. For individual outcomes may differ, even if the overall frequencies are the same. Of course, if we could take for granted that identical relative frequencies entail identical probabilities, then we could say that the *probability* of success is the same for both bets (either in the single case or in the

long run). But to take this for granted is to assume what we are trying to establish: the supervenience of probabilities on the non-probabilistic.

To escape this objection we have to strengthen our premisses. We need to suppose that set-ups A and B are non-probabilistically identical in such a way as to justify the following claim: Any trial on either set-up *would have had* the same outcome if performed on the other. It then follows that in any given case (and at any given odds), a bet on electron emission would have the same result on either set-up. And this conflicts with our preference for a bet on A, stemming from the assumption that there is a higher chance of electron emission in A than in B.

Thus, on sufficiently strong assumptions, we find that single case probabilities do supervene on non-probabilistic features of the world. As in the moral case, the argument makes essential use of a special association between a certain sort of mental attitude – here (what the factualist will regard as) a belief with probabilistic content – and choice behaviour. We saw that this association may itself seem problematic for the factualist: it is the basis of the confidence problem. However, the radical factualist's reply was that the so-called 'association' is actually an identity. We now find that in the probabilistic as in the ethical case, such a strategy is not to be dislodged by the observation that the disputed 'facts' appear to supervene on facts of other kinds. The radical factualist explains this observation as well as anybody.

5

Fact and Psychological Function

We found that in the ethical and probabilistic cases, non-factualism can be driven to seek its foundations in the head – to try to analyse in psychological terms the semantic distinction between fact-stating and non-fact-stating discourse. Statements thus come to be characterized as (what are normally) expressions of belief; non-statements – or those that the non-factualist is concerned with, at least – as the (normal) expressions of other sorts of propositional attitude. And this appeal to psychology is no mere expository gloss. In order to defeat the radical factualist, the non-factualist must claim to be able to distinguish beliefs from other attitudes, in advance of settling the semantic issue.

We have not shown that the non-factualist's dependence on psychology goes beyond the ethical and probabilistic cases. All the same, it is clear that non-factualists about various other topics do in fact appeal to psychology. One important case is the non-factualist approach to indicative conditionals. This relies on the observation (originally due to Ernest Adams 1965) that the subjective assertibility of an indicative conditional goes by the subjective conditional probability of its consequent given its antecedent; rather than say by the absolute probability of the material conditional. The view thus involves that indicative conditionals are typically expressions of high subjective conditional probabilities; and – crucially, if the account is to be non-factualist – that these cannot be construed as a kind of belief. Another case is the so-called 'epistemic' treatment of possibility and necessity, and the related treatment of laws of nature as generalizations towards which we hold a certain sort of attitude (a view advocated by Ramsey 1978, chapter 6B, for example).

I think we are thus entitled to conclude both that non-factualism has often in fact put its faith in psychology; and that at least in certain central cases it is bound to do so. Our final sceptical task is to show that this faith is misplaced. Beliefs cannot be adequately distinguished from other propositional attitudes; not at any rate unless we are prepared to rely, as the non-factualist here cannot, on a characterization of this psychological distinction in semantic or metaphysical terms.

The chapter is in three main parts. I first consider the Humean approach to the belief/desire distinction, based on Hume's account of the origins of action. This account continues to be widely taken for granted, even in these anti-Humean times. As a result, the task of explaining action in psychological terms may seem a promising source of the sort of distinction we are looking for. I shall argue that the promise is empty, however. In search of the source of the belief/desire distinction the modern philosopher ends up, as Hume himself did, appealing once more to semantics: claiming that beliefs, unlike desires, have truth conditions.

The second part of the chapter reaches a similar conclusion along a more general, more contemporary and perhaps initially more promising route. It considers the suggestion that we look to recent cognitive theory for an account of the *content* of a propositional attitude. A species of propositional attitude is a class of beliefs if and only if there is a class of contents such that the species concerned is the set of beliefs with contents of that class. (This is the psychological analogue of the condition for a species of utterance to be represented as a class of assertions.) Hence a basis for the belief/non-belief distinction might be expected to be a corollary of an adequate account of content. I shall argue, however, that contemporary accounts of the nature of content do nothing to settle this issue – or rather they settle it, if at all, in the semantic or metaphysical terms that the non-factualist has already renounced.

Finally, I shall turn to some arguments claimed to show that indicative conditionals are not expressions of genuine beliefs. In view of the logical role of conditionals, this is potentially the most significant non-factualist theory of all. My approach will again be deflationary: I shall try to show that

these arguments parallel some striking but clearly inconclusive objections to a factualist treatment of ordinary probabilistic judgements.

5.1 DISTINGUISHING DESIRES: THE HUMEAN STORY

The Humean argues that reason is incapable of action – in modern form, that beliefs are motivationally inert – and hence that an action explanation is incomplete unless it appeals, at least implicitly, to some sort of desire or motivating attitude. The view is as orthodox as it was once controversial. Indeed, it is at the heart of so-called ‘folk psychology’ (a fact that should not be uncritically attributed to the folk influence of Hume, one suspects). I want to show that the Humean argument is inadequate, however, even in its most charitable contemporary reconstruction. It fails to show that belief cannot motivate action; and hence it fails to show that moral commitments (even if agreed to motivate actions) are not beliefs.

Our interest in the Humean argument stems from its application to these semantic questions. But it may be as well to point out that in challenging the principle that belief is non-motivational, we are not committed to rejecting totally the orthodox view of the explanation of action. For in one crucial respect Hume’s conclusion is stronger than the orthodox belief/desire model requires. It would be in keeping with the spirit of this view to say that the explanation of action involves two sorts of belief: beliefs about the desirability, or value, of certain states of affairs; and beliefs about the bearing of the action in question on the obtaining of those states of affairs. This modified view could agree, in any actual case, on what constitutes an explanation of an action. It differs from the Humean only as to how the attitudes mentioned by such an explanation should be described.

This difference is crucial to the use made of Hume’s principle by non-factualists in ethics. The non-factualist needs to show that moral statements lie on the right side of a psychological distinction – in other words, that they do not express beliefs. Given Hume’s principle, it only needs to be shown that moral commitments do motivate actions; for this puts them on the desire side of Hume’s distinction. Clearly, however, someone unconvinced by the Humean argument could quite well endorse the standard model of action explanation; accept that moral judgements express mental states on the active or motivating side; and yet insist that moral commitments are genuine beliefs about moral facts. This was the response of the radical factualist we met in chapter 4.

Such a view admits two possible accounts of the place of desire. One account would recognize a category of desires, distinct from beliefs, but claim that motivation is not the exclusive preserve of attitudes of this category (moral commitments being one sort of exception). The problem will then be to say what distinguishes beliefs from desires, given that it is not the power to motivate action. This problem might incline us to the second sort of account, recognizing no significant foundation for a distinction between beliefs and desires. This view would further divide into two possibilities: the substantial claim that all desires are beliefs; and the weaker sceptical claim that there is simply no adequate justification for not so representing an arbitrary propositional attitude. Our interests again lie with the sceptic.

Let us turn then to the case for Hume’s principle. Why should not beliefs be motivational? The Humean’s first line of attack might be something like the following argument. It is intended to show that the belief that a contemplated action A would have some property F (or would bring about an outcome having property F) cannot in itself constitute a reason for an agent to perform A. If there were such a case (the argument claims) we could imagine a similar agent who would dislike F, and would hence be motivated not to do A by a desire that not F. This being so, the action A cannot in fact have been caused by or be explained in terms of the agent’s belief that A would be F. It is also relevant that the agent desired that F, or at least did not desire that not F. In other words, the causal or explanatory story needs to mention the agent’s desires.

I shall call this *the imagined dislike argument (IDA)*. Hume himself appeals to IDA, or something very like it, in this notorious passage:

'Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger. 'Tis not contrary to reason for me to chuse my total ruin, to prevent the least uneasiness of an *Indian* or person wholly unknown to me. (1978, p. 416)

Hume's immediate point is that reason does not control the passions. But given that the passions do control action, the argument becomes IDA – as indeed the latter example illustrates, in speaking of choice rather than preference. Hume's suggestion seems to be that the belief that an action will prevent his total ruin will not explain his performing that action, because he might have desired total ruin (in preference, at any rate, to some other consequence of preventing ruin, such as the unease of another party).

IDA has two major flaws, however. The first is to rely on the principle that if a causal explanation would be inadequate in the *presence* of some extra condition, then it is inadequate unless it mentions the *absence* of that further condition. This principle is no more plausible here than it is for causal explanations in general. In almost every case, causes are at best sufficient in *the circumstances* for their effects; and to insist on an explicit description of the relevant circumstances would be to cry for the moon. (We rely on *ceteris paribus* clauses because we cannot fill in these gaps.) In the present case, to admit such a principle would be to abandon any systematic approach to the psychological explanation of action. After all, there is no end to the things that even the most fastidious agent does not dislike.

IDA's second fault is just as damning: given that the issue is whether desires are anything but a special class of beliefs, the argument simply begs the question. For if some beliefs are desires, to believe that an outcome would be F might be the very same thing as desiring that outcome. In this case the possibility that the argument imagines – that of someone holding the belief, but desiring an outcome such that not-F – would be that of a person who desires what he takes to be undesirable. An advocate of the view that desires are a species of belief seems entitled to regard this as no possibility at all.

Neither of these objections is particularly profound. So it may seem doubtful whether IDA does justice to Hume's distinction, or explains its contemporary appeal. However, I think the argument does point us in the right direction. It seems to derive its initial plausibility from a model of the mind which we find intuitively appealing. Perhaps it thus assumes what it is supposed to establish. But at least it has the virtue of directing us to the crucial assumption.

We may think of the issue in terms of a simple model of the mental structure of an intentional agent. The functional role of desire seems to be to connect two other components: on the one hand, a unit (*Cog*) responsible for cognitive deliberation; on the other, a unit (*Will*) whose function is to initiate bodily action, in response to the outputs from *Cog*. There are several possible models of *Will* and its relationship to *Cog*. At one extreme we could regard it as a kind of store, or processing house, for intentions already formed in *Cog*. At the other extreme we could take it to be the unit in which judgements arriving from *Cog* are matched with a store of desires to yield intentions. However, since we want a model in which the action that interests us takes place at the boundary between *Cog* and *Will*, let us avoid both extremes. Let us view *Will* as a structure on which intentions 'condense', as a consequence of suitable states of *Cog*.

In these terms, Hume's principle is that no intention can condense in the absence of a desire – that desires are a necessary catalyst to the intention forming process. IDA rests on the claim that whatever beliefs might be present in *Cog*, we could imagine a desire whose presence would prevent a given intention condensing from those beliefs. The argument is a poor one, but catalytic view of desire has a lingering appeal.

However, we know that in the chemical case, altering either the composition of a vapour or the nature of the condensation chamber may make a catalyst redundant. So the analogy suggests that we consider the effect of changes in the corresponding aspects of the mental model: to Cog on the one side, and Will on the other. As yet we have no reason to regard these factors as cognitive constants. So there is nothing to stop us trying to imagine agents who differ in these respects, from each other, and more importantly from the way the Humean view takes us to be. The appeal of the orthodox view may rest on the fact that we have not yet noticed the alternatives.

There are two main possibilities to consider: the causal gap supposedly filled by desires might be bridged from either side. The first possibility would be to expand Cog, so as to admit beliefs with the capacity to produce intentions in a Will of the orthodox kind. The second would be to endow Will with the disposition to form intentions in the presence of certain orthodox beliefs. (I ignore the obvious hybrid possibilities.)

In the former case, we want a belief whose capacity to rationalize an action does not depend on an agent's possession of certain desires. It seems to me that we can get a good idea of what such a belief would be like, by reflecting on familiar expressions of moral or rational necessity – 'I am bound to tell the truth', for example, or 'I must go down to the sea again.' Superficially, at any rate, utterances such as these can be regarded as expressions of beliefs that possible actions have certain properties. Moreover, such beliefs (if they really are beliefs) seem to provide reasons for actions which are insensitive to an agent's desires about actions with the corresponding properties. The belief that one must go down to the sea seems capable not merely of making one do so, but of overriding, in the process, considerable distaste for messing about in boats.

These points are notoriously disputable. There are two lines of attack. One will urge that believing that I must do something, or am bound to do it, only provides me with a reason for doing it because, like everybody else, I have an implicit desire to do what I must. The other will take the opposite tack and argue that because utterances such as these express reasons for actions, they do not express beliefs. However, both lines of argument seem to depend on the very claim in question: namely, that beliefs alone cannot motivate actions. To settle the issue, we would need a criterion with which to distinguish beliefs from desires. Hume's argument claims to offer such a criterion, saying that desires are distinguished in virtue of their active or motivating role in the production of action. But it now appears that the argument itself depends on the very criterion it is intended to supply. As McDowell puts it,

the idea of the world as motivationally inert is not an independent hard datum. It is simply the metaphysical counterpart of the thesis that states of will and cognitive states are distinct existences; which is exactly what is in question. (1978, p. 19)

Thus it seems that we have no non-circular reason to deny the possibility of an intrinsically rationalizing belief: a belief whose capacity to motivate and rationalize a contemplated action does not depend on co-present desires (even if its motivational effect could be overridden by certain such desires, were they to be present). And if we have reason not actually to affirm the possibility, it lies partly in the fact that the distinction between the belief and desire is itself so obscure.

I said that there are two possible ways to bridge the claimed gap between Cog and Will. The second is to imagine Will itself equipped to handle a somewhat broader range of inputs. Why shouldn't Will, or the process by which an agent forms intentions to act, simply be built to respond to certain beliefs about the properties of actions and states of affairs?

Consider, for example, a classical economic rationalist, whose actions are based entirely on the profit motive. Such an agent forms intentions to perform those actions he believes will maximize his net monetary return. One might object that the rationalist *wants* to make a profit. But why should he? Why should his mind not simply be organized so that believing that an action would maximize his

profits normally leads him to form the intention to undertake it (in just the way, and to whatever extent, belief/desire combinations are ordinarily supposed to lead to intentions)?

One possible answer appeals to Davidsonian themes. We ourselves wouldn't be motivated to act as the rationalist does unless we wanted to make a profit, so we can only make sense of his behaviour – rationalize it – by ascribing to him such a desire. This objection might lead us too far astray. In making the ascription of attitudes to cognitive systems radically dependent on the ascriber's own system of attitudes, it threatens the entire project of speculative psychology. Fortunately, we can evade it. We want to know, *inter alia*, whether our own Wills are ever activated by beliefs alone. But in the face of a purported real case of this sort of thing, the Davidsonian move will simply be powerless. It cannot insist that *our* rationalization of the action in question would need a desire, for that will be the very point at issue.

I conclude that motivating beliefs are a conceptual possibility. If we begin with an orthodox Humean model of ourselves, we can imagine two kinds of psychological modification that would allow beliefs to motivate actions. The effect of this is not simply that the practical question of the nature of the motivation of action remains unsettled, but rather that the question itself is under threat. Hume's principle not only entails that action requires desires, as well as beliefs; it also offers a criterion for distinguishing desires from beliefs. Without it, the non-factualist lacks a foundation for that distinction.

Like Hume himself, the non-factualist may therefore feel inclined to try another tack – to say that desires are distinguished from beliefs in virtue of the fact that it is 'impossible ... they can be pronounced either true or false' (1978, p. 458). This would be to distinguish beliefs in virtue of their semantic properties; and it is a popular move. However, we shall see that it takes the non-factualist back to positions already abandoned – positions whose untenability forced the non-factualist to seek a psychological basis for the fact/non-fact distinction.

This may seem an immediate consequence of such an appeal to semantic properties. After all, the attempt to analyse assertion in terms of truth was the first target of our sceptical challenge. For the moment, however, a significant possibility remains open to the non-factualist. Psychological truth might succeed where linguistic truth failed. If semantic properties apply primarily to mental states, and only derivatively to language, then the arguments of chapter 2 may have missed the mark. Truth might indeed be insubstantial, but nevertheless play a crucial role in establishing a psychological basis for the fact-stating/non-fact-stating distinction.

To exclude this case, we shall need to consider some possible foundations for psychological semantics. This brings us to the concerns of contemporary cognitive theory. And it brings us to another possible response to the argument of this section. We have seen that the requirements of action explanation do not yield the psychological distinction the non-factualist is looking for. On reflection, however, this avenue was always rather limited. At best it offered a distinction between beliefs and the range of motivating attitudes we have loosely called 'desires'. Even if this had met the needs of the ethical non-factualist, it would have been largely irrelevant to other non-factualists. If beliefs do not motivate actions, then neither, for example, do the partial beliefs invoked by the non-factualist account of probability.

From this point of view, our mistake may seem to have been to concentrate on too small an aspect of the functional organization of the mind – to concentrate on the initiation of action, to the exclusion of other functions of the mind. As it happens, this is exactly the mistake of which contemporary functionalist accounts of the mind accuse their behaviourist predecessors. In the next section we want to enquire whether these contemporary theories succeed where Hume failed, in distinguishing beliefs from non-beliefs.

5.2 FUNCTION AND COGNITIVE CONTENT

We may seem to be faced with a daunting diversity of contemporary views on the nature of the mind. One of Professor Fodor's relatives 'sighs for the days when well-brought-up philosophers of mind kept themselves occupied for hours on end analysing their behavioural dispositions' (Fodor 1985, p. 76). But things are not as bad as they seem. For one thing, as Fodor puts it, 'the chaotic appearances are actually misleading. A rather surprising amount of agreement has emerged, if not about who's winning, at least about how the game has to be played' (1985, p. 76). We can exploit this agreement – and Fodor's guide to the options – for our present purposes.

More importantly, however, it turns out that there is very little to say: none of the contemporary options tries to analyse the distinction we are looking for, or indeed makes a point of claiming to have the resources for doing so. Most contemporary theories simply take for granted the Humean model of action explanation – in the guise, as they put it, of 'folk psychology'. This provides the framework within which they tackle such questions as the nature of content. And those that do not assume the belief/desire model of action explanation are those that reject folk psychology, and with it the apparatus of propositional attitudes. So on neither side is there a developed claim to provide a basis for the distinction between beliefs and other sorts of attitudes; nor is there any prospect of such a basis.

This claim may seem implausible. After all, we know that we can think of the question as to whether desires (say) are a kind of belief, as that of the existence of an appropriate species of *content*. Roughly, desires can be represented as beliefs if and only if there is a content modifier *Des* such that the desire that P can be represented as the belief that *Des*(P). And cognitive science has had a great deal to say about the notion of content. Surely there is something that will tell us whether there could be such a *Des*?

It seems not. The point is perhaps most clearly illustrated in what Fodor calls 'the Representational Theory of Mind' (*RTM*), according to which propositional attitudes are relations to representations in a 'language of thought'. Fodor gives the following 'canonical formulation of this view':

For any organism O and for any proposition P, there is a relation R and a mental representation MP such that: MP means that (expresses the proposition that) P; and O believes that P iff O bears R to MP. (And similarly, R desires that P iff O bears some *different* relation R', to MP. And so forth ...) (1985, p. 88)

Clearly, the distinctions between the propositional attitudes are here built in from the start: the desire that P is interpreted in terms of the relation R' to MP, rather than in terms of the relation R to DES(MP) (DES being the mental representation of the content modifier *Des*).

From the Humean point of view this looks perfectly natural. If we question the Humean categories, however, it begins to seem rather arbitrary. For mental representations and the attitude relations R, R', etc., are both individuated in the same kind of way: in terms of their causal or functional roles. One of the great attractions of the RTM is its ability to exploit the computer metaphor, and conceive of mental processes as 'transformations of mental representations' (Fodor 1985, p. 95). Such transformations are causal processes whose outcomes depend on two factors: the syntax of the mental representations with which the process begins, and the mode in which those representations are initially stored – whether they are 'believed', 'desired' or whatever. But why not reduce these two factors to one, by regarding the mode of storage as an aspect of syntax? After all, in any actual case the available syntax will be capable of representing an enormous range of distinctions. It is surely capable of a handful more.

It is true that if we knew the syntax of the language of thought – in advance, as it were – then the question of representing mode of storage as a syntactical distinction would not arise. The syntactical guidebook would tell us that there was no way of representing certain distinctions within the language; leaving us no choice but to regard their role as that of a mode of storage of a sentence of the

language. The non-factualist cannot hope for an answer of this kind, however. For what is the semantic role of the sentences of the language of thought? Simply that of the fact-stating sentences of natural language – or so the non-factualist must say, for it is these sentences he takes to express beliefs. In other words, the task of delineating the syntax of the language of thought is simply the original task of delineating fact-stating discourse: the very task for which the non-factualist now wants to distinguish beliefs from non-beliefs.

So the non-factualist is not entitled to appeal to the syntax of a language of thought, in order to distinguish beliefs. Indeed, no one is entitled to appeal to the syntax of a language of thought in order to determine which functional roles should be counted content-determining and which attitude-determining. For (as Mellor 1988 emphasizes) semantics must precede syntax, in this sense: our grounds for regarding a system of states and causal processes as a syntactic system of representations and transformations must be that we assign its elements a semantic role. The title ‘linguistic’ has to be earned.

There is little agreement on how it should be earned. Fodor again: ‘Of the semanticity of mental representations we have, as things now stand, no adequate account’ (1985, p. 99). It seems to me that proponents of the RTM have not always been sufficiently aware of the extent to which it depends on the assumption that this problem is solvable, at least in principle. For how else could an opposing view – that of Stich (1983) – present itself as a *purely* syntactic account? Stich says that the

core idea of the STM [Syntactic Theory of Mind] – the idea that makes it *syntactic* – is that generalizations detailing causal relations among the hypothesized neurological states are to be specified indirectly via the formal relations among the syntactic objects to which the neurological state types are mapped.’ (1983, p. 151)

But how is this different from a formal description of (an aspect of) the causal structure – say an account of the operation of a telephone exchange that describes it in terms of a formal model of its switching operations? In the absence of the semantic notion of representation, the label ‘syntactic’ is quite inappropriate.

Stich evidently regards the significance of the ‘syntactic’ character of his theory of mind as more than that of a mere formal description. This shows up in the fact that he entertains (though he does not insist on) versions of the STM that follow folk psychology in postulating ‘two distinct classes of states underlying behaviour, belief-like states and desire-like states, with state types of *both* classes mapped to a single class of formal objects’ (1983, p. 151). Clearly, if the syntactical story was simply a formal model there would be no reason not to treat the belief/desire distinction as itself syntactical (in the manner, presumably, of the belief/disbelief distinction). Belief-like states and desire-like states would then get mapped to different but related classes of formal objects (and we would again have no reason not to regard desire-like states as a special kind of belief-like state). From a formal point of view, this is the obvious approach; and only the vestiges of the semantic point of view of folk psychology could recommend the alternative.

I emphasize that this is a criticism only of Stich’s ‘syntactic’ characterization of his theory, and not of the theory itself. On the contrary, I think Stich’s articulation of a semantics-free approach to the mind is of considerable dialectical value to the proposal I want to develop in part II (where the task, in a sense, will be to explain how the semantics gets in). The effort would be misguided, to say the least, if it could not begin from a foundation that does not already include what we want to put in. Stich has made it very much less implausible that there could be such a foundation.

Back to the RTM: it might be objected that in suggesting that this view leads us back to the claim that desires are distinguished from beliefs in virtue of their lack of truth conditions, I have ignored the RTM’s account of content. Given this account, it might be suggested, the RTM can characterize the sentences of the language of thought without recourse to the semantic distinctions of natural language

(and hence, for our purposes, without leading us back to our starting point). In other words (the suggestion would go) the RTM can explain the fact that it does not represent the desire that P in terms of DES(P): namely by showing that its account of content does not leave room for the necessary content modifier, *Des*.

This objection might be more forceful if there were an agreed approach to the problem of the content of propositional attitudes. As it stands, the closest thing to an agreed position is that content is individuated, in part at least, by the causal or functional roles of the attitudes concerned. This suggestion trades on the appealing hypothesis that the inferential role of propositional contents – the sort of thing that is traditionally supposed to individuate *propositions* – is modelled in the causal role of the relevant mental states. This association is at the heart of the computer metaphor: ‘computers ... just *are* environments in which the causal role of a symbol token is made to parallel the inferential role of the proposition it expresses’ (Fodor 1985, pp. 93–4).

Appealing as this suggestion is, however, it does nothing to justify the belief/non-belief distinction. On the contrary, it again takes for granted a distinction between those aspects of the functional role of a state token that determine its content, and those that determine that it is a belief, a desire, or whatever – a distinction that again looks particularly odd, if we bear in mind the ratio of content differences to attitude type differences in the finished theory (and take off our folk-psychological blinkers). It is as though the Taxation Department kept separate databases for its male and female clients: we would want to know why gender could not be represented explicitly, like everything else that the Tax Department knows about us. Why buy a filing cabinet, if you intend to keep everything in two or three colossal files? Or to put it another way, why start a separate file for desires, but not start one for disbeliefs, say, or beliefs about nutrition?

One last suggestion: perhaps we have concentrated on the wrong aspect of content. Many cognitive theorists think that there are two aspects to content: the internal aspect, that (more or less) matches internal functional role; and an external aspect, that depends on the actual reference of the attitude concerned – on what in the world it is in fact an attitude *about*. (See McGinn 1982, for example. Fodor himself now favours the view that content is entirely a matter of the external aspect; see Fodor 1987.) With the external aspect taken into account, the range of contents is circumscribed by the range of facts. But clearly this takes us back to our starting point. We wanted the belief/non-belief distinction in the first place because the direct metaphysical attacks looked so unpromising.

In summary, then, we find that if content is individuated in terms of functional role, there is no apparent reason to distinguish more than one type of propositional attitude; and that if it is individuated in terms of external semantic relations, the limits of belief depend on the limits of fact. Neither possibility provides what the non-factualist now needs in order to answer the sceptic or the radical factualist: a psychological criterion with which to distinguish beliefs from other attitudes, and thereby to distinguish fact-stating from non-fact-stating discourse.

*5.3 CONDITIONAL FACTS

There is one other important case in which what can be taken to be functionalist arguments have been widely thought to support a non-factualist conclusion. The case is that of indicative conditionals, and the arguments rest on Ernest Adams’s (1965) hypothesis, now widely accepted, that the ‘assertibility’ of an indicative conditional goes by the conditional probability of the consequent, given the antecedent (rather than by the probability of the corresponding material conditional, say).

To see what is meant by Adams’s claim, let us look at the ordinary case, with which Adams contrasts the behaviour of indicative conditionals. In what circumstances, in general, will a speaker be willing to assent to a statement S? As we have seen, the standard answer is something like this: When he or she believes that S, or at least has a partial belief (or degree of confidence) that S that in the context in

question is sufficiently close to a full belief that S. This answer is of course very rough, but not too rough. On the one hand, it is clearly a subjective condition, not immediately concerned with objective justification for assent to S. On the other, it does not imply that speakers have to think *about* their beliefs, in order to assent to S. We do as a matter of fact tend to assent to those propositions in which we have high degrees of confidence, the claim goes, even if we are rarely aware that this is what is going on.

We saw in chapter 4 that a determined factualist about probabilities will equate degrees of partial belief, or degrees of confidence, with beliefs about probabilities. The normal rule of assertibility then becomes the principle that assertibility goes by increasing probability. ‘Assertibility goes by increasing *subjective* probability’ is the way it is often put, but this can be a source of confusion. If initially applied to a degree of confidence, the term ‘subjective probability’ encourages the (factualist) view that for any partial belief there is an associated full belief about a probability. What probability is this? Why, the *subjective* probability! The term ‘subjective probability’ thus comes to refer ambiguously to the original degree of confidence, now thought of as a belief about a probability, and to the probability that figures in the content of this belief. In fact, of course, if degrees of confidence are beliefs about probabilities, then these probabilities need be no more subjective than the subject matter of any other belief. While the beliefs themselves are subjective only in the same sense: namely, that they are our beliefs, rather than somebody else’s. For this reason, I shall avoid ‘subjective probability’, instead referring to the mental states concerned as *credences*. The term ‘credence’ leaves it open as to whether, with the radical factualist, we should think of credences as beliefs about probability.

Adams’s hypothesis about indicative conditionals amounts to the claim that what in general determines whether we assent to such a conditional is not our credence in, say, the corresponding material conditional, but rather what we may call our *conditional credence* in the consequent, given the antecedent. For a factualist about probability a conditional credence is a belief about a conditional probability; and Adams’s hypothesis comes to the claim that the assertibility of an indicative conditional goes by perceived conditional probability, rather than by the perceived probability of a truth-functional conditional.

It is harder to say what a non-factualist about probability should make of the notion of conditional credence, and hence of Adams’s hypothesis. For one thing, if absolute (that is, non-conditional) credences are not beliefs about probabilities then, given Adams’s hypothesis, the orthodox definition of conditional probability –

$$(5.1) P(C/A) = P(C\&A)/P(A), \text{ so long as } P(A) \text{ is non-zero –}$$

seems to provide no suitable account of conditional credences. The main problem is that we often assent to an indicative conditional – and therefore, on this story, have a conditional credence – when we clearly have no corresponding absolute credence in the antecedent, the consequent, or their conjunction. I am confident that if it is raining in Moscow then the domes of the Kremlin are wet; but I simply have no idea as to whether it is raining there. (For this and related arguments, and a discussion of what the non-factualist should make of them, see Price 1986.)

However, our present concern is to show that despite certain arguments to the contrary, a factualist about probability can be a factualist about conditionals. Given the factualist’s view of absolute credences, it is easy to account for the above cases, and yet to reconcile Adams’s hypothesis with the standard definition of conditional probability. For we can believe that the *ratio* of two probabilities is high, without holding an opinion as to the value of either. The factualist about probability thus regards a conditional credence of degree *n* in C given A, as a belief that $P(C\&A)/P(A) = n$; and interprets Adams’s hypothesis as the claim that the assertibility of the indicative conditional ‘If A then C’ goes by the value of this ratio. Roughly, we assent to such a conditional when we believe that this ratio is sufficiently close to 1.

We can think of such principles of assertibility as fragments of a functionalist account of the mind. They belong to that part of such a theory that relates mental states to linguistic behaviour. As such, moreover, these principles suggest a ground for a non-factualist treatment of conditionals. For they suggest that the general functional account of assent behaviour, summed up in the principle that assertibility goes by degree of (absolute) credence, does not apply to conditionals. Conditionals seem to need a functional account of their own – in which case we have a theoretical basis for distinguishing them from utterances to which the general account does apply.

The argument rests on the assumption that the principle of assertibility for conditionals is not a special case of the general principle. In other words, it requires that conditional credence is not simply a species of absolute credence. If there were a conditional, ‘ \rightarrow ’ such that

$$(5.2) P(A \rightarrow C) = P(C/A)$$

the factualist could endorse Adams’s hypothesis, and yet, simply by claiming that ‘If A then C’ is to be interpreted as ‘ $A \rightarrow C$ ’, be able to say that for indicative conditionals, as for statements of fact in general, assertibility goes by absolute credence.

This suggestion was an early response to Adams’s (1965) formulation of his hypothesis (for example by Stalnaker 1970 – 5.2 is sometimes referred to as ‘Stalnaker’s hypothesis’). But there are two arguments that have been claimed to show that there can be no such conditional ‘ \rightarrow ’. The first is due to Lewis (1976) and the second to Carlstrom and Hill (1978). I want to show, however, that in both cases the supposed problem for this reading of conditional credence is already a fairly obvious but non-fatal feature of a factualist view of absolute credence – and a feature that a factualist about conditionals can accommodate in much the same way as a factualist about probabilities.

Lewis’s argument offers a *reductio* of the assumption that there is a conditional ‘ \rightarrow ’ whose probability is given by 5.2. The key is a standard principle of decomposition by cases – a straightforward consequence of the standard axioms of probability:

$$(5.3) P(X) = P(Y)P(X/Y) + P(\sim Y)P(X/\sim Y).$$

Substituting $A \rightarrow C$ for X and C for Y, 5.3 gives us

$$(5.4) P(A \rightarrow C) = P(C)P(A \rightarrow C/C) + P(\sim C)P(A \rightarrow C/\sim C).$$

Assuming for the moment the equality

$$(5.5) P(X \rightarrow Y/Z) = P(Y/X \& Z),$$

from 5.4 we derive

$$(5.6) P(A \rightarrow C) = P(C)P(C/A \& C) + P(\sim C)P(C/A \& \sim C)$$

and hence, since $P(C/A \& C) = 1$ and $P(C/A \& \sim C) = 0$,

$$(5.7) P(A \rightarrow C) = P(C).$$

From 5.2 it now follows that

$$(5.8) P(C/A) = P(C).$$

This conclusion is perhaps absurdity enough. It amounts to the result that any pair of propositions X and Y are probabilistically independent, provided only that $P(X \& Y)$ and $P(X \& \sim Y)$ are non-zero. But Lewis derives two further results: that there cannot be three incompatible propositions X, Y and Z that all have non-zero probability; and hence that the probability function has at most four values.

Lewis's argument uses 5.5, which he derives quite straightforwardly from the standard definition of conditional probability (5.1) and the assumption, in effect, that rational agents change their credences by *conditionalizing*: i.e., that on coming to believe that X, a rational agent assigns to any proposition Y a probability equal to the conditional probability she previously ascribed to Y given X. But the details need not concern us, I think, because 5.5 is more general than the argument requires, and we can get what we need for 5.6 a little more directly.

Given the conditionalization assumption ('COND', for short), it follows that $P(A \rightarrow C/C)$ equals $P_C(A \rightarrow C)$ – that is, the probability that (as rational agents) we would ascribe to $A \rightarrow C$, if we came to believe that C. But by 5.2,

$$(5.9) P_C(A \rightarrow C) = P_C(C/A)$$

and since $P_C(C) = 1$,

$$(5.10) P_C(C/A) = 1.$$

This shows that $P(A \rightarrow C/C) = 1$, and similarly we can use the fact that $P_{\sim C}(C) = 0$ to show that $P(A \rightarrow C/\sim C) = 0$, giving 5.7.

In order to show how the factualist might respond to Lewis's argument, I want to exploit certain similarities it bears to a classic argument for fatalism. This argument is generally agreed to be fallacious. But the key to the fallacy is the key to the factualist's response to Lewis (which is not to say that Lewis's argument is fallacious – more on this in a moment).

The traditional function of the argument in question is to encourage an audience to risk its life on the proponent's behalf. In the following fine example (for which I am indebted to Hugh Mellor), the proponent is Henry V, admonishing Westmoreland for wishing for reinforcements for their impending battle with the French:

... No, my fair cousin:
 If we are mark'd to die, we are enow
 To do our country loss; and if to live,
 The fewer men the greater share of honour. (Henry V, Act IV, Sc. III)

Philosophers are not insensitive to the pragmatic virtues of this style of argument. Leibniz (1951, p. 153) remarks that 'the sophism which ends in a decision to trouble oneself over nothing will haply be useful sometimes to induce certain people to face danger fearlessly.' He comments that 'it has been applied in particular to Turkish soldiers', but has doubts about its efficacy: 'It seems that hashish is a more important factor than this sophism, not to mention the fact that this resolute spirit in the Turks has greatly belied itself in our days.' Pragmatics aside, however, Leibniz, like most philosophers, is convinced that the argument is fallacious.

Its similarity to Lewis's argument stems from two factors: the first that Lewis's principle of decomposition (5.3) can be thought of as a generalization of the logical principle of argument by Dilemma, used by the fatalist; and the second that both arguments involve hypothetical judgements about conditionals. In the case of the fatalist argument, the latter characteristic may not be

immediately apparent. It is essential, however, for the fatalist must convince his victim not only that bravery is better in either eventuality, but also that the victim's choice will not otherwise determine the course of events. In the above case, for example, Henry must convince Westmoreland that whether (A) they fight without reinforcements is irrelevant to whether (C) they are victorious. For otherwise Westmoreland will rightly point out that their choice is not between A and not-A (A being better whether C or not-C). Rather it is between (A and not-C) and (not-A and C) – the latter presumably being preferable.

The fatalist thus argues as follows:

- 1 $C \vee \sim C$.
- 2 $C \rightarrow (A \rightarrow C) \& (\sim A \rightarrow C)$ – i.e., if C, then if A then C and if not-A then C, so that A is irrelevant to whether C.
- 3 $\sim C \rightarrow (A \rightarrow \sim C) \& (\sim A \rightarrow \sim C)$ – i.e., if not-C, then if A then not-C and if not-A then not-C, so A is again irrelevant to whether C.
- 4 Hence, by Dilemma, A is irrelevant to whether C.

It is premisses 2 and 3 that involve hypothetical assessments of conditionals.

As well as these two similarities, of course, there are some very significant differences between this argument and Lewis's. The two arguments are quite different in intent. Lewis's is a *reductio*. Indeed, the absurdity that Lewis derives from 5.2 parallels the fatalist's final conclusion. In arguing for 5.5, moreover, Lewis justifies premisses whose analogues (premisses 2 and 3) the fatalist takes for granted. Nevertheless, I think that the arguments have enough in common to allow us to use our grounds for rejecting the fatalist's as a guide to a possible answer to Lewis's.

Let us focus on premisses 2 and 3, which compare with Lewis's use of the equalities $P(A \rightarrow C/C) = 1$ and $P(A \rightarrow C/\sim C) = 0$. At this point both arguments rely on the procedure of assessing a conditional under an assumption concerning the truth of its consequent. I think we can get a good idea of what a factualist might object to about this procedure by constructing a parallel involving, in place of conditionals, a claim about probabilities themselves.

Let PROP be the proposition that $P(C) = 1$. The following argument claims to show that $P(\text{PROP}) = P(C)$. By 5.3,

$$(5.11) P(\text{PROP}) = P(C)P(\text{PROP}/C) + P(\sim C)P(\text{PROP}/\sim C).$$

Now by COND, $P(\text{PROP}/C)$ is the probability we would ascribe to PROP, if we came to believe that C. In these circumstances, we would ascribe probability 1 to C itself (this follows from COND and the fact that $P(C/C) = 1$). In other words (now assuming a factualist's view of credence), we would come to believe that $P(C) = 1$ – that is, that PROP. This doesn't yet tell us what *probability* we would ascribe to PROP; but applying COND once more, this time using the fact that $P(\text{PROP}/\text{PROP}) = 1$, gives us what we need: in these circumstances, $P(\text{PROP}) = 1$. Similarly, we can show that $P(\text{PROP}/\sim C) = 0$. By 5.11, this gives us $P(\text{PROP}) = P(C)$.

Somewhat less subtly, this time following the fatalist rather than Lewis, we can 'show' that PROP is materially equivalent to C itself:

- 1 $C \vee \sim C$.
- 2 Given that C, then there is no possibility that not-C, and so $P(C) = 1$ – i.e., PROP is true, as is C itself.
- 3 Given that not-C, on the other hand, it is certainly not certain that C, so PROP is false – again like C itself.

4 Hence, by Dilemma, PROP always has the same truth value as C.

It is clear that on any ordinary view of probability, both conclusions are absurd. The first rules out cases (easily found in physics, for example) in which the probability of an outcome necessarily lies in some range that excludes 1; while the second rules out all probabilities other than 1 and 0. But it is easy to see on what grounds a factualist will find these arguments fallacious. They assume, in effect, that PROP has a fixed meaning, such as to allow us to assess its truth value or probability in hypothetical contexts. Whereas in fact the sense of a probabilistic judgement depends very much on its context. The fallacy becomes apparent (the factualist will say) if we make explicit the normally implicit indexicality of such a judgement. If we think of PROP as

(5.12) Given the evidence to hand, $P(C) = 1$,

for example, then it becomes clear that in assessing PROP under the assumptions that C and that not C – in adding each assumption to the available evidence, in effect – we assess a different proposition in each case. We are interested in whether, given the evidence *actually* to hand, $P(C) = 1$; it is no help to be told that if *further* evidence were to hand (namely the fact that C), it would then be the case that $P(C) = 1$.

In the long run, we should clearly expect from the factualist a more precise account of the indexicality of probabilistic judgements. We shall return to this topic in chapter 7. For the moment the important thing is that from the factualist's point of view an appeal to implicit indexicality seems to work as well for conditionals as for probability. If we think of $A \rightarrow C$ on the lines of

(5.13) Given the evidence to hand, C follows from A,

then both Lewis's argument and the argument for fatalism commit the same indexical fallacy as the probabilistic arguments. They show that on the assumption of certain *further evidence* – the hypotheses that C and that not-C – A is irrelevant to C; but in general these facts are themselves irrelevant to the question at issue, which is whether *on the actual evidence* C follows from (or is probabilistically dependent on) A.

Against this indexical view of indicatives, it might be objected that it violates the presumption, as Lewis puts it, that 'our indicative conditional has a fixed interpretation, the same for speakers with different beliefs, and for one speaker before and after a change in his beliefs.' How else, Lewis asks, 'are disagreements about a conditional possible, or changes of mind?' (1976, p. 301). The thought is thus that if conditionals involve an implicit indexical reference to evidence, then as evidence varies between speakers, and for a given speaker over time, so will the content of conditionals. This seems contrary to the intuition that different speakers, or the same speaker at different times, can disagree about a conditional, without, as it were, arguing at cross purposes. For the moment, however, it is sufficient answer to this objection that speakers apparently cope with the same difficulties in the case of probabilistic judgements (and certainly do so with many other sorts of indexical expression). The important thing is that the two cases seem to be in the same boat.

Let us turn then to the second argument claimed to show that Adams's hypothesis cannot be treated as a special case of the general principle of assertibility for statements of fact: that of Carlstrom and Hill. Here we cannot do better than to reproduce Carlstrom and Hill's own presentation of the argument (with minor notational changes):

We will derive a contradiction from the assumption that truth condition *s* can be assigned to the indicative conditional $A \rightarrow C$, for arbitrarily chosen A and C, in some possibly complex way such that the probability that $A \rightarrow C$ is true is identical with the conditional probability $P(A \& C)/P(A)$. Let A and C be contingent and logically independent sentences, and let A, A&C, and $A \rightarrow C$ be true

at some possible world W_1 . The truth value of $A \rightarrow C$ at a world is not fixed by the truth values of A and C , for none of the sixteen equivalence classes of Boolean sentences constructible from A and C has a member with the property that its probability of truth is always identical with $P(A \& C)/P(A)$. Thus there are worlds W_2 and W_3 with the following properties: $A \rightarrow C$ is false at W_2 but true at W_3 , and each of A and $A \& C$ has the same truth value at both worlds. We can imagine two possible ways in which probabilities are assigned to the class of worlds. Under the first assignment, one of the two worlds W_1 and W_2 is probably the case, but each is about as likely to be the case as the other. In such a situation a sentence true at both worlds has a probability of nearly one of being true, a sentence true at just one of the worlds has a probability of about one half, and a sentence true at neither world has a probability of about zero. The second assignment is just like the first except that the worlds W_1 and W_3 are considered in place of W_1 and W_2 . Now since $A \rightarrow C$ is true at W_1 and false at W_2 , its probability of truth under the first assignment is about one half. Under the second assignment its probability is nearly one. But the probabilities of both A and $A \& C$ are about the same for both assignments, and in both cases, the probability of A is considerably larger than zero. Hence, the conditional probabilities are about the same. Although the probabilities that the conditional is true are quite different, the conditional probabilities are very close. Conditional probability is not probability of truth. (1978, pp. 156–7)

I want to show that as with Lewis's argument, the problem that this argument identifies is already a feature of a factalist interpretation of non-conditional credences. In this case the relevant point is so obvious that it may seem trivial. For example, let B be the proposition that the probability of a given proposition A is 0.9. Since $P(A) = 0.9$ does not imply that A , it is possible that B is true and A false. Let W be a world in which this is the case. Now imagine an assignment of probabilities to worlds that gives W a probability very close to one. In terms of this assignment it is true that $P(A)$ is close to zero, and yet also very probable that $P(A) = 0.9$. In other words, this assignment justifies the assertions both that B is false, and that B is very probably true.

The fallacy is fairly obvious: if we want to assign probabilities to possible worlds, and hence to think of the probability of a proposition as the sum of the probabilities of the worlds in which that proposition is true, then we must take it that judgements about probabilities are not true or false at individual worlds, but rather in virtue of facts about the space of worlds as a whole. Probabilistic judgements must thus be viewed as modal judgements.

Again, our concern is not with the adequacy of this form of factualism about probabilities, but rather with the point that the same kind of answer defeats Carlstrom and Hill's analogous objection to a factalist treatment of indicative conditionals based on Adams's hypothesis. I think the most striking way to illustrate this is simply to *define* ' $A \rightarrow C$ ' as ' $P(C/A) = 1$ '. It is true that in one sense this violates Adams's hypothesis – more on this in a moment – but it preserves something similar. We have already noted that it is a consequence of the factalist view of credence that for any proposition X , to believe (or be disposed to assent to) X is to believe (or be disposed to assent to) $P(X) = 1$. Let us express this consequence as

$$(5.14) \text{Bel}(X) \leftrightarrow \text{Bel}(P(X) = 1).$$

Given Adams's hypothesis, this means that if we accept 5.2, we are in any case committed to

$$(5.15) \text{Bel}(A \rightarrow C) \leftrightarrow \text{Bel}(P(C/A) = 1).$$

The present suggestion is in effect that we should regard 5.15 as a consequence not of Adams's hypothesis and 5.14, but simply of the fact that ' $A \rightarrow C$ ' and ' $P(C/A) = 1$ ' are synonymous. The consequence of 5.14 will rather be that

(5.16) $\text{Bel}(A \rightarrow C) \leftrightarrow \text{Bel}(P(P(C/A) = 1) = 1)$.

To the extent that it tells us why someone who believes that the probability of C given A is (effectively) 1 also believes that if A then C, this suggestion agrees with – indeed, *explains* – Adams’s hypothesis. But it will see it as a mistake to regard Adams’s hypothesis as an assertibility condition, in the above sense: the same kind of mistake as we would make in claiming that ‘X is a bachelor’ is governed by the rule that one should be prepared to assent to it just in case one believes that X is an unmarried male.

In the present case, the difference shows up in cases in which $P(C/A)$ is assigned a value other than 1. Suppose, for example, that we believe that $P(C/A) = 0.75$. On Adams’s hypothesis, this implies that we take ‘If A then C’ to be assertible to degree 0.75; whereas on the present suggestion, since $P(C/A) = 0.75$ implies that $P(C/A)$ is not one, ‘ $A \rightarrow C$ ’ will simply be believed to be false. As it happens, this suggestion thus makes better sense than Adams’s hypothesis of at least some uses of indicative conditionals. Consider this exchange, for example:

S: If his heart has stopped, then he is dead.

R: That’s not true – some people have survived cardiac arrests, so even if his heart has stopped, he may still be alive.

In such cases conditionals seem to be rejected outright, even if the corresponding conditional probability is quite high.

However, our present concern is not to defend this view of indicative conditionals, but simply to make the point that for a factualist about absolute credences, the Carlstrom and Hill argument, like Lewis’s, raises no new problem for a factualist view of conditionals. Carlstrom and Hill’s *reductio* depends on imposing on conditional truths a constraint the factualist is already bound to reject for probabilistic truths. This is really all we need here, though the above suggestion goes further. It proposes a factualist interpretation of indicative conditionals, automatically available to factualists about non-conditional credences, that makes sense of the grounds, if not the letter, of Adams’s hypothesis: the intuition that assent to an indicative conditional requires a high conditional credence.

I conclude that the arguments of Lewis and of Carlstrom and Hill do not prevent a factualist from endorsing Adams’s hypothesis (or something like it), and yet treating the resulting assertibility rule for indicative conditionals as a special case of the rule for statements in general. Earlier we construed these assertibility rules as fragments of a functionalist account of mind and linguistic behaviour. It seemed that indicative conditionals might be a case in which functionalism could provide the belief/non-belief distinction required by a psychologically grounded non-factualism. But it now appears that for conditionals, as in general, the functionalist is unable to justify a division between those functional differences that reflect differences of content, and those that reflect a difference in the type of attitude concerned.

It should be mentioned here that some factualists about conditionals – notably Lewis himself and Frank Jackson – have responded to Adams’s hypothesis in a very different way. Broadly, they accept that indicative conditionals employ a special rule of assertibility, but consider this to be an addition to the appropriate case of the rule for truth-conditional utterances in general. On this account indicative conditionals have truth conditions – those of the material conditional, in fact – and their rule of assertibility requires, *inter alia*, a belief in their (probable) truth. But for reasons to do with the conversational role of such utterances, it also requires that the corresponding conditional credence be sufficiently high. (Lewis 1976, Jackson 1987.)

The major difficulty for such an approach is to explain the significance of the claim that conditionals have truth conditions. It cannot be to account for ordinary practices of acceptance and rejection of

such conditionals; for to recognize a need for Adams's hypothesis is to recognize that the truth conditions of the material conditional do not explain ordinary usage. Moreover, we concluded in chapter 3 that if we are prepared to explain away usage, any meaningful utterance can be construed as truth-conditional. In a parallel fashion, any utterance can be construed as expressing a belief. But the construction does not guarantee that such a belief will play any non-eliminable role in a functional account of the use of such utterances. In a similar way, it seems to be the functional story associated with Adams's hypothesis – the special rule of assertibility – that does the work in a Lewis or Jackson account of the use of indicative conditionals. Lewis and Jackson accept that in virtue of Lewis's *reductio* argument (and perhaps also Carlstrom and Hill's) this part of the story cannot be interpreted in terms of the general account of statement of fact. The above argument suggests that this is a mistake. Provided that we are factualists about probability, and in particular interpret non-conditional credences as genuine beliefs, we can also be factualists about indicative conditionals.

What, incidentally, if we are not factualists about non-conditional credences? Are we then prevented by Adams's hypothesis and the Lewis and Carlstrom and Hill arguments from being factualists about conditional credences? It seems to me not. I mentioned earlier that in this case there are strong grounds for rejecting the usual definition of conditional probability in terms of non-conditional probability. A conditional credence cannot be thought of as the possession of a pair of absolute credences of appropriate content and degree, for in most cases in which we need a conditional credence (according to Adams's hypothesis), we simply do not have the relevant absolute credences.

In the (1986) paper in which I offered this argument, I suggested that we should therefore think of conditional credences as what I called 'inferential dispositions': dispositions to adopt a consequent credence of a certain degree, on acquiring an antecedent belief. This is not the only possibility. For example, Ramsey (1978, chapter 3) thought of both conditional and absolute credences as theoretical psychological entities, to be characterized in the first instance in terms of their effects on behaviour. Both can be measured in terms of choice of betting odds. The difference is that conditional credences call for conditional bets.

The various possibilities agree on one thing, however. The equality 5.1, normally used to define conditional probability, is no longer sacrosanct. At best, it is likely to be what it is for Ramsey: a derived result about the relationship between the absolute and conditional credences of a rational agent. The present significance of this new attitude to 5.1 lies in the fact that it suggests another answer to the arguments of Lewis and Carlstrom and Hill. In so far as these arguments rely on the usual definition of conditional probability, it is open to someone who rejects this definition – instead regarding conditional credences as an independent class of psychological entities – to object that Lewis and Carlstrom and Hill misuse what is at best a contingent connection between conditional and absolute credences.

The objection is actually more appropriate for the Carlstrom and Hill argument than for Lewis's. We saw that Lewis's argument can be expressed without reference to 5.1. In the Carlstrom and Hill case, however, 5.1 is crucial. In this case the claimed absurdity was just that it should be possible for different systems of belief to assign the same probability to A and $A \& C$, and yet different values to $P(A \rightarrow C)$ and hence to $P(C/A)$. But if 5.1 is at best dependent on an assumption of rationality, then there is nothing contradictory in the possible existence of such systems of belief. (Should irrationality be *impossible*?) There might be a contradiction if the Carlstrom and Hill argument showed us that rational systems of belief could differ in this way, but the argument offers nothing of this kind.

One might think – I speak here from experience, as my (1986) paper shows – that in taking a non-factualist view of absolute credences, and hence being led to view conditional credence as a primitive psychological category, one thereby adopts a non-factualist view of conditionals themselves. But on the contrary, it now appears that a primitive view of conditional credence in fact makes things easier for a factualist account of conditionals. Such a view can already help itself to our earlier responses to

the Lewis and Carlstrom and Hill arguments; and now, in virtue of its new attitude to 5.1, has another very promising objection to the latter argument.

If we characterize conditional credences as inferential dispositions, or something similar, then at first sight it seems that we assign them a different role from belief in a functionalist theory of the mind. Conditional credences become dispositions to add to our beliefs – to move from one belief to another. But at second sight we should appreciate that there is nothing in the functionalist theory to show that this functional role is not itself performed by a class of beliefs. It is the familiar problem: as psychologically inclined non-factualists, we must do more than show that a class of utterances are the expressions of mental states with a special functional role; we must exclude the possibility that the functional distinction in question marks a difference in attitude content, rather than a difference in attitude type.

PART II

Truth and Its Limits: An Explanatory View

6

The Function of Truth

For the benefit of readers who may have just joined us, let me summarize events so far. We have looked at a range of attempts to analyse the notion of statementhood – to say what it is about an utterance that qualifies it to be ranked in the fact-stating part of discourse. However, none of these attempts seemed able to exclude the possibility that all discourse is fact-stating discourse. None therefore meets the needs of either of the programmes we identified in chapter 1 – of non-factualists or marker theorists, who have a common interest in a non-trivial fact-stating/non-fact-stating distinction, even if conflicting views as to where it should be drawn.

One option at this point would be to walk away from the problem, concluding that non-factualists and marker theorists are simply misguided in expecting language to divide in this way. I think this would be premature, however. Before we blame the problem for the failure of the analytic solution, we should consider the possibility of other solutions. In the remainder of the book I want to argue for another approach. I want to show that so long as we concentrate on *explanation* rather than *analysis*, we can find considerable illumination in the concerns of non-factualism.

As I indicated in chapter 1, this explanatory approach will be founded on a novel theory of truth. The attractions of this theory are not limited to its ability to make sense of non-factualism, in my view, significant as I think this is. So, as a guide particularly to the reader whose interest is in the more general issues, let me indicate the structure of the remaining four chapters.

In this chapter I first characterize the explanatory approach, abstractly and by example; and then try to head off some objections to its application to a linguistic distinction of the kind we are interested in. I shall talk at this stage of the fact-stating/non-fact-stating distinction, but parallel remarks would apply to a distinction between the truth-bearing and non-truth-bearing parts of language. In any case, I then concentrate on truth, proposing an account of its function in discourse. The chapter concludes with some remarks on the relation between this account and other recent ‘explanatory’ approaches to truth.

Chapter 7 begins with a defence of the proposed account against a number of objections. From section 7.6 I turn to the account’s application to non-factualism. Chapter 8 continues this topic, mentioning also the non-indicative moods. Chapter 9, finally, returns to general issues. I discuss the connection between the proposed account of truth and Fregean general semantics, and relate it to other recent positions on truth, realism and objectivity.

6.1 EXPLANATION WITHOUT ANALYSIS

As theorists we often find ourselves with the task of explaining distinctions marked in the ordinary lives of ourselves and others. We notice that in cases of a certain kind, our subjects treat some instances in one way, others in another. Sometimes the behavioural difference will amount to application or non-application of a descriptive term. Sometimes it might consist in some other aspect of our linguistic behaviour. Or sometimes it might be entirely non-linguistic. Whichever it is, however, these behavioural distinctions call for explanation. Such patterns do not seem arbitrary – we want to know how they arise.

One sort of explanation tries to identify some feature of the two sorts of case, to which our subjects’ behaviour can be seen as responsive. This is the analytic or reductive approach. On this view, to put it crudely, their choice between behaviour X and behaviour Y is governed by their ability to recognize the presence of certain properties P_X and P_Y , and their conformity to the rule that they do X when

they detect P_X and Y when they detect P_Y . In the particular case in which the behavioural difference amounts to the application or non-application of a descriptive term, this amounts to the view that they use the term to refer to one of the properties in question; its use again being dependent on their ability to detect such a property in their environment.

An analytic account should thus be explanatory; the appeal to underlying properties should be capable of accounting for the behavioural difference with which we began. An adequate explanatory account need not be analytic, however. There are ways of accounting for behavioural distinctions that do not depend on an appeal to recognition (even implicit recognition) of underlying properties. This is perhaps most obvious in the case of non-linguistic behaviour – that of other species, as well as some of our own. Why do we tend to fall asleep every evening, for example? Recognition of the time of day obviously has less to do with it than has the physiological basis of our circadian rhythm.

In linguistic cases it is perhaps less obvious that the appropriate explanation may be non-analytic. All the same, some such accounts are quite familiar. They arise most naturally from a detached perspective: that of the anthropologist, say, or the historian of science. For example, we would not expect a fully recognitional account of the origins of a system of religious or defunct scientific distinctions; for we want to say that there may be nothing in the world that matches the distinctions concerned.

Some terminology: from now on I shall use ‘explanatory account’ as an abbreviation for ‘non-analytic explanatory account’. This is not only shorter, but gets the emphasis right. Explanation comes second in an analytic account. Explanatory approaches put explanation first.

I want to suggest an explanatory approach to two linguistic distinctions that are very far from defunct; two distinctions marked in our ordinary use of the notions of truth and falsity. One of these is the distinction whose analytic treatment part I was concerned to discredit – that between statements and non-statements, or truth-bearing and non-truth-bearing uses of language. The other is the distinction we mark within the truth-bearing part of language, in applying the descriptions ‘true’ and ‘false’ selectively. I want to explain the former distinction in terms of an account of the latter. In other words, I want to show that we can explain the limits of truth in language in terms of an account of its function in central cases.

In both cases the explanation is intended to be non-analytic. I want to show that our use of truth and its limits can be explained without being taken to reflect or be responsive to any such real properties of linguistic items. Because these distinctions are so much a part of our own practice, however, the detached explanatory perspective does not come easily. To soften the target a little, I want to begin with an example in which there seems to have been an analogous shift from an analytic to an explanatory approach to a perceived distinction in human life. It is the case of social class.

Social stratification is a striking feature of most societies. It is not simply a theoretical notion, important only to anthropologists and sociologists, but part of the lives of ordinary people. Many aspects of a person’s life depend on his or her caste, or class, as it is perceived by the rest of the community in which he or she lives. Ordinary folk apply such notions to themselves, and the theorist begins with the notions of folk sociology.

Folk sociology takes folk usage at face value. After all, what is folk sociology except folk usage about society? Folk usage applies class properties to individuals. Hence it gives the impression that the way to a theory of class is to *analyse* these properties – to discover, for example, the distinguishing characteristics of the aristocrat and the commoner, the prince and the untouchable.

The simplest such theory perhaps takes the analysis no further than the folk class properties themselves. Class is thus itself regarded as part of the natural order of the world. An individual is seen

to belong to a class in much the way that an animal or plant belongs to a species. We now find such a view absurd and repugnant. We should not underestimate its plausibility, however, particularly in societies without significant class mobility, and in combination with a suitable account of the origins of such an order.

A more sophisticated approach, less clearly repudiated in contemporary thought, is to analyse class distinctions in terms of more basic properties of individuals. So long as these properties are taken to be changeable, this approach allows for class mobility. Unlike the first, it can also distinguish between real and apparent class, and hence allow for the class exiles of fiction and fact – lumberjack lords, Downing Street grocers, and the like. Moreover, it promises an explanation of some of the social manifestations of class. Indeed, it may seem to offer a justification of social stratification – the idea that because the rich are better at something, they not only get but deserves the gravy.

Attractive as they have often seemed, such accounts of class founder on one of three main fallacies. For a start they may appeal to underlying properties that although genuinely individual, are not unevenly distributed between the social classes. Appeals to biological properties of individuals are likely to fail for this reason, for example. Unlike carthorses, the working classes are not innately adapted to the labouring life.

The second kind of fallacy is to call on properties that although both individual and appropriately distributed, bear the wrong explanatory relationship to the relevant social distinctions. This can happen in several ways. The properties concerned may be effects rather than causes of the class distinctions they are supposed to account for – think of the correlation between access to education and social class, for example. Alternatively, their correlation might stem from a common cause, or simply be accidental.

The third fallacy is perhaps the most pervasive, however. It is to appeal to properties that are themselves constituted in social terms. It would be involved, for example, in an attempt to explain class distinctions in terms of differential ownership of capital or possession of political power. Property and power are themselves in some sense social constructs. We may well clarify the problem of class by describing it in these terms – come to understand that some social distinctions are of more significance than others. But to clarify a distinction in this way is not to explain how it comes to exist in the first place.

In rejecting such accounts of class, we recognize, I think, that an adequate theory of class would not be simply an analysis of the class properties of folk sociology. We distinguish the task of describing the class structure of a given society from that of explaining the existence of such distinctions. As noted, the descriptive task may have an analytic component. But we would not expect the explanation of social stratification to appeal to non-accidental differences between individuals.

What the explanation of class should appeal to is a matter for some debate. By way of illustration we might mention two approaches, both in some sense evolutionary. On the first, class would be viewed as a product of notions of kinship – of their tendency over the generations to accentuate, rigidify and defend what were originally minor and accidental variations in social role. Classes would condense like galaxies from clouds of dust; or perhaps more accurately, like organisms from the primordial soup. On the second view, in contrast, a social class would be compared more to an adaptive variation than to an organism or species in its own right. Class distinctions would thus be explained in terms of their functional roles, and hence their contributions to the reproductive success of a community as a whole.

Whatever the appropriate story, however, the class case makes two important points. It shows that the use of a distinction that seems initially to reflect intrinsic differences between individuals may turn out to require a non-analytic style of explanation. And it illustrates that in the grip of the analytic perspective, we are liable to take as explanations what are at best redescrptions of the distinctions we

want to account for. It seems to me that the sorts of views we considered in part I make the same kind of mistake about the fact-stating/non-fact-stating distinction. Naïve theories of class attempt to explain class distinctions in terms of intrinsic differences between individuals. Analogously, these linguistic views try to account for the distinctive features of fact-stating discourse in terms of some intrinsic property, supposedly distinctive of fact-stating uses of language. As in the case of class, the approach fails for two main reasons. Either the claimed underlying property is not appropriately distributed (or cannot be shown to be so distributed, independently of the distinction it is supposed to explain). Or the supposed intrinsic property is itself a manifestation of the difference of usage that calls for explanation.

However, I think it must be conceded that although the class analogy might suggest this possibility, it does little to support it. It is clear, I take it, that the proper approach to class is explanatory; but far from obvious that the linguistic distinction might even admit the same kind of treatment. This difference seems to stem from an important disanalogy between the two cases. We incline to the lottery view of class selection: whatever the origins or function of social stratification, an individual's class is decided, in the main, by the accident of birth. The lottery view is much less plausible in the linguistic case, however. The 'individuals' differentiated by the linguistic distinction are speech acts, linguistic tasks, or something of the kind. They are not biological individuals, whose interests we know to be often subordinate, from a biological point of view, to those of a species or community. They compare instead to enduring characteristics of a species. We want to explain some striking differentiation of the tasks for which we employ language – the distinctions marked in English by the use of the indicative mood, or the application of the notions of truth and falsity, for example. As a result, the linguistic distinction seems more tied to an explanatory base than is true of the class case.

It is rather as if we had observed that a community uses two quite different methods of performing the task of moving themselves from place to place. Sometimes they 'go on foot', at other times they 'take the Tube'. The community may have little sense that these activities have something in common. They may not realize that in principle anywhere that can be reached by taking the Tube could also be reached on foot. As theorists, however, we would expect an explanation of this behaviour not simply to account for the fact that the movement task is differentiated in this way, but to explain why particular movement tasks tend to be performed in one way rather than the other. We note for example that from Highgate to High Barnet, almost all our subjects 'take the Tube'. We would be surprised if this were simply an accident – if the habit were not explicable in terms of certain features of that particular movement task (such as the distance from Highgate to High Barnet, and the availability of the Northern Line).

In the linguistic case, similarly, we expect that as an enduring feature of human behaviour the use of the statement form will turn out to be well adapted to the needs of particular linguistic tasks – to the sociolinguistic function it has presumably developed to serve. I think this explains why the analytic or reductive approach is less clearly inappropriate here than in the case of class. If an account of fact-stating discourse needs to appeal to some pre-existing difference between linguistic tasks, in order to explain the selective application of the features of usage characteristic of statements, then this difference seems itself to form the basis of a reductive account. For can we not say that statements are those utterances that are used to perform tasks having the underlying property that explains the assertoric form of usage?

I think there are three things that may seem to recommend this suggestion. The first is the lingering appeal of the analytic approaches discussed in part I. In the face of the evidence, our intuitions doggedly maintain that we can explain the assertoric form in terms of its use in expressing beliefs, carrying truth values or representing states of affairs. Putting this aside, two factors remain: first that it is difficult to see how we could explain the development of the assertoric form, except as a manifestation of the kind of base property that would enable the suggestion to go through; and secondly that an appealing account of the use of language seems to require that speakers be aware of some such base property, in order to explain their choice of particular forms of words.

In the next section I shall try to cure any lingering attachment to intuitions of this kind. Readers not so afflicted may move directly to section 6.3. But as an indication of the kind of resolution these factors allow, consider again our anthropological subjects who ‘go on foot’ or ‘take the Tube’. I said that although these people might have little conception that these were different ways of doing the same thing, we would expect to find explanatory differences between the ‘movement tasks’ for which they typically choose one mode rather than the other. But we would recognize, I think, that the correlation between these underlying properties and our subject’s choice of transport might be far from perfect. Being creatures of convention and habit, they might sometimes take the Tube when (as we can see) it would be better for them to walk, and *vice versa*. These correlation failures might be either type-type, as when it would generally be better to make the journey in question by the other means; or type-token, as when the usual advantage of the chosen mode is outweighed by special circumstances (frozen points at Camden Town, say). From the explanatory point of view these failures are of little significance. An evolutionary account relies on long run advantages, and tolerates an element of historical accident. But clearly they defeat any simple reduction of the acts of taking the Tube and going on foot to the underlying properties that explain their differentiation in nature. And the same may be true in the linguistic case. Whatever the usual function of fact-stating or truth-bearing discourse, we should not be surprised to find cases (tokens and types) in which this function goes unserved. This possibility prevents any simple analysis of assertion in terms of the underlying properties being held to account for its form and existence in language.

*6.2 ADAPTION AND LINGUISTIC CHOICE

We want to explore the idea that the statement form has developed to serve some identifiable sociolinguistic function. We want to resist, however, the resulting pressure to *characterize* statements as speech acts that serve this function. I mentioned two intuitions that might seem to favour this move: first, that it would allow us to explain the development of the assertoric form as a manifestation of the relevant functional property; and secondly that an adequate account of the use of language might seem to require that speakers be aware of some such property, in order to explain their choice of particular forms of words.

Concerning the differentiation of language into assertoric and non-assertoric forms, there seem to be two possible accounts. One would allow that the development of the assertoric form has been at least partly guided by speakers’ own intuitions about the appropriate form of their language, in the light of its functions. The other would insist that such development is essentially ‘blind’, not significantly influenced by speakers’ understanding.

The latter account seems bound to resort to an evolutionary story, explaining the distinctive characteristics of assertoric and non-assertoric discourse as the result of a gradual process of adaption and selection. The same will be true of the former account, to whatever extent it relies on ‘blind’ development. In neither case, of course, need ‘evolutionary’ imply a genetic basis for the variation and retention of linguistic characteristics. Nor is it clear that linguistic selection need depend on reproductive success. In humans as in certain other animals, new patterns of behaviour can spread through a population in very much less than a generation; and hence very much faster than could be explained by their contribution to fertility.

All the same, the likely evolutionary character of the development of language does a lot to undermine the intuitions that seemed to favour the analytic approach. For one thing, an evolutionary theory is capable of explaining the development of a behavioural distinction, without treating it as the manifestation of some pre-existing distinction. This is one of the great achievements of Darwinian theory: the fact that it can account for functional differentiation in nature, without taking the development of species and other biological kinds to consist, as it were, in the teleological colonization of previously vacant roles. On the contrary, the process of development creates the roles as it goes along. Hence the vacuity of the doctrine, sometimes advanced, that evolution ‘fills every

available niche'. Within certain physical constraints, evolution creates the niches as it creates their occupants.

Take the case closest to hand. Darwinism frees us of the idea that the human hand is a directed solution to a pre-existing functional problem – an idea embodied not only in 'designer' theories of creation, but also in the Lamarckian view that adaptation occurs as organisms strive to reach certain goals. After Darwin we can say in all seriousness that the functional role of the hand – its biological niche – is the performance of manual tasks. This truth is as trivial as it seems. Only a teleological view requires that the functions of the hand have anything more in common than the fact that they are the functions of the hand. For a Darwinian the interesting truths lie elsewhere: in the story concerning the development of the hand from its precursors, and the contributions of its developing functions to survival and success. And so much, I think, for the feeling that an analytic account is required to explain the development of the assertoric form of usage. It rests on a failure to appreciate the potential of non-teleological accounts of the development of biological characteristics.

The second intuition also rests on a teleological viewpoint, but at a different level. Whereas the first intuition concerned the development of forms of usage themselves, the second has to do with their application on individual occasions of use. It stems from the feeling that any significant feature of an utterance should in principle be explicable in terms, in part, of the speaker's *understanding* of the relevant part of the language. We think of the choice of a particular form of words as an intentional action, and therefore as something that ought to be explicable, in part, in terms of the speaker's beliefs about the significance of those words. To accept otherwise would seem to be to consign the relevant feature of linguistic competence to the ranks of the instinctive, the habitual and the 'wooden' (Blackburn 1984, p. 96).

I think we are thus inclined to think of the choice of the assertoric form as something like the choice between synonyms – what goes on when a competent speaker chooses to say 'Tube' in preference to 'Underground', 'Metro', or 'Subway', for example. Such a choice is presumably based on a prior understanding of the nature of the speech act in which the chosen word will play a part. It is difficult to shake off the idea that the 'choice' of the indicative mood and other features of the assertoric form is similarly based on an implicit grasp of an underlying property of an intended communicative act. Since such an underlying property would seem to provide the basis for an analysis of the notion of statement of fact, this 'choice' model of utterance lends credence to the analytic approach.

Appealing as the choice model may seem, it cannot be applied without restriction. Unlike Humpty Dumpty, we find our choice of words constrained by our intended meaning. Of course, these constraints usually allow us a good deal of freedom. Thanks to the existence of synonyms, there are usually many ways of 'saying the same thing'. But the fact that there are such constraints means that we may in fact have no choice about the assertoric form, and hence no need to be aware of an underlying property in virtue of which that form is appropriate.

It might be objected that even where language gives us no choice, we have in some sense to know our intended meaning, in order to be in a position to determine the correct words for the job. If so, then linguistic intentions are once again prior to the selection of the assertoric form. But there are big problems for such an account. In particular, it seems to depend on the existence of judgements that are not linguistic in form. For what does a speaker conclude, when he judges that the sentence 'P' is the appropriate way to express what he means to say? That 'P' expresses the proposition that P? This judgement is not tautologous, as of course we discover if we translate it into a language other than English. But it is not the sort of thing that a competent English speaker could be in any doubt about. So to the extent that reasoning is conducted in English, this judgement cannot be a significant prerequisite for the decision to say 'P'.

What if we move to another language – perhaps the language of thought? Then we do have a non-trivial judgement, of the kind of thing that a would-be English speaker would have to know. Let us

suppose that in the language of thought, this judgement is expressed by 'Q'. In other words, we are supposing

R: In the language of thought, 'Q' expresses the proposition that in English, 'P' expresses the proposition that P.

The suggestion is that in order to account for our speaker's utterance of 'P', we have to admit that he judged that 'P' is an appropriate way of saying that P. Since this judgement is trivial in English, we are supposing that he expresses it in another language – the language of thought, for example. (Interposing other languages at this point would only delay the inevitable.) Presumably he makes this judgement by tokening 'Q' in an appropriate 'judgement register'. In effect, he says 'Q' to himself. But why 'Q'? Is *this* choice guided by the judgement that R? If so, then the problem recurs again. If not, then why is there a problem in the first place?

Thus at some point this view seems committed to the existence of non-linguistic judgements. This takes the view perilously close to the instinctive, habitual and wooden. The mark of 'non-woodenness' seems to be our ability to reflect on our own beliefs and practices. But how can we reflect on what we cannot express? In other words, it looks as if the preference for the choice model of the use of the assertoric form (and hence for an analytic account of statement of fact) rests on an unattainable ideal. In the use of language, as elsewhere, our freedom of choice cannot be boundless. Our choices are always constrained by the terms in which the options present themselves. We cannot choose what we cannot conceptualize. The above argument suggests that in the case of 'choosing one's words', the options present themselves in linguistic form. Choosing one's words is not a matter of first settling on the particular (fully specified) speech act one wants to perform, and secondly selecting the appropriate words for the job. There is only one kind of choice; though it may go in several steps, as one refines one's intentions, in the process of settling on a particular form of words.

There is a nice parallel here with accounts of the origin of life. Like teleological misinterpretations of Darwin, creation myths tend to embody an unwarranted conceptual distinction between a biological kind and its 'place' or niche in nature. For example, the biblical account may leave us with the impression that when God created a species, He made two decisions: to include a particular niche in the Great Plan, and to create the creature that fills that niche. These decisions might be taken in either order, and we find suggestions of both orderings in the biblical account. The more obviously absurd is that of creation before assignation to a biological role. This is suggested, for example, by God's famous injunction – to creatures already equipped, presumably, with their reproductive instincts and apparatus – to 'Go forth and multiply.' The reverse ordering is suggested, for example, by the account of the creation of Eve. We are given the impression that God first decides to provide Adam with a mate, and then decides to create Woman for that purpose. But there is room for two decisions here only to the extent that the first decision is incomplete. God might have initially decided that Adam should have a mate without making up His mind on some of the details. (Should Eve eat Adam after mating, for example?) This leaves room for a series of decisions. However, to the extent that God has filled in such details on the Great Plan, the actual creation of Woman is not a separate decision but merely the fulfilment of a prior intention. A place in the Great Plan is delineated by the properties of its occupant. To decide to fill the place is thus to decide to create the occupant.

Analogously, the above view of linguistic competence involves what amounts to a two choice account of the 'creation' of an utterance: the idea that in deciding to speak, we first decide what kind of linguistic act we wish to produce – what niche we want to fill – and secondly choose the appropriate sentence for the job. It seems to me that this model involves the same sort of misconception as the two-choice view of the creation of life. Just as the common feature of the hand's activities is simply that they are *manual*, so what unifies the speech acts for which we use the statement form is just that they are utterances in the statement form. We cannot choose the statement form on the basis that we recognize that an intended speech act belongs to a certain kind, because to get that far is already to have made a decision about the use of the statement form.

Given two sentences of whatever form, we can of course choose to utter one, both, or neither. In doing so we rely, presumably, on some conception of the effect of an utterance of either kind. What we must resist, I think, is the idea that this understanding of the role of a potential utterance is somehow independent of our awareness of the linguistic expressions whose use is at issue. In the end, we must live with the fact that linguistic acts present themselves to us as *linguistic* acts, and not in terms of some pre-linguistic characterization.

All the same, in this and the next two chapters it will sometimes be convenient to speak as if we took it that language had been developed to serve a range of pre-defined functions; in particular, to express a range of different sorts of mental states – beliefs and suchlike. As is often the case in biology, a teleological perspective will thus be a useful expository device. This psychological scaffolding will be thoroughly dispensable, however. The work will be done by the notion of behavioural dispositions, and we won't need to assume that such dispositions serve to delineate the familiar psychological categories – in advance, as it were, of the linguistic distinctions we shall use them to explain.

6.3 TRUTH: THE EXPLANATORY PERSPECTIVE

In chapter 2 I mentioned Strawson's (1949) 'amen' account of truth, which makes a theory out of the observation that one function of the term 'true' is to provide a uniform means of endorsing a statement made by another speaker. Strawson argues that the description 'true' is primarily a device to confirm or corroborate a previous utterance. He doesn't say what 'false' is, but there are at least two possibilities: that it 'confirms' the negation of the original utterance; or, without assuming an account of negation, that it indicates some sort of rejection or 'disconfirmation'.

Strawson's theory is not an analysis of truth. It is rather an account of the function of the term 'true' in discourse – an account that does not presuppose that the function of such a term must be primarily referential. Strawson does not tell us what truth is, but offers a theory as to why we should talk as if there were such a property. In our terms it is thus an explanatory theory of truth.

It is not a very satisfactory theory, however. For we noted that if corroboration were the main function of 'true' it would be unclear why the term should not be applied more widely – to questions and commands, for example. It would often be useful to have a simple device for putting one's weight behind a question asked by someone else: for indicating, in effect, that one is asking the same question oneself. Similarly, there are cases in which it would be useful to be able to say in a simple way that a question did not have one's endorsement. The same is true of commands, requests, indeed of practically any sort of utterance. This amen theory of truth does not explain why we should be allowed the conversational convenience in some cases but not in others. So if a functional theory of truth is to account for the non-indicative moods, the function of 'That's true' had better be more than that of 'Amen'.

A more promising approach lies in the observation that to call an utterance 'true' is to say that it is *correct*, or *right*; to call it 'false' is to say that it is *incorrect*, or *wrong*. As Dummett puts it, 'the roots of the notions of truth and falsity lie in the distinction between a speaker's being, objectively, *right* or *wrong* in what he says when he makes an assertion' (1978, p. xvii). We might be wary of Dummett's use of the notions of assertion and objectivity – these are the sorts of notions we want to explain – but the emphasis on the *normativity* of judgements of truth and falsity is a step in the right direction.

Whatever their origins, normative terms seem to function as verbal carrots and sticks, with which we can reward and punish certain sorts of behaviour. As with more tangible incentives, such rewards may be actual (for past behaviour) or merely conditional (on future behaviour). But unlike that of carrots and sticks, the reward or punishment bestowed in a normative judgement is in a sense self-inflicted. This not because we ourselves wield these verbal implements – that is true of carrots and sticks – but because their effect is conditioned and conventional. A child who has not learnt the meaning of 'good' and 'bad' cannot be encouraged to modify his behaviour by being told that it is good to eat his

breakfast and bad to jump on the cat's tail (which is why parents are often tempted to resort to less convention-governed incentives).

Non-verbal incentives can also depend on convention, of course. A child who has not learnt the relevant conventions will not be moved by the offer of a fiver if he eats his breakfast. And just as we might be taught to be 'bad', or to regard 'Everything you say is false' as a pat on the back, so we might learn to attach monetary value to almost any kind of physical object.

In general, at any rate, the use of such conventional rewards and punishments presumably serves some point. Why else should it have developed? In the case that interests us, we might hope and expect that the point of use of the normative terms 'true' and 'false' is such as to explain the ordinary limits of their application. We might hope to show, in other words, that the application of these terms to assertions serves a purpose that would not be served by applying them to, for example, questions and commands.

Again, we would thus be explaining 'true' and 'false' not analytically, but simply in terms of their functional role. We would be asking not what these terms refer to, but rather what they do for us (and hence why we might have developed them). This approach would be uncontroversial for a notion such as monetary value (or the valuations inherent in class distinctions, to return to our earlier example). However, it is apt to seem misguided in the case of normative notions such as truth, moral value and rationality. In the monetary case, we recognize that monetary value is conventional not simply as regards the objects we choose to attach it to, but also in the sense that the very notion of monetary value is clearly a social product (and hence conventional). But we are reluctant to allow that normative terms involve both these kinds of conventionality. The first kind is obvious enough: it is simply the point that we could have meant by 'bad' what we presently mean by 'good'. But we are inclined to assimilate this with the point that, for example, we could have meant by 'kipper' what we actually mean by 'sardine'. That is, we resist the suggestion that what we would be talking about in using these different words is also a social construct, rather than a natural feature of the world.

Accordingly, I think, we consign ourselves to some of the great treadmills of Western philosophy – the attempts to analyse the concepts of Truth, Reason, and the rest, and hence to understand their normativity and the role they play in our lives. These well-worn steps continue to seem more attractive than the conventionalist alternatives. A striking indication of the depth of this preference lies in the tendency on the part of philosophers who abandon the attempt to analyse truth in terms of correspondence to reality, to seek to analyse it in some other terms. We thus have the pragmatist tradition of coherence accounts of truth, of attempts to identify 'true' with 'useful', 'accepted at the end of inquiry', and the like. True, in the pragmatist tradition we also find the view that philosophy should not attempt to analyse truth. But even here there seems to be reluctance to take on the task of *explaining* our use of the notion of truth. Instead such pragmatists 'would simply like to change the subject' – to 'suggest that we do not ask questions about the nature of Truth' (Rorty 1982, p. *xiv*). This is the quietist tendency I referred to earlier, to blame the problem for the failure of the analytic solution. We shall return to it in chapter 9; but it is further evidence of the near invisibility of the explanatory alternative in contemporary thought.

Perhaps we suspect that in taking the explanatory approach, the conventionalist buys the easy life at the cost of what gives point to life, namely the very values in question. In other words, we fear that if we allow ourselves the explanatory perspective, allow ourselves to ask why we talk in these terms, we lose our entitlement to use them, to continue to behave as if they matter. We can cheerfully regard money as a social convention, because we could countenance getting by without it. But how could we be philosophers, or indeed persons, without a sense of the reality of Truth, Reason and the Good?

At best, this could amount to an argument for not attempting to develop explanatory account of normative terms – a version of the argument that there are some questions it is safer not to ask. Even so, I think it exaggerates the risk. In other areas of life, we seem to be able to combine an awareness

of the functional role of appetites and values with an undiminished ability to feel their force. Abstinence does not seem to be assisted by the knowledge that one's craving has a biological basis. It might be objected that there is a difference between these cases and those involving the use of normative terms such as 'true' and 'good': namely, that since the latter depend on linguistic convention, it is easier for us to give them up, or to cease to feel their force. But if the use of the word 'convention' is thought to carry this implication, we can do without it. The explanatory approach need not claim that with the benefit of an explanation of our use of these normative terms, we could (let alone should) do anything to change the practice. Some linguistic practices may perhaps be accessible to revision. But many, perhaps most, may simply be wired into us, in circuitry over which we have no control.

6.4 THE CASE FOR SEMANTIC INCENTIVES

We have digressed, however. The suggestion was that we regard the normative notions of truth and falsity as the core of a conventional system of punishment and reward, and hence try to explain the limits of their application. In other words, we should try to show that the application of these terms to assertions serves a purpose that would not be served by applying them for example to questions and commands.

What purpose could this be? I suggest that the normative character of the notions of truth and falsity provides speakers with an incentive to justify their utterances to others, and hence to participate in an important form of argument or dispute. At its crudest level, it ensures that speakers are rewarded when their utterances are endorsed by others, and punished when their utterances are rejected. Thus it gives everyone an interest in participating in and attempting to 'win' the relevant disputes. It gives us all an interest in persuading others to endorse our own utterances (and hence, of course, in making sure that our own utterances are such as to be endorsed by others).

The normative nature of the terms in which such disputes are conducted gives them short term value to individuals, but cannot constitute their long term value to the linguistic community as a whole. On the contrary, the development of such a system of linguistic incentives suggests an underlying behavioural advantage – something that would explain the linguistic development in evolutionary terms. This advantage presumably turns on the behavioural consequences of the mental states exposed to criticism in such linguistic disputes. In particular it might lie in the survival advantages of basing one's mental attitudes on as wide a body of experience as possible.

Rough as it is, the approach already suggests an explanation for the fact which embarrassed Strawson's theory – namely that 'true' and 'false' can be used to endorse and reject only certain sorts of utterances. Perhaps not all utterances serve a function such that there is reason to seek or expect agreement between speakers. Consider requests, for example. Imagine that the two of us are talking to a third person, Fred. I ask Fred to buy us each a cup of coffee. Fred looks at you, but you say nothing. Our utterances on the subject conflict, in the sense that no one person could have both asked and not asked Fred to buy a cup of coffee. But clearly this 'conflict' may be entirely benign. There may be no reason to expect us to make the same requests, of Fred, or in general – and hence no advantage (quite the contrary) in a linguistic device that would encourage us to do so.

The proposal is thus that some utterances normally express states of mind with respect to which there is survival value in encouraging speakers to dispute, in the above sense. For these states of mind, two heads are better than one, and hence there is a general advantage in exhibiting differences, and in providing incentives to encourage the spread of well-justified beliefs. The use of the normative terms 'true' and 'false' is an important part of the apparatus that language has developed to exploit this advantage. However, the apparatus is not deployed where it is generally not appropriate. The roles of utterances such as questions, commands and requests are not such as to provide an advantage in seeking unanimity. On the contrary, these utterances are such that different speakers may reasonably conflict (in the above sense), even if fully acquainted with each other's point of view. So ordinary

usage does not provide the incentives for disputes in such cases – perhaps marking their immunity with the non-indicative moods.

We thus have the beginnings of a functional account of truth, grounded on the claim that its primary role is to encourage and facilitate a useful form of linguistic dispute. The present sketch leaves a lot to be desired, of course. How, for example, is the claimed difference between indicatives and non-indicatives to be reconciled with the results of chapter 3 – with the apparent possibility of treating non-indicatives as disguised assertions? I shall return to the non-indicatives, and to this query in particular, in chapter 8.

It may seem also that the proposed account leaves no room for non-factualism. An interesting non-factualist thesis is bound to concern a case in which utterances are ordinarily treated as true and false. The non-factualist claims that although such utterances appear to be statements of fact, they are really something else. But the non-analytic approach leaves no room for a distinction between appearance and reality in this case – no room for categories underlying use in terms of which the non-factualist's claim could be expressed. We shall see, however, that the account leaves room for a distinction of a different kind. In particular, the complexities of the relevant form of linguistic dispute give rise to a range of cases in which the optimum application of the normative notions of truth and falsity is far from straightforward. The ordinary linguistic apparatus is stretched to its limits, and the strain shows up as doubts as to whether the utterances in question 'have truth values' – whether, in other words, these notions are applicable at all. Not surprisingly, these are just the cases in which non-factualism has seemed attractive. Far from excluding non-factualism, this approach thus does a fair job of explaining its continued appeal.

It is true that the approach entails that traditional non-factualism is misconceived, in taking the analytic stance – in claiming to discern a categorical distinction underlying use. But in place of this misconception, the account offers two distinct sorts of question: the first as to whether the use of a given class of utterances is fully assertoric, or rather involves the identified departures from the canonical use of truth and falsity; and the second as to whether the actual use is the most appropriate, in view of the account's understanding of the function of the notions of truth and falsity. Traditional non-factualist arguments can be expected to bear on both questions. Indeed, in the next two chapters, as we examine the structure of linguistic disputes, we'll repeatedly appeal to the sorts of considerations that have been taken to support non-factualist treatments of particular topics.

In the next chapter our first job will be to show that the relevant notions of dispute and disagreement do not themselves depend on the notions of truth and falsity. If we want to explain the notion of truth in terms of its role in fostering certain activities, our characterization of these activities cannot depend on truth. Before that, however, I want to distinguish this approach to truth from another approach, also described as 'explanatory', which has received attention in recent years.

*6.5 TRUTH: ANOTHER 'EXPLANATORY' PROGRAMME

We have come to the task of explaining truth through that of accounting for the distinction between fact-stating and non-fact-stating discourse. We are interested in explaining rather than analysing truth because we have concluded that unless an analysis of truth draws on a prior understanding of the nature of statement, then, like Strawson's amen theory and the disquotational accounts of truth, it will be unable to explain the fact that not all utterances are treated as truth-valued. In explaining the existence of a more than disquotational (and hence non-universal) notion of truth, we hope to explain, in effect, the existence of the non-universal linguistic category of statement of fact.

In recent years several writers have noted that although the ordinary notion of truth (or at least the philosopher's ordinary notion of truth) seems stronger than a disquotational notion, it is hard to see why we need a stronger notion. The disquotational notion (or something of similar 'strength', such as a Tarskian truth theory) seems to be all that we need for a theory of linguistic understanding. And in

psychological theory, it seems to be the ‘internal’ cognitive role of an attitude, rather than any ‘external’ relation of correspondence with the world, that does the explanatory work. The recent appreciation of these points, and of their significance for an account of truth, is particularly due to Hartry Field (1972, 1978).

Field asks ‘why we need to be able to apply semantic notions like truth and reference to the sentences that people believe and desire’ (1978, p. 48). This question seems to me to involve very much the same explanatory perspective as the approach suggested above. It sees the task as that of explaining the ordinary use of the notion of truth. (I think Field’s answer is inadequate, however – more on this in a moment.) But in the discussion to which Field’s work has given rise, the phrase ‘explanatory theory of truth’ has come to mean something different. The emphasis has shifted from explanation *of* (the use of) the notion of truth, to explanation *by means of* the notion of truth.

This shift is associated with a slightly different starting point: a philosophical commitment to a ‘realist’ world view, and hence an interest in justifying, rather than merely explaining, the use of the kind of notion of truth that such a view seems to require. This approach to truth is particularly due to Putnam (1978), who suggests that a realist notion of truth may be required to account for the fact that linguistic behaviour contributes to our overall behavioural success: ‘The *success* of the “language using program” may well depend on the existence of a suitable correspondence between the words of language and things, and between the sentences of language and states of affairs’ (1978, p. 100).

Putnam’s approach makes truth a theoretical notion, an element in the kind of story that a theorist would have to tell about language and its place and nature. Hence there is at least a danger that this theoretical notion will have nothing to do with the ordinary notion of truth. In what sense is it the same notion? And even if it is the same notion, why do ordinary speakers need it? These questions emphasize the difference between an explanatory theory of truth in Putnam’s sense, and a theory that seeks to explain the ordinary use of ‘true’ and ‘false’. In principle a theory could be explanatory in both senses. Truth might be a necessary part of an essential folk theory of language and the world. (Devitt 1984, chapter 6, seems to take this view, though not on Putnam’s lines.) But the two approaches are independent. Putnam’s does not guarantee ours, and ours does not depend on Putnam’s – folk notions may be explicable without being explanatory.

All the same, we can distill several proposals for an explanatory theory of truth in our sense from the work of Putnam, Field, and others. If any of these succeeds in explaining the fact that the ordinary notion of truth is not merely disquotational, then presumably it will explain (where the disquotational theory could not) the fact that truth is restricted to a particular sort of discourse. Our own suggestion would thus be unnecessary, or perhaps repetitious.

Putnam’s own suggestion has been criticized (McGinn 1982, p. 222; Devitt 1984, p. 90) on the grounds that the explanation of the success of linguistic behaviour does not require a realist or correspondence notion of truth; a notion of warranted assertibility will do the same job. As Devitt puts it:

The role claimed for truth in [Putnam’s] explanation must be ... as follows: the expectations that a belief gives rise to tend not to be disappointed if the belief is true but tend to be disappointed if it is false. True beliefs tend not to face recalcitrant experience. The problem is that warranted or justified beliefs tend not to either. (1984, p. 90)

It seems to me that there is a more serious objection, however. Not only does the explanation of success not need a correspondence notion of truth; all it needs is a disquotational notion. For the central claim, in effect, is that beliefs tend to lead to success if and only if they are true. And this is just the sort of generalization that has long been recognized to provide gainful employment for a disquotational notion of truth. The classic example is Ramsey’s: ‘He is always right’, or ‘Everything he says is true.’ Ramsey points out (1978, p. 45) that if we allow ourselves to quantify over

propositions (or sentences, or some such), we can read this as ‘For all P, if he says that P then P.’ (This and more complicated cases are discussed at length by Grover, Camp and Belnap 1975.)

In the present case the claim at the heart of the appeal to success is something like

(6.1) True beliefs tend to lead to successful behaviour.

Applying Ramsey’s technique gives us

(6.2) For all P, if P then the belief that P tends to lead to successful behaviour.

If 6.2 does indeed mean the same as 6.1, then the notion of truth cannot really be doing any explanatory work. This is not to deny that we can explain our success by means of the hypothesis that most of our beliefs are true. It is simply to point out that in the formulation of this hypothesis, ‘true’ may be nothing more than a useful syntactical device, as the disquotational theories claim. This does not make the hypothesis empty. There is plenty of room for doubt as to whether

(6.3) For most P, if we believe that P then P –

or, as it might be better put, as to whether

(6.4) For all P, if we believe that P then it is probable that P.

If it is any good, the argument from success provides reasons to allay this doubt. But what we thus have reason to accept does not concern a more than disquotational notion of truth.

Moreover, it seems to me that this objection applies equally to the accounts that McGinn and Devitt suggest in place of Putnam’s success-based explanation of a realist notion of truth – as it does to the following suggestion from Field, from which McGinn and Devitt develop their accounts:

I think that the reason why we need to be able to apply semantic notions like truth and reference to the sentences that people believe and desire is that we hold the theory that people’s beliefs are, in many circumstances, reliable indicators about the world; and the only way to state this theory is to use the notion of truth *Moreover, this theory is not a piece of gratuitous metaphysics that could easily be dispensed with:* it is central to our getting information about the world, for we are constantly using our opinions about other people’s beliefs in forming opinions about the world These inferences evidently proceed by means of certain *reliability principles*, principles that say under what conditions a person’s beliefs about certain things are likely to be true. (Field 1978, p. 48)

A typical reliability principle is presumably something of the form

(6.5) Everything mother believes is true.

However, we have just noted that there is no need to regard such a use of ‘true’ as anything more than disquotational. What matters is that

(6.6) For all P, if mother believes that P, then P.

It is true that eliminating apparent references to semantic properties is more difficult when we believe only that mother is reliable on a limited range of topics. As Field says, ‘we want to be able to say in the theory that some people have very reliable beliefs *about physics* but very unreliable beliefs *about*

the state of the economy, and so forth' (1978, pp. 48–9). If mother knows her onions, we may want to say that

(6.7) Everything that mother believes *about onions* is true.

This becomes

(6.8) For all P, if mother believes that P and P is about onions then P.

But we are still left with the notion of *aboutness*, that Field takes to require a theory of reference. However, there is scope for treating 'about', like 'true', as an eliminable syntactical convenience. If we allow ourselves to quantify over predicates, for example, we can get close to 6.7 with

(6.9) For all *F*, if mother believes that onions are *F* then onions are *F*.

So it looks as though Field's reliability principles require neither truth nor reference, in the realist's sense.

McGinn and Devitt both modify Field's proposal by placing more stress on a broader notion of communication. Devitt says that we need to explain the fact that speakers both learn from and teach each other, and that 'it is these explanations that require talk of truth' (1984, p. 92). In so far as his account concerns learning, Devitt follows Field, emphasizing our need for judgements about the truth of the beliefs of others. As I have argued, these judgements need no more than a disquotational notion of truth (which is to say that they don't need truth at all, so long as they are prepared to put up with some syntactical inconvenience).

Concerning teaching, Devitt notes also that a speaker S must believe that an audience A will take an utterance or other sign as an indication that S has a certain belief (that it is raining, in Devitt's example). What matters is that A

takes the behaviour to 'mean that' it is raining; or, better, he takes it to be true if and only if it is raining. He takes S to be sincere, and so infers ... that S has a belief that is true if and only if it is raining. He takes S to be reliable and so comes to believe that it is raining. (1984, p. 93)

Putting aside reliability, the question is whether A's belief about the significance of S's behaviour (or S's belief about that belief) needs to mention non-disquotational truth. If we are allowing non-linguistic behaviour, it is surely misleading to say that A must believe that S 'means that' it is raining. If S means that it is raining then he is saying so, and his behaviour amounts to a speech act. In this case what A must believe is first that S has said that it is raining; and secondly (since A believes S sincere) that S believes it is raining. The non-linguistic case requires even less: simply the belief that the behaviour in question indicates that S believes that it is raining. Thus, there doesn't seem to be a role even for a disquotational use of truth.

It may be that the point Devitt is aiming for here is the one suggested by McGinn. McGinn proposes

that we locate reference in the point of communication – in the intentions with which assertions (and other kinds of speech act) are made. A hearer understands a speech act as an assertion just if he interprets it as performed with a certain point or intention – viz. to convey information about the world. ... To fulfil the intention to communicate knowledge (or belief) about the world in language the speaker must exploit signs standing in representational relations to things in the world. (1982, p. 226)

However, this suggestion seems to involve the view we rejected in chapter 2: namely, that assertions are distinguished from other kinds of speech act in virtue of being ‘aimed at truth’. McGinn seems to be proposing that in order to understand the point of an assertion, we must see it as made with such an aim; and hence that we require a notion of truth, thought of in terms of ‘representational relations to things in the world’. The lesson of chapter 2 is that as an explication of the notion of assertion, this gets the relationship between truth and assertion the wrong way round. If we try to characterize assertion in terms of the notion of being aimed at truth, we cannot say why not all utterances should be regarded as assertions.

The conclusion of chapter 2 was thus that in order to understand truth, we must first understand the nature of assertion. That is, we need to be in a position to understand why it is assertions, particularly, that admit the standards of correctness and incorrectness embodied in the notions of truth and falsity. At the time, the possibility remained open that this called for an independent analysis of the notion of assertion, on the basis of which we could hope to explain the function of ‘true’ and ‘false’. We have now rejected the analytic approach, however. As a result, we have to think not so much of the role of truth in the individual case, as of the function of the assertoric form of discourse as a whole. We have to think of truth not as a property uniquely possessed by linguistic entities of a certain kind, but as the most striking component in a pattern of linguistic behaviour, whose significance we want to explain. Instead of explaining truth in terms of assertion, we want to explain the assertoric pattern itself, of which truth seems simply the most striking facet.

Another part of the same pattern is perhaps the notion of representation itself: the very idea of ‘signs standing in representational relations to things in the world’, as McGinn puts it. Traditionally the notion of representation has seemed to be prior to truth. Truth has been viewed as *correct* representation. In order of development I think that these notions might be better taken the other way. Truth might arise first, in virtue of its role in argument; and representation follow, in virtue of our attempts to understand the nature of the ‘correctness’ of true beliefs and utterances. However, the important point is that explaining truth and explaining our idea that signs represent the world are part of the same enterprise. So although it is true of truth, as McGinn says of reference, ‘that finding a place for [it] forces us to take a world-directed view of communication’ (1982, p. 226), this is not the basis of an explanation of truth, but simply a comment on the nature of the explanatory task.

Thus it seems to me that the recent search for an ‘explanatory theory of truth’, stemming from the work of Field, does nothing to provide an explanation of truth in our sense. Nor, incidentally, does it seem to make any progress towards a theory that would be explanatory in Putnam’s sense – that is, in which truth is itself an explanatory notion. As we saw, the purported applications of such a notion seem to need nothing more than the disquotational property of truth.

The Power of Negative Thinking

I have suggested that we try to explain the existence and application of notions of truth and falsity in terms of their functional role. In particular, I suggested that their main function might be to encourage a useful form of linguistic behaviour, namely reasoned argument. It seems that in virtue of their normativity, these notions enable and encourage us to favour opinions with which we agree and to disfavour opinions with which we disagree. Disagreements thus become socially unstable, and we are led to defend our opinions, to argue and to seek agreement. In the long run, I suggested, such disputes can be expected to have a beneficial effect on the behavioural dispositions of individual speakers. They help to ensure that as individuals we hold and act on attitudes that reflect, to some extent, the combined wisdom of our linguistic community. Our behavioural dispositions can thus be tested against those of other speakers, before they are put to use in the world. The guiding principle is that it is better to be criticized for claiming that tigers are harmless than to discover one's mistake in the flesh.

Obviously, the notions of disagreement and dispute do a lot of work in an account of this kind. A natural suspicion is that these notions themselves need to be explained in terms of truth and falsity. This would put paid to the proposal to base a functional account of truth on its role in promoting such disputes. In this chapter our first task will be to explore this objection. We shall see that there are several forms it might take, but that none is inescapable. In the process we shall be led to a clearer account of the origins and conduct of such disputes and a better understanding of their behavioural advantages. That in turn will prepare the ground for the chapter's second main concern: to show how this functional approach to truth leaves room for non-factualism.

7.1 TRUTH AND THE TRIGGERS OF DISSENT

In its simplest form, the circularity objection would be the claim that a disagreement (of the relevant sort, at any rate) about an assertion P is a disagreement about the *truth* of P. However, in this form the objection is easily dismissed: for a dispute about the truth of P would have to be a dispute about the truth of 'P is true', and so on, *ad infinitum*.

A more interesting form of the circularity objection: How is it that we recognize the existence of a disagreement (and hence the grounds for a dispute) with another speaker, if not in virtue of the perception that he or she has said something *false*? The proposed account claims to explain our use of 'true' and 'false', in terms, primarily, of their role in these disputes. What it claims to explain, in other words, is our habit of describing the grounds for argument as being that our fellow disputee has said something *false*. The claim is that our use of this normative term serves the function of providing an incentive to the other speaker to enter into a dispute.

The account thus aims to explain the ordinary speaker's perception of the basis of a disagreement: the perception that someone else has spoken *falsely*. In order to explain it, however, the account needs to appeal to some more basic feature of the cases in which it takes this perception to be appropriate – to say that (perhaps unconsciously) we recognize the grounds for a dispute in terms more basic than the perception that someone else has spoken falsely. Similarly, the cue that a dispute is resolved cannot simply be the perception that the other disputee now speaks the truth, for the account wants to explain our inclination to describe it in these terms. So what could be the cues of argument, if not our judgements concerning the truth and falsity of the utterances of others?

A note here on terminology: I take dissent to be analogous to judgement rather than to assertion – that is, it happens in the head. A dispute begins when someone expresses dissent. Presumably some of the

utility of truth and falsity lies in their encouraging dissenters to speak up. But even silent dissenters may profit from the experience, by reconsidering their own beliefs.

A speaker makes an assertion, another expresses dissent, and a dispute is under way. We want to know how the latter speaker came to believe that dissent was appropriate. And we want to avoid the suggestion that it is a matter of coming to believe that the former speaker spoke *falsely*.

The mistake thus avoided would be that of pitching the cues of dissent too high – of making them depend on perceptions of truth and falsity. We should also be careful not to pitch them too low. We want to say that the biological utility of linguistic disputes stems from the behavioural consequences of the states of mind that come under challenge (albeit indirectly, in the form of their linguistic proxies) in such a dispute. So it is important that dissent be cued to something of behavioural significance. However, these behavioural consequences are long run affairs, and no doubt hideously complicated. It would be absurd to require that in individual disputes, speakers are cued by an understanding of the behavioural significance of the opposing points of view. Moreover, this would actually discourage disputes in cases in which the behavioural significance of differing beliefs is balanced by differing desires: I want to see Fred, and believe he is in the kitchen; you want to avoid him, and believe he is not in the kitchen; we both head for the kitchen, but one of us will be disappointed. So it is important that dissent be cued not by an awareness of the behavioural consequences themselves, but by some manifestation of the mental states that, in the long run, have those consequences. We need to separate the practice of dissent from the underlying facts that explain its development.

We want to be able to say that the triggers of dissent lie some- where between these two extremes – between perception of truth values themselves, and perceptions of behavioural consequences. The answer might seem obvious: shouldn't a speaker simply look for cases in which he asserts 'P' and another speaker asserts 'Not-P' (or *vice versa*)? This is too quick, however. For one thing, speakers may express conflicting points of view in completely different words. There need be no easy syntactical clues. More importantly, your 'Not-P' may not conflict with anything I have *said*; but if it conflicts with something I *believe* (i.e., that P), then I ought to dissent. So it appears that we need to credit speakers with the competence to discern that an utterance expresses a mental state incompatible with their own beliefs. There seem to be two components to such a competence. The first is the ability to *interpret* an utterance – to see what a speaker means. The second is the ability to detect incompatibilities between one's own beliefs on the one side, and a claim embodied in a received utterance on the other.

On some views of meaning and interpretation, the former ability depends on the notion of truth – for to interpret an utterance is to know under what conditions it would be true. However, we have seen that such views seem to require only a thin, disquotational notion of truth. This was the basis of our objection in chapter 2 to the proposal to derive a substantial theory of truth from a Tarskian truth theory, constrained by the needs of radical interpretation. So the ability to interpret an utterance does not seem a source of circularity in the proposal to base a thick notion of truth on its role in encouraging argument. (I do not mean that it is clear what the notion of meaning and interpretation will come to under such a proposal; I discuss this in chapter 9.)

This leaves the ability to detect incompatibilities between beliefs and hypotheses. The present view is not alone in crediting speakers with an ability of this kind. The question is not whether we have such an ability, but whether it depends on a grasp of the kind of notion of truth we want dispute behaviour to explain. As far as I can see, such a view could only arise from the idea that in judging that beliefs X and Y are incompatible, we are really judging that these beliefs could not both be true. This use of 'true' is less obviously thin and disquotational than some we have encountered, but only because it is applied to beliefs rather than to sentences. But why suppose that incompatibility judgements need to be *about* beliefs and hypotheses? What speakers need is not the ability to reason about their own mental states, but rather the ability to employ these mental states in reasoning about the world in

general. Prompted by the suggestion that Q, they need not to decide that the hypothesis that Q is incompatible with their belief that P, but simply to employ their belief that P in deciding that not-Q. It is a nice question just how we notice such incompatibilities. But it is hard to see why anyone should think that perceiving that P and Q could not both be true is any more easy, or primitive, than perceiving that P implies not-Q, or than utilizing one's belief that P in 'perceiving' that not-Q.

Thus for the moment I take it that what triggers dissent from the assertion that Q is simply the belief that not-Q. (Here 'belief' need mean nothing more than 'disposition to assert'.) This invites the objection that a grasp of negation itself depends on a grasp of the notions of truth and falsity, along the standard truth-functional lines. I have some sympathy with this objection, to the extent that I doubt whether we are entitled to help ourselves to the notion of negation in this way. I think it would be more accurate to say that negation has its origins in the expression of dissent. However, I doubt that this makes dissent dependent on falsity. Instead the view might be something like this: in acquiring a language, we learn to assent to sentences in some circumstances and dissent from them in other circumstances; dissent, like assent, is triggered by the activation of the sorts of habits thus acquired. Such a view of linguistic competence has independent advantages. (In Price 1983a I argue that it enables us to block Dummett's move from manifestation constraints on linguistic knowledge to a non-classical logic.) But I cannot defend it here.

7.2 TRUTH AND THE WORTH OF BELIEF

Let us turn to another form of the circularity objection. According to the proposed account, the utility of argument (and hence our use of 'true' and 'false') lies in its tendency to 'improve' the commitments with which we meet the world. Thus it rests on the principle that some commitments are 'better' or more valuable than others. But what is the measure of the worth of a belief, except its truth or falsity? Isn't a belief poor, or incorrect, just in case it is *false*? The objection is thus that in explaining the function of 'true' and 'false', the proposed account will have to appeal to the idea that commitments can be true or false in some prior sense.

The way to meet this objection is to note that on any plausible account, truth cannot be the only measure of the worth of belief. For imagine for a moment that there were a substantial property of truth, possessed by certain of our mental states. Why should we aim for beliefs that possess this property? Not for their own sake, presumably, but because such beliefs are valuable in some other sense – because they dispose to behavioural success, for example. Vague as this notion is, there is clearly no simple correlation, on anyone's view, between truth and behavioural appropriateness. False beliefs can lead to good outcomes, true beliefs can lead to bad ones, and many beliefs in fact have no behavioural consequences. Unlike truth, behavioural 'rightness' is bound to be a long run matter.

So from the point of view of behavioural utility, true belief is not an end in itself. Truth is at best an end in virtue of its usual consequences – something we should aim for in virtue of the long run behavioural advantages of doing so. If one already takes for granted a view of truth as a substantial property of mental states, this will be the natural way to construe its role in discourse; though it will need to be explained why this property should be a good guide to behavioural success. This is not the only possible view, however. Our proposal seeks to *explain* the ordinary notion of truth in terms of its role in discourse. Instead of beginning with the conviction that truth is a substantial property, we begin with the observation that ordinary speakers display such a conviction (and then seek to explain this fact). We thus regard truth as a 'mythical' goal of enquiry, whose popularity is to be accounted for in terms of the benefits of subscribing to such a myth. Of course, it remains to be shown that the main characteristics of the notion of truth can be explained on such a basis. However, the one characteristic that is immediately explained is the fact that truth is treated as the ordinary goal of enquiry. Here both sides agree on the appearances; and only a prior commitment to a more substantial view of truth could support the objection that our view fails to make sense of the fact that truth is the real end that we aim for in revising our beliefs.

The above objection calls also for the following reply. In the end, far from construing the benefits of argument as the production of true beliefs, our account should not depend on the notion of *belief* at all. As I said in chapter 6, our use of belief at this stage is no more than an expository convenience. It is dispensable, in principle, in favour of talk of behavioural dispositions. If this were not so then the present objection would be rather beside the point. A well-founded notion of belief, or indeed of representational attitudes in general, would bring with it not only an account of truth but also a basis for the fact-stating/non-fact-stating distinction. So relying on a substantial notion of belief would be damning in itself. The worst that could be said of relying on a substantial notion of *true* belief would be that it opens the stable door after the horse has bolted.

Thus far in this chapter we have been mainly concerned to show that the notions of disagreement and dissent do not themselves depend on those of truth and falsity. It is time to pay more attention to disputes themselves. We are interested in the nature and origins of their advantages, and hence in showing that in order to exploit these advantages, the goal of argument must have more or less the characteristics of the ordinary notion of truth. And we want to be on the lookout for the origins of non-factualism.

7.3 THE VALUE OF ARGUMENT

Arguments arise from disagreement and are resolved by agreement. In some sense, they thus serve to diminish disagreement among the members of a linguistic community. What could be the point of such a function? Why should disagreement be collectively or individually disadvantageous?

Let us for a start approach these questions in terms of an orthodox notion of belief. On such a view, beliefs have truth values, and a disagreement entails that at least one of the speakers involved has a false belief. This suggests that the disadvantageousness of disagreement rests on that of false belief. Resolving disagreements helps to correct false beliefs – not infallibly, of course, for a dispute may lead both parties to false beliefs – but perhaps more often than not.

But what is wrong with false beliefs? Here, as we saw, the answer will presumably be that false beliefs tend to lead to inappropriate behaviour. ‘Tend to’ is as strong as the claim can go – as we have seen, it is possible for mistakes at this level to cancel out, in virtue of the holistic nature of decisions to act. So on the standard view the justification for argument is something like this: verbal disagreements – for simplicity, pairs of utterances of the forms P and not-P – tend to indicate that one participant holds a false belief. (This ‘tend to’ stems from several sources, but perhaps most significantly from the fact that even the most well-intentioned conversationalists may not ‘mean the same thing’ – express the same belief – by saying P.) False beliefs tend to be a behavioural handicap. And argument tends to replace false beliefs with true beliefs.

I want to suggest that the orthodox notion of belief – of a truth-valued mental state – plays no significant role in this account. What matters is a certain correlation between utterance and action in general: roughly, that within a speech community, utterances of a given sentence tend to be correlated to mental states with similar causal or functional roles in the determination of behaviour. An utterance type is thus correlated with a class of mental states that share what I shall call the *Same Boat Property* (SBP). A class of mental states have the SBP if their typical behavioural consequences are such that their behavioural appropriateness, or utility, is predominately similar across a speech community. If a mental state has the SBP, then if it is appropriate for any one of us, it is appropriate for all – we are all in the same boat.

The notion of the SBP is far from precise; and not, I suspect, capable of being made significantly more precise. But provided that we do not expect it to bear too much weight, I think it will be useful at several points. Later it will help us to explain some of the variations from the statement form that are associated with non-factualism. More immediately, it clarifies the role of truth in the above account of the benefits of argument. For given the SBP, we can offer a justification for argument very

much like the one outlined above: verbal disagreements tend to indicate that one party is behaviourally disadvantaged (or potentially so); arguments tend to improve matters. This does not yet establish that in the earlier version, the standard conception of belief is redundant. After all, it might be the standard conception that entitles us to the SBP (for the class of mental states typically correlated with a given utterance). Indeed, it may seem that the required case of the SBP is a trivial consequence of the truth-conditional view of belief; that beliefs with the same truth condition are clearly appropriate in the same circumstances – namely, when and only when their truth conditions obtain. However, this ignores the fact that there is no simple correlation between truth and appropriate behaviour. So the orthodox account above of the value of agreement does not *derive* the SBP (from sameness of truth conditions); it assumes it, in the form of the assumption that throughout a community, false beliefs tend to be a behavioural liability.

The principle that some mental states have the SBP can be regarded as a theoretical hypothesis. The account requires that it hold to a sufficient extent, in some form, to explain the utility of a procedure whose function is to encourage resolution of certain disagreements between the commitments of different speakers. In the end, if need be, it can derive its plausibility partly from that of the theory that requires it; and partly from the demonstration that particular violations of the principle turn out to be correlated with departures in usage from the procedure that it explains.

The step from the hypothesis that certain of our mental states have the SBP to the claim that argument is advantageous depends on two main assumptions. One is that this is a case in which individual interests recommend co-operation, which of course is far from true in general, even within a single species. But this is hardly likely to be controversial, given that language itself is a co-operative and social activity. The co-operative basis of dispute behaviour seems simply a special case. (After all, we are free to avoid disputes – to decline to co-operate – if it seems to our advantage.) More importantly, the step depends on the assumption that in general, at any rate, pushing for agreement does more good than harm. Clearly disputes can be counter-productive, leaving both participants with inappropriate commitments. However, it seems reasonable to assume that on average argument does improve matters, both by encouraging reconsideration of individual grounds for judgement, and by the effect of pooling of evidence.

This account of the function of disagreement and argument thus rests on the notion that in virtue of their behavioural consequences, direct and indirect, some commitments are more appropriate than others; and moreover, that the appropriateness of a given commitment is more or less constant throughout a linguistic community. These factors enable individual speakers to benefit from a social activity, made possible by language: the drive to seek agreement. This drive is fuelled by the conventional notion of an external standard for belief and judgement, in virtue of which disagreements are socially unstable. The concept of truth is the catalyst that drives our systems of commitment through reactive argument, in search of the equilibrium that comes with agreement.

So why does disagreement matter? The account gives two answers. Internally, from the participant's perspective, disagreement matters because it is the cue for a mode of criticism – we criticize judgements with which we disagree and are criticized ourselves in turn. Externally, from the explanatory perspective, disagreement matters because in the long run it is an indication of behavioural liability. How are these perspectives related? The account suggests that the external explains the existence of the internal – or more specifically, of the special notion of *correctness* on which the internal perspective rests.

Note that agreement is not an end in itself, on this view, on either the internal or external levels. Externally, agreement need have no value of its own, but only the contingent virtue that stems from being the normal result of an otherwise useful process. And internally, it is simply the state in which disputants accept that they have achieved their common goal – that is, as they see it, the truth of the matter.

Indeed, it is crucial that disputants should not treat agreement as an end in itself. In the absence of any felt allegiance to an external goal, agreement is easily engineered – for example, as in many animal societies, by a system of deference to authority. I don't mean to suggest that argument is free of deference to authority, and such like. Clearly it isn't, in practice. Nor do I deny that such ways of reaching agreement can be valuable, even in human societies. The point is rather that there is a quite different benefit, that stems (in practice imperfectly) from the conception of an external goal for assertion and judgement. This benefit comes not from the social benefits of agreement itself, but from the improvements in individual behavioural performance that stem from the reactive process of which agreement is the stable state.

Imagine, for example, a community of creatures who use sounds as signals for a variety of purposes. Certain sounds might be uttered in response to significant features of their environment – food, dangers of various kinds, and so on; others, perhaps, in advance of significant kinds of behaviour, not otherwise predictable by other members of the community. In both cases it is easy to see how agreement – uttering the same sounds at any given time – might be valuable in itself. To the extent that the community's success depends on co-ordinated behaviour, it will be important that the members of the community are disposed to the same behaviour at the same time.

This aside, however, there might be a quite separate point to noting and resolving disagreements. It might enable individuals to improve their own behavioural commitments, even when these commitments have no role to play in the co-ordinated behaviour of the community as a whole. Such resolutions could be achieved in a number of ways. One possibility, obviously appropriate in some cases, would be to defer to experts. On visual matters, bats would do well to defer to hawks; on auditory ones, the reverse. (This doesn't mean that hawks and bats need co-ordinate any resulting behaviour, of course.) On some matters, however, the community might possess no experts. It would then seem advantageous for everybody to be encouraged to have their say; in other words, for disagreement to be resolved by argument among equals. I am suggesting that it is in this context that the normative notions of truth and falsity are important. They are at least part of what it takes to encourage the members of a community to argue when disagreements are noticed. The method is not infallible, of course. Once in place, for one thing, truth itself becomes a matter on which authoritarians may exercise authority. But an imperfect system of norms may be very much more useful than no system at all.

7.4 TWO COMPARISONS: ELLIS AND HABERMAS

The above point marks one of the main differences between the account of truth I am proposing here and one of the very few explicitly explanatory approaches to truth in the recent philosophical literature: the account suggested by Brian Ellis, in a paper called 'Truth as a mode of evaluation' (1980). Ellis distinguishes 'objective and subjective modes of speech':

To say that '*p*' is true, or more conventionally, just *p*, is for me to express my belief that *p* objectively. But for me to say that I believe that *p* is to express my belief that *p* subjectively In speaking in the objective mode, the belief or attitude we express is disembodied and put forward neutrally for discussion as though it were a thing that had some independent existence The participants in the discussion see themselves as arguing about this neutral detached object, and deciding whether to accept or reject it. (1980, pp. 88–9)

Ellis associates truth with the objective mode of speech. He leaves it unclear, I think, why this mode should employ both the forms '*p*' and "'*p*' is true.' That point aside, he suggests that objectification is a

manifestation of a process, ... necessary for human communication, which makes it possible for human beings to develop what is roughly a *communal* system of beliefs and attitudes, which in turn is necessary for co-ordinated and co-operative action We are social beings,

and our survival depends on cooperative activity. How we act normally depends on our desires, values, and belief systems And we may suppose that intersubjective agreement in our belief and value systems at least *facilitates*, even if it is not a necessary condition for, such activity. (1980, pp. 89–90)

Thus it seems that Ellis sees the advantage of the objective mode almost entirely in terms of benefits of agreement *per se*. It is true that for social creatures agreement or commonality of purpose is often useful, indeed indispensable. As I have said, however, I don't think that this can explain why agreement should be sought in argument, rather than say imposed by authority. In so far as what matters is simply that we all do the same thing, it would be more economical for evolution to arrange that we all defer to a leader. The fact that we take the trouble to argue suggests a further benefit, not guaranteed by easier paths to agreement. I have suggested that this benefit lies in the tendency of argument to improve a participant's individual adaptation to his or her environment. I think there is something of this idea at only one point in Ellis's paper:

Basically, we are people with belief systems who are concerned to interact with each other, *and so to widen the bases of experience upon which our belief systems depend*, and to reach intersubjective agreement about things. (1980, p. 98, my italics)

In virtue of its emphasis on *agreement*, Ellis's approach to truth seems vulnerable to the usual objection to coherence theories. It is often felt that such theories fail to make sense of the intuition that truth has something to do with answerability to some *external* standard – the intuition, to put it crudely, that we can agree until we are blue in the face, yet still be *wrong*. Correspondence theories try to make sense of this world-imposed 'rightness', as the basis of an *analysis* of truth.

Like Ellis's, the present approach rejects this attempt at analysis, in favour of an explanation of our possession of the notion of truth. Unlike Ellis's, however, my account retains the intuition that world-imposed rightness has something to do with truth. On my account it is crucial to the explanation of truth that some commitments are (or tend to be) more appropriate than others. This 'appropriateness' is world imposed. Behaviour that is successful in one world would be disastrous in another. This notion of appropriateness is not an analytic base for truth, however, but part of the explanatory base. In the terms I distinguished above, appropriateness belongs to the external perspective, whereas truth belongs to the internal perspective.

I have emphasized the argumentative role of truth and its contrast to authority as a means of achieving agreement. These two points may call to mind another similarity the present account bears to contemporary work on truth. Both are prominent in the theory of truth developed in recent years by Jürgen Habermas. Here I can do little more than to note this connection. Two cautions seem to be in order, however. For one thing, it is unclear to what extent Habermas's theory is intended or can be regarded as an explanatory account of truth. Certainly it has been interpreted as a variant of a Peircean consensus *analysis* of truth (see Pettit 1982, for example). Secondly, it is far from clear that Habermas would countenance the naturalistic and evolutionary flavour of the account here proposed; Hesse (1978, p. 390) suggests that 'any such interpretation of human evolution seems to fall foul of Habermas' own strictures against Chomsky's naturalism [cf. Habermas 1970] – it is itself a social theory and hence presupposes its own normative standpoint.' (I return to this kind of objection in chapter 9.)

7.5 ARGUMENT AND THE SOURCE OF BELIEF

We have been discussing the value of disagreement and argument, and hence trying to account for the nature of the ordinary notion of truth. So far, we have characterized the benefits of argument in terms of its effect on the very commitments whose linguistic expressions are the overt matters of dispute. It is time to recognize another benefit. An inappropriate commitment may be not only a behavioural

liability in itself, but also a manifestation of a potentially more serious liability – a sign of a flaw in the system of background beliefs, rules or habits of inference, from which the commitment in question arose. It seems that disputes have the function not merely of correcting individual mistakes, but also of encouraging improvements in the general beliefs and dispositions that lead to those mistakes.

It will be helpful at this point to appeal to a simple model of the origins of belief. Let us say that a person typically forms a belief in virtue of two things: a prior *evidential* belief, and a general *inferential* disposition to adopt a belief of one kind, given a belief of another kind. Since our use of this model will be largely heuristic, accuracy need not be one of its virtues. To our folk-psychological intuitions, however, many cases of belief formation do seem to be of much this kind.

The model allows disagreements to be traced to two possible sources. If we disagree about the truth of some proposition P, we may have had different evidential beliefs, or we may have employed different rules of inference. Our disagreement may turn out to be *evidential* or *inferential* (or a combination of both, of course). And our dispute, if we engage in one, has the potential not only to rid one of us of a behaviourally disadvantageous attitude to P, but also to correct the source of that mistake.

I think it would be an interesting and profitable task to examine the structure of such arguments in more detail. The model seems to throw light on the connection between truth, inference and validity. It seems that in the process of argument, speakers attempt to trace the claimed ‘incorrectness’ of their opponent’s assertions to its roots in any proffered defence. At the same time they try to import ‘correctness’ from the claimed roots of their own position. Central to this is the notion of an inference that claims to preserve ‘correctness’, and can itself be challenged on the grounds that it fails to do so. However, our immediate concern is with another application of the model. As yet there is nothing in the notions of disagreement and argument to indicate that there is a place for non-factualism. This model will help us to explain the limits of truth, and hence the origins of non-factualism.

In section 7.3 I emphasized the importance of the SBP (‘Same Boat Property’), only in virtue of which can it be true that if two people hold conflicting commitments, then one of them is likely to be behaviourally disadvantaged. Without the SBP there would be no reason to encourage agreement, and hence none to treat disagreement in evaluative terms – as an indication that someone has made a *mistake*.

Given that we do think in these terms, argument seems to be an attempt to trace the origins of the disagreement to some evidential or inferential basis. In effect, argument consists in a search for what both parties will agree to be the original mistake. In general, in other words, we argue as we should if the justification for evaluative treatment extends to the precursors of the belief at the centre of an initial disagreement; as we should if the SBP applies as much to evidential beliefs and rules of inference as to the judgements that lead us into conflict.

However, it turns out that this extension of the SBP is not always in order. There are cases in which it seems proper that speakers be considered to be *initially* in different boats – and therefore *initially justified* in disagreeing – even though not justified in preserving that disagreement, when each becomes aware of the other’s point of view. We shall see that such cases populate the margins of factual discourse. It is in these ill-defined fields of usage that non-factualism finds its place.

7.6 PROBABILITY AND THE END OF ARGUMENT

The most striking cases, in my view, are those involving probabilistic judgements. Such judgements reflect evidence that is less than complete, in the following sense. If I am interested in whether P, then I make a judgement such as ‘It is probable that P’, or ‘There is an *n*% chance that P’, only because my

evidence does not permit me actually to determine whether P. I do the best I can with the limited evidence available to me.

However, my evidence may differ from yours. It need not be that either of us is mistaken about our evidence, but simply that we have access to different facts. So my best judgement about the probability that P may differ from yours. But when I declare that it is probable that P, you may well reply that on the contrary, it is rather unlikely that P. There appears to be a disagreement, the basis for an argument. When the facts emerge, however, we needn't insist that in the first place one of us must have been mistaken. On the contrary, I may well recognize that in your place I would have made your assessment of the probability that P, and *vice versa*.

It is important to appreciate that the SBP does still apply in a limited form. Once we become aware of the difference between our judgements of the probability that P, then we are in the same boat. On the one hand, the behavioural consequences of probabilistic judgements are the same for both of us (subject to the usual qualifications). And on the other, since in principle we now have access to each other's evidence, there is no longer a systematic difference between us, such as to justify different judgements about the same probability. So there is a reason for us to seek agreement, and hence for convention to apply the normative incentives that encourage us to do so.

The fact that the SBP does not generalize to the *initial* disagreement suggests that in a properly organized linguistic community, speakers would be prepared to withhold criticism of each other's initial utterances. Indeed, they may wish to commend, rather than criticize, if it becomes apparent that the initial difference of opinion rests on their each having access to a different body of evidence. So if 'true' and 'false' are evaluative terms whose function is to encourage useful revision of judgements and patterns of inference, then we should expect that speakers who discover that a disagreement is of this kind would be reluctant to say that the other party's initial claim was *false*. And this does seem to be what happens. When we discover that our disagreement has this sort of basis, we are content to say that we were both correct, given the evidence we then had; and we go on to consider the bearing of the combined evidence.

Some terminology: I shall say that in such a case the initial disagreement *evaporates* or becomes *insubstantial*. A disagreement is *substantial* so long as the parties concerned regard each other as mistaken, and evaporates when such evaluations are no longer felt to be appropriate. We shall see that this can happen in several ways.

In the probabilistic case, disputes seem to run as follows. Any disagreement is initially treated as substantial. If it then transpires that the disagreement rests simply on the fact that the disputees originally had access to different bodies of evidence – in other words, that each recognizes that the other has made best use of his or her own evidence – then the original disagreement evaporates. Once both parties have access to their combined evidence, however, any remaining disagreement is treated as substantial.

What determines whether a disagreement about a particular sort of judgement should always be treated substantially? The following answer might suggest itself: a disagreement is or at least should be treated as substantial if it is a dispute about a matter of fact – if it answers to an objective reality, in virtue of which one party must be mistaken, as long as they disagree. As will by now be apparent, I think that this gets things back to front. I suggest that judgements of a certain kind are factual just to the extent that we do treat all disputes involving such judgements as substantial. Evaporative disagreement marks the limits of what we might characterize as the factual pattern of usage. In other words, the crucial feature of the factual pattern is that it requires us to say that of any pair of conflicting judgements, at least one must be false.

In chapter 6 I distinguished the descriptive and explanatory components of a non-analytic approach to the notion of statement of fact. This characterization belongs on the descriptive side. The job of the

descriptive part of such a theory is to characterize the manifestations of the distinctively fact-stating form of language; and to do so in such terms as to enable us to explain its existence. I think that by focusing on dispute behaviour – especially on the extent of substantial disagreement – we can do much to explain the relevant distinctions we find in usage in terms of the value and limits of argument. At the same time, I think we can do considerable justice to non-factualism, at least in its more important applications. For we can show that the intuitions that seem to support non-factualism are very much those that incline us to leave room for evaporative disagreements. Non-factualism seems largely a product of these intuitions, viewed through the lens of philosophical analysis. The present approach rejects the analytic perspective, but it enables us to explain the intuitions that that perspective distorts.

In the remainder of this chapter I want to defend the suggestion that factual discourse be characterized in these terms against a variety of objections. In particular, it needs to be shown that the distinction between substantial and insubstantial disagreements puts the limits of factuality in roughly the right place. By this I don't mean that we can tell in advance where the right place is; but simply that it would be a *reductio* of the proposal to characterize statements in this way, to show that evaporative disputes commonly occur about utterances whose factual nature is not in doubt. So to begin with we need to show that what gives rise to evaporative disputes in the probabilistic case is not a feature of assertoric utterance in general. We shall also want to show – anticipating a likely objection – that evaporation cannot adequately be explained by the supposition that probabilistic judgements are indexical.

7.7 COMMITMENT AND THE CASE FOR TOLERANCE

We have observed that probabilistic disagreements are liable to evaporate when they turn out to stem from the fact that speakers have relied on different evidence. The same might seem true of almost any kind of judgement. It often happens that although we disagree with another speaker, we recognize that in her position, with her evidence, we would have come to the same conclusion ourselves. Accordingly, we are inclined to excuse her mistake. Doesn't this show that almost any assertoric utterance can give rise to an evaporative dispute?

I think that these cases are actually quite different from those that can stem from probabilistic judgements. The difference shows up in the use of the labels 'true' and 'false' in the two sorts of case. In the probability case, we observed that we are reluctant to say that a probabilistic judgement is false when it turns out to have been properly based on different evidence from our own. On the contrary, we are likely to say that it was 'true, given the speaker's evidence', or something of the sort. But now consider a case of the other kind. In fading light, a guest mistakes my rabbit for a rat. 'There's a huge rat in your garden!', he cries. I reassure him, and politely excuse his mistake: 'That rabbit *does* look ratty at this time of night.' Excusing the mistake is not the same as acknowledging that no mistake has been made, however. If it came to the point, then, clearly, I would continue to maintain that the guest's assertion was false. From his point of view, it was understandable, perhaps, but certainly not true. There is no room for points of view on the question as to whether a creature is a rabbit or a rat.

The special character of probabilistic judgements shows up particularly in the following fact. It can be reckoned correct to make a probabilistic judgement even if one knows of the existence of evidence which, if one had it to hand, might make one revise that judgement. For example, a doctor might say, 'You are probably not infectious, but if we had your test results we would have a better idea'; or even, 'You are probably not infectious, but the test results will tell us one way or the other.'

Contrast these examples with the following: 'You are not infectious, but if we had the test results, we would have a better idea'; or 'You are not infectious, but the test results will tell us one way or the other.' (Or: 'That's a rat, but in better light it might turn out to be a rabbit.') Such assertions take away with one hand what they have given us with the other. Utterances such as 'You are not infectious', or 'There's a rat in the garden', seem to be expressions of 'final' commitment, in the

sense that to accept such an assertion is to accept that further evidence is unnecessary. If there is a rat in the garden, then any evidence that seems to suggest the contrary is misleading. As long as we are committed to claim that there is a rat, we are committed to the view that the case is closed – additional evidence is mistaken or irrelevant, if not confirmatory. The above examples violate this ‘closure principle’, explicitly rejecting the commitment to ignore further evidence.

Gilbert Harman has discussed this feature of commitment. In a recent book on reasoning he observes ‘that people find it difficult to accept things only tentatively ... There is a strong tendency quickly to convert such tentative acceptance into full acceptance’ (1986, p. 49). Harman characterizes ‘full acceptance’ as follows:

Belief in or full acceptance of P involves two things. First, one allows oneself to use P as part of one’s starting point in further theoretical and practical thinking. Second, one takes the issue to be closed in the sense that, when one fully accepts P, one is no longer investigating whether P is true. ... Full acceptance of P may therefore involve an implicit commitment to the claim that further investigation of P would not be worthwhile, for example, in leading one to discover further evidence that would lead to a reassessment of one’s conclusion. (1986, pp. 47–8)

It is a nice question why we make such final commitments. Clearly, they are often unjustified. The very fact that we can appreciate this, that we can change our mind, shows that in a wider sense our actual commitments are not final. On the whole, we do not close off the possibility of conflicting evidence. Why then don’t we express ourselves in a form that allows for this possibility? Why doesn’t my startled guest cry, ‘There is probably a large rat in the garden!’? Or more interestingly, why isn’t ‘There is a large rat in the garden’ simply understood in this way, so as to allow for insubstantial disagreements? (The interpretation would not have to be probabilistic. ‘It seems that P’ has much the same effect.)

A possible answer is that in the development of language full acceptance and assertion simply came first. Perhaps we could do without it, at least in theory. But in practice it developed before the various forms of qualified assertion, and so we have both. Harman suggests a less contingent constraint: that in practice there may be no alternative to full acceptance, for creatures of our limited mental abilities.

If one had unlimited powers of record keeping and an unlimited ability to survey ever more complex structures of argument, replies, rebuttals, and so on, it would be rational always to accept things only tentatively as working hypotheses, never ending inquiry. But since one does not have such unlimited powers ..., one is forced to limit the amount of inquiry in which one is engaged and one must fully accept most of the conclusions one accepts, thereby ending inquiry. Tentative acceptance ... cannot be the general rule. (1986, p. 50)

These factors suggest that most commitment is full commitment, and hence that most expressions of commitment are normally expressions of full commitment. But they do not seem to entail that ordinary expressions of commitment must always be taken as expressions of full commitment. That is, they seem compatible with the possibility that all or most assertions can in exceptional circumstances be the basis of an evaporative disagreement. I do not want to deny this possibility, which in my terms is the suggestion that the (descriptive) distinction between statements and non-statements is always a matter of degree. But the issue is somewhat premature, until we have a clearer idea of the range of possible ways in which utterances can give rise to evaporative disagreements. For the moment I want simply to make the point that probabilistic judgements are subject to evaporative disagreements, for a reason not characteristic of assertoric judgements in general.

This reason can now be described as follows. Probabilistic judgement is a species of intrinsically ‘non-final’ judgement. Probability has its origins in situations in which we feel that final commitment is unwarranted. In this context not only is judgement sensitive to evidence, but there is no warrant for

ruling out further evidence. Suppose for example that we are interested in whether P, and that the context is such that we make judgements about the probability that P. These judgements are based on our evidence about P – and we would not make such probabilistic judgments if we thought that the evidence was sufficient to warrant a final judgement that P (or that not-P). In other words, we hold open the possibility that further evidence would make a difference to our views as to whether P. But this means that whatever judgement we do make must be non-final – we must hold open the possibility of revising it, if new evidence comes to hand. So the non-finality of judgement about probability stems from its origins in the expression of non-final judgements about other topics.

It is illuminating to contrast this explanation to a characteristic argument for non-factualism about probability, that we encountered in chapter 4. That argument focused on the same feature of probabilistic judgement – its origins in the expression of qualified judgement about other matters – and hence observed that probabilistic assertions typically express ‘partial beliefs’, or degrees of confidence. The non-factualist wanted to infer from this that such assertions do not express full beliefs, and hence that probabilistic utterances are not statements of fact. However, we saw that a factualist need not accept this inference; for the non-factualist has not shown that the same mental state could not be both a partial belief that P and a full belief that it is probable that P. On the contrary, the factualist might suggest, the notion of a degree of confidence, functionally construed, provides a characterization of the content of a belief about probability.

The present argument does not depend on this questionable step. It is not intended to show that probabilistic statements are not factual, but rather to explain the fact that their usage differs in a significant respect from that of typical non-probabilistic statements. The argument does not take this difference to prove that probabilistic utterances are not ‘genuinely’ assertoric, descriptive, or whatever; for it belongs with a point of view that rejects such underlying distinctions. It simply attempts to explain, in terms of the most evident function of probability in discourse, a subtle but important aspect of the application of the notions of truth and falsity to probabilistic judgements.

It may seem surprising that a disagreement about the very same sort of judgement can be sometimes substantial and sometimes insubstantial. There is a simple explanation, however. In the probability case, for example, I said that disagreements evaporate in cases in which it turns out that agents are initially relying on different pieces of evidence. Cases like this are more common in some areas than in others. In scientific discussions, for example, it can usually be taken for granted that people share the same evidence, and moreover that new evidence will not come to hand, at least in the context of discussion. Probabilistic judgements can then be treated as final, and evaporation will be rare.

Even in more ordinary contexts, there seem to be good reasons for treating a probabilistic dispute as substantial, until it proves not to be. For one thing, it helps to ensure that the important disputes – the cases in which somebody has made a mistake – don’t slip through the net. It is better to risk prosecuting a few innocent differences than to risk excusing a guilty one. Secondly, we noted that if the grounds for an insubstantial disagreement are that speakers have initially relied on different evidence, then those grounds are effectively cancelled by the communication that calls attention to the disagreement. Communication puts the SBP into effect – if two people can disagree, then they can compare and combine their evidence. In order that they have an incentive to do so, it seems important that any current disagreement be treated as substantial. Only if such a disagreement is treated as if there is a single correct (or ‘true’) answer, will participants be encouraged to reassess their positions, to re-examine the latest evidence, and hence to strive for agreement.

Thus in the probability case it seems that only non-current disagreements should be allowed to evaporate. (Not all non-current disagreements, of course, but only those that turn out to have stemmed from access to different bodies of evidence.) The basis for such disagreements is automatically cancelled by communication. It is natural to wonder whether this is a peculiarity of the probability case; or whether, on the other hand, all current disagreements should be regarded as substantial. To settle the question, we shall need to examine some non-probabilistic sources of evaporative

disagreement. That will be the main task of chapter 8. Before turning to non-probabilistic cases, however, I want to deal with the objection that a factualist can explain the existence of evaporative disagreements about probabilities by appealing to indexicality – by allowing that probabilistic judgements make implicit indexical reference to the evidence on which they are based.

7.8 THE INDEXICAL OBJECTION

An evident feature of indexical judgement is that apparent disagreements may turn out to stem from a misunderstanding as to the intended reference of an indexical term. For example, the judgements ‘It is cold here’ and ‘It is warm here’ are incompatible, in the sense that no one could correctly assent to both at the same time; but obviously both may be correct if said by speakers at different places or times. In chapter 5 we saw that a probabilistic factualist has other reasons to think of ‘It is probable that P’ along the lines of ‘*Given the present evidence*, it is probable that P.’ Does not such a reading, coupled with the above feature of indexical judgement in general, explain what happens when two speakers base judgements about the probability of the same state of affairs on different bodies of evidence?

There is an important difference between the two cases, however. Imagine for example that you phone me and say ‘It is snowing here today.’ I will disagree with you if and only if I think it is not snowing where you are; and agree if and only if I think that it is snowing there. But in a probabilistic case I may disagree with you even if I think your evidence does support your assessment of the probability at issue; and agree with you even if I think that it does not. In other words, a disagreement may exist in the probability case, even if both parties acknowledge that neither has incorrectly assessed his or her own evidence. There is no analogous possibility in the usual indexical case.

The objector will reply that in the probability case this possibility depends on a shift in the reference of the indexical term, in the course of a conversation. It is simply the possibility that, for example, I might agree that given *your evidence* it is probable that P; and yet disagree that given *the presently available evidence* it is probable that P. Clearly such changes of reference are likely to occur in conversation. At least in principle, conversation allows us access to each other’s information on the topic at hand. We may begin with different bodies of ‘available evidence’, but conversation ensures that what is now available to each of us is our combined evidence.

But can this account for the facts? The impression that it does so seems to me to trade on the vagueness of the notion of ‘available evidence’. There are various readings of this phrase, ranging from the completely subjective – ‘The evidence of which I am actually aware’ – to the fairly objective – ‘The evidence available in principle.’ At the subjective end of the scale, we certainly explain how apparent disagreements about probability can evaporate; but at the cost of making it mysterious why they should be treated as disagreements in the first place. Take a forensic example: you say that your evidence points to organized criminals with contacts in high places; I think that mine suggests a clever twelve-year-old with a personal computer. My initial reaction is not to think that we disagree but to assume that we must be relying on different evidence. Moreover, since there is no evident disagreement between us there is no apparent motive for us to take matters any further. What needs to be explained is the fact that in a case like this we should and do go on to consider our combined evidence. On my account the motive is provided by our treatment of such a case as a substantial disagreement, answerable to a normative notion of correctness. (This is the *internal* motive, of course; the external justification comes separately. More on the latter below.) But on this subjective version of the indexical interpretation, the sanitized exchange provides no motive for anything more than polite acknowledgement of different but thoroughly compatible judgements.

So much for a subjective interpretation of the supposed reference to ‘available evidence’. What if this phrase is read as something less unambiguously subjective – for example, as something like ‘the evidence reasonably accessible’? The problem is now a dilemma. On the one hand, if this reading is taken sufficiently strongly to exclude the possibility that two parties to a conversation have different

‘available evidence’, then it excludes the very cases the indexical treatment is supposed to explain – those that seem to allow evaporative disagreements about probability. If on the other hand the reading allows for differences between speakers in their ‘available evidence’, then the above problem recurs. The supposed disagreement doesn’t get started, because speakers who understand probabilistic assertions in these terms will be inclined simply to attribute what would otherwise be a conflict to differences in the evidence that each refers to.

Thus it seems to me that an indexical reading cannot explain this distinctive feature of probabilistic judgement. The impression that it does so rests on a systematic equivocation concerning the supposed indexical reference – in effect, on a vacillation between the horns of the above dilemma.

7.9 EVAPORATION AND THE PROPERTY OF TRUTH

At this point a determined opponent, though conceding that the indexical interpretation does not provide an orthodox explanation of evaporation, might hold out the possibility of some other such explanation – some other way of accounting for evaporative disagreements within an orthodox truth-conditional interpretation of probabilistic judgement. I want to finish with an argument that seems to count against this possibility.

The analytic view thinks of truth and falsity as properties; the goal of analysis is an understanding of the nature of such properties. Suppose that P is a sentence (or statement, proposition, or whatever) of the kind that bears these properties. ‘P is true’ is then a respectable assertoric sentence in its own right – it can be used to say that a certain linguistic entity has a certain property. Moreover, since it is not probabilistic it cannot be the subject of the kind of evaporative dispute described above.

Imagine now that P itself is probabilistic. The analytic view faces a dilemma. On the one hand we have just observed that if two speakers disagree about a simple predication such as ‘P is true’ then one of them is mistaken, whatever their evidence – one of them has spoken falsely. Yet on the other hand to disagree about ‘P is true’ is to disagree about P. If the former disagreement entails that one party is mistaken, then so does the latter. If we allow evaporative disagreements about P, we must allow them about ‘P is true.’

If truth and falsity are properties, in other words, their literal ascription must be confined to sentences not subject to the ‘no fault’ evaporative disagreements characteristic of probabilistic judgements. It follows that our ordinary treatment of probabilistic judgements is not entirely compatible with the truth-conditional model. From the orthodox point of view, either we misuse such judgements, or they do not possess genuine truth conditions; for their use is incompatible with their having the property for possession of which truth conditions are conditions.

This point has been obscured, I think, by the assumption that the features of use on which it rests can be explained by indexicality – by some sensitivity of probabilistic truth to evidential context. However, we have seen that the appeal to indexicality itself rests on an equivocation in the relevant notion of ‘available evidence’. In any case, the result is an embarrassment for the analytic approach. In one way or another, it detaches the analyst’s desired ascription of truth conditions from its evidential basis in usage. It implies that we either use truth-bearing sentences as if they were not truth-bearing, or use non-truth-bearing sentences as if they were truth-bearing.

Note that the present approach encounters no such problem. An explanatory view is not committed to the idea that truth is a property, seeking merely to explain the fact that ordinary speakers are inclined to treat it as such. If it turns out that ordinary usage is sometimes incompatible with a treatment of truth as a property, then so much to the credit of the explanatory view – at least so long as it can account for this fact.

I have argued that in the probabilistic case the appropriate account turns on the role of probabilistic judgement in contexts of incomplete evidence. In chapter 8 I want to show that by appealing to different factors in the same general spirit, it is possible to explain the peculiarities of certain other areas of discourse – areas which, like probabilistic usage, have been popular targets for the non-factualist approach.

The Limits of Intolerance

We are trying to account for the ordinary use of the notions of truth and falsity in terms of their function in encouraging argument. In view of our original concern with non-factualism it is important that the account should work ‘at the edges’ – in the borderline cases in which talk of truth and falsity seems to some degree inappropriate, and in which non-factualism has hence seemed particularly attractive. We looked at one such case in chapter 7. I argued that our occasional reluctance to apply the notions of truth and falsity to probabilistic judgements can be correlated with a systematic failure of the normal functional justification for the use of truth and falsity.

In this chapter I want to examine some further cases, and show that again the functional theory of truth can explain the kinds of peculiarities that have been taken to favour non-factualist treatments of the topics concerned. I’ll pay most attention to moral and conditional judgements (the topics whose non-factualist treatment we considered in part I). But I shall also comment on the account’s potential in other areas in which non-factualism has been appealing. And I’ll finish, as promised, with some remarks on the non-indicatives.

Perhaps I should say again that in attempting to explain the appeal of non-factualism, I won’t be trying to make a case for non-factualism as normally construed. Non-factualism normally takes for granted the analytic approach. It sees itself as concerned to determine the true extent of the property that distinguishes fact-stating discourse. Part I discredited that endeavour. As we’ll see, however, there is considerable scope for reconstruction of non-factualist arguments along explanatory lines. A subsidiary concern of this chapter will be to distinguish several stances that remain open in this connection, when the analytic project has been closed down.

8.1 RELATIVISM AND EVAPORATIVE ARGUMENT

In chapter 7 we drew a rough distinction between two sorts of linguistic dispute, according to the root of an initial disagreement: in an *evidential* dispute the root lies in speakers’ evidential beliefs; in an *inferential* dispute it lies in the rules or habits of inference that lead from evidential beliefs to the conflicting judgements that are the immediate grounds for disagreement. We noted, however, that disagreements can have multiple roots; and the distinction between evidential beliefs and rules of inference is in any case tendentious.

At that stage we had not raised the possibility of evaporative disagreements. However, the one sort of evaporative disagreement we have so far considered is clearly evidential rather than inferential. It turns on the fact that in making probabilistic judgements, speakers may have access to different bodies of evidence. So long as the speakers concerned do not disagree with each other’s evidential beliefs or probabilistic inference rules, we have an evaporative evidential disagreement.

Can there be evaporative inferential disagreements? I think that there can, and indeed that some of the best-known tenets of the non-factualist movement amount to the endorsement of this possibility. For consider moral or aesthetic relativism – the view that different individuals or groups may hold to incompatible but equally ‘valid’ aesthetic or moral principles. Such a possibility would reveal itself in a case in which the two groups or individuals make conflicting ethical or aesthetic judgements, on the basis of agreed non-moral or non-aesthetic facts. Since the non-evaluative evidence is agreed, the disagreement must stem from the general evaluative principles, in virtue of which each group or individual infers moral or aesthetic conclusions from such a basis. In other words, in our present

terminology, the disagreement has an inferential basis. The relativist claims that in such a case, neither side need have grounds for regarding the other as mistaken. But in our terms, such no fault disagreement is the mark of insubstantiality; and hence the relativist seems to be urging the possibility of evaporative inferential disagreements.

This case bears out some points I made earlier. It illustrates both the element of reconstruction of non-factualism that follows the rejection of the analytic viewpoint, and the fact that the reconstruction leaves open several distinct kinds of non-factualist conclusion. In appealing to the possibility of relativism, non-factualists in ethics and aesthetics have on the whole been concerned to argue that moral and aesthetic truth *is* relative – that it is relative, as a matter of fact, and therefore should be treated as relative. It follows from this view that if a community treats moral truths as absolute, its members are mistaken about the nature of moral judgement. In other words, the non-factualist traditionally begins with claims about the ‘real’ status of moral judgements, and proceeds to prescriptions concerning the use of moral utterances. In rejecting the analytic point of view, we seem bound to reject this form of non-factualism. We have given up the idea that there is an independent standpoint, with respect to which there is anything more to being a statement of fact than being treated as such. We may now describe usage and hence explain it, but there is no longer a place for trying to analyse usage and hence prescribe it.

I think that a good deal of contemporary non-factualism can be reconstrued in this descriptive sense, admittedly at the cost of systematic disrespect for the intentions of its authors. The trick is simply to reinterpret what the analytic non-factualist normally appeals to by way of evidence – such things as our ordinary intuitions about the truth and falsity of the kinds of judgement in question. From the non-analytic point of view this appears not as evidence for the presence or absence of some underlying property, but rather as the basis of a descriptive taxonomy of use. Much of orthodox non-factualism is thus directly accessible to our first kind of reconstructed non-factualist, whose role is to catalogue the subtle variations of usage on the boundaries of the statement form.

Perhaps this kind of non-factualism need not be merely descriptive. It cannot prescribe ordinary usage as such, but it might seek to correct any corruption of usage that could be attributed to the influence of a mistaken philosophical theory. All the same, to a traditional non-factualist this looks like a dull substitute for prescription grounded on analysis. Is there room for anything more colourful – for a non-analytic non-factualism that would seek to revise ordinary usage itself? I think there are two possibilities, stemming from two kinds of reflection on the use of the class of judgements in question.

8.2 SECOND ORDER EXPLANATORY NON-FACTUALISM

The first possibility would involve the recognition that actual practices of disagreement simply fail to resolve certain sorts of dispute. In the moral case, for example, it would flow from the perception that moral disagreements may simply reach a stalemate, in which neither party can do more than to continue to affirm some disputed moral principle. This perception needn’t itself recommend relativism. But the inherently unstable nature of such a stalemate, in which each party retains the conviction that the other is at fault, might seem to recommend a change in usage – a degree of relativistic tolerance.

The second possibility is perhaps only a more theoretical version of the first. It would concentrate not so much on the actual procedures we employ in argument as on the role of a type of discourse in our lives, and that of argument in general. From such a standpoint it might be argued that particular types of utterance would be better treated in one way rather than the other – ‘better’ being cashed in terms of the overall value to the individual, community or species of such forms of behaviour. Given the present functional account of truth, for example, it might be argued that we would do better to allow for evaporative moral and aesthetic disagreements.

Thus there seems to be room for a non-descriptive non-factualism – a second order theory whose prescriptions concerning usage would flow from a view of the significance of factual usage in general, and an understanding of the role of the particular area of discourse whose status is in doubt. I shall not directly attempt to decide to what extent existing arguments for or against relativism about morality and other topics can be interpreted in this light. But I do want to illustrate the range of considerations that might be relevant to such a second order non-factualism; and some of these, as we shall see, call to mind standard moves in the discussions of relativism about morality and other matters.

Let us begin at one extreme, with a challenge to the conceptual possibility of the second order stance. In the moral case, for example, the stance seems to depend on the claim that it is possible to discuss the merits of moral principles – of different rules of inference from non-moral beliefs to moral conclusions. This seems to depend in turn on an understanding of the meaning of moral judgements which would be prior to the adoption of a particular set of acceptance rules for such judgements. After all, these acceptance rules are what the discussion is about. Now the objection: such an understanding is impossible, because what we mean by moral terms is a function, *inter alia*, of what we take to be their assertion conditions. In a slogan, meaning *is* use, and therefore cannot prescribe use.

In effect, this is the objection that relativism collapses into the indisputable but uninteresting possibility that different speakers may simply use the same word with different meanings. The proper reply is obviously that some uses are more important than others. In the moral and aesthetic cases, at any rate, the feature of use that does most to fix meaning seems to be not the adoption rules in virtue of which speakers acquire commitments, but the ‘application rules’ – the conventions in virtue of which commitments have behavioural consequences. It is relatively easy to imagine someone who adopts moral and aesthetic commitments in different circumstances from ourselves. But as we have already had cause to observe, it is difficult to make sense of someone who means what we mean by ‘good’, ‘beautiful’, and so on, and yet who is not motivated much as we ourselves are by commitments involving these terms. There are certain exceptional cases, of course – for example cases in which the effect of such a judgement is overridden by some other motivation – but we would be inclined to regard the fact that someone lacked that pull *in general* as evidence that he meant something different from us by the terms concerned.

Conversely, it seems that these definite behavioural consequences prevent arguments over matters of moral and aesthetic principle from degenerating into terminological border disputes. If a person’s moral and aesthetic commitments – those of her commitments that are expressed in what are for us moral and aesthetic terms – have for her the same kind of behavioural consequences as such commitments would have for us, we take this as evidence that she means the same by these terms as we do. We are then in a position to compare her moral and aesthetic sensibilities to our own; and free to conclude, with the relativist, that although such a person’s sensibilities differ from our own, she need not be *mistaken*.

To show that relativism is a conceptual possibility is not to show that we are or should be relativists, however. The probabilistic case illustrates this point. As we observed in chapter 4, simple probabilistic judgements also have fairly definite consequences for behaviour. Roughly, judgements of increasing probability are reflected in an increasing willingness to risk a given loss for a given gain, in an action that depends for success on the state of affairs whose probability is in question. If someone consistently failed to act in this way, having sincerely professed to what appeared to be probabilistic commitments, we would be inclined to say that he did not mean what we mean by ‘probable’ and associated terms. Conversely, if someone does apply what appear to be probabilistic judgements in this way, then we are inclined to assume that she does mean what we mean. And we continue to assume it, even if we disagree with the statistical principles on the basis of which she makes her judgements. Probabilistic relativism is thus a conceptual possibility. However, it seems to have little of the attraction of its moral and aesthetic counterparts. Once evidential differences have been sorted out, we don’t say that *chance* lies in the eye of the beholder.

If we were to try to justify this difference between the probabilistic and ethical cases, I think it would have to be in terms of the claim that barring evidential differences, probabilistic judgements involve no systematic failure of the SBP. Presumably our ability to make probabilistic judgements has some survival value, and presumably its value would vary if we varied our probabilistic inference rules. (It is easy to think of cases in which the wrong rules would regularly lead to disaster.) Access to evidence aside, there seems to be no difference between speakers that could explain how the optimal probabilistic inference rules could be different in different cases. And in the absence of such a difference, there may seem to be no case for probabilistic relativism.

However, this is to overlook the possibility that there might simply be no optimal set of probabilistic inference rules. Any such set of rules is a strategy for making judgements on the basis of less than conclusive information. In principle there might be a range of such strategies, none clearly better than any other. Probabilistic relativism then begins to look appealing, I think, on inferential as well as evidential grounds. This suggests that the apparent lack of evaporative inferential disagreements in ordinary probabilistic usage may simply reflect the fact that ordinary probabilistic reasoning is not complex enough to have encountered the difficult cases. (If so then we might do better to look at the reasoning of professional statisticians. I am not sure to what extent differences between statistical schools of thought can be viewed in these terms.)

8.3 RELATIVISTIC PRACTICE: FOR AND AGAINST

The point remains that in the moral and aesthetic cases, as for probability, relativism is neither entailed nor recommended simply by the fact that it is a conceptual possibility. One further factor is whether moral and aesthetic judgements have the SBP. The relativist's traditional appeal to subjectivity can be seen as an argument that the property does not hold in these cases (and hence, in our terms, that moral and aesthetic disagreements are properly treated as potentially evaporative). The subjectivist claims that moral and aesthetic utterances are expressions of desires, motivations, or something of the kind. Ideally such attitudes should be correlated with dispositions to feel pleasure or displeasure in certain states of affairs. (Masochism is not a particularly good survival strategy.) Hence if different speakers can have different such dispositions, the SBP need not apply: if pleasure lies in the eye of the beholder, then so (we should allow) does beauty.

This sort of argument can be challenged at a number of points, however. In particular, it seems to depend on the view that appetites or motivations are simply givens, and not themselves contingent products of cognitive processes. If motivations are allowed to be changeable, then there are at least two lines of argument in favour of non-relativistic moral and aesthetic practices.

The first argument would rest on the claim that some practices are in some sense 'better' than others, and hence that there is a place for the kind of improvement that is encouraged by the notion that expressions of motivation are answerable to an external standard. 'Better' might be cashed in several ways. Perhaps most plausibly, it might be argued that some tastes and appetites are potentially more rewarding than others. Hence there would seem to be an individual advantage in a social process of moral and aesthetic discussion, argument and refinement. The more educated one's palate, the greater one's enjoyment, on this view – and hence we all gain, on the whole, by the process of comparison, criticism and improvement that is encouraged by the practice of treating moral and aesthetic judgements as true and false.

This argument seems more plausible for the aesthetic than the moral case. For the moral case, however, the recognition that motivations are changeable states suggests another line of argument. For it suggests that we regard moral judgements not as expressions but as sources of motivational states. Roughly, we tend towards what we take to be good not because 'taking to be good' is a manifestation of a pre-existing tendency but because 'good' functions as a conventional reward for action. Clearly, this is a view of moral discourse that parallels the present account of truth and falsity. It is a likely development of the theme, familiar in ethical theory, that moral concepts play a

regulative role in society. The view would thus be that morality provides a system of incentives that functions to encourage socially ‘desirable’ forms of behaviour (‘desirable’ being cashed in non-moral terms, of course).

However, such a system of incentives would be useless if everyone was held to be entitled to a reward, in virtue of satisfying his or her own individual standards. Relativism thus threatens the social function of morality – which perhaps explains why it is such a disturbing thesis, at least in extreme forms. The practice requires that the system of rewards be seen as objective, as answerable to an external standard; or in other words, that speakers regard differences in their application of moral descriptions as indications of mistakes.

Thus we have two possible arguments for disallowing evaporative inferential disagreements on moral and aesthetic matters. The first is based on the claimed advantages of educated aesthetic and (less plausibly) moral tastes. The second turns on the regulative role of moral and (less plausibly) aesthetic judgement. Both, I think, leave scope for limited relativism. The advantages of conformity may sometimes be outweighed by the disadvantages of intolerance. ‘Live and let live’ may sometimes be the optimal survival strategy. Perhaps we should behave as if everyone in our own community is in the same boat, but allow that foreigners do things differently.

I do not propose to examine the above justifications for moral and aesthetic objectivity in any detail. Our present purpose is simply to illustrate the variety of factors that, from the non-analytic perspective, may be held to explain or recommend a particular pattern of usage. However, the arguments do call for a number of comments. The first of these concerns the parallel between the latter approach to morality and our suggested account of truth.

8.4 TRUTH AS INCENTIVE – A CIRCULARITY?

We described ‘true’ and ‘false’ as incentives, whose purpose is to encourage speakers to seek agreement, and hence to participate in linguistic disputes. The second argument above suggests that in the moral case, such a system of incentives requires that speakers treat moral truth as objective – in other words, that they treat moral disagreements as substantial. This seems to imply that in the case of truth and falsity themselves, the claimed system of incentives would itself require that speakers treat disagreements about the truth of utterances as substantial. If so, isn’t the account in trouble? Doesn’t it commit us to an endless hierarchy of disagreements, corresponding to iterations of the predicates ‘... is true’ and ‘... is false’?

This is a nice point, to which I want to return in the next chapter. The brief answer is that although it is true that if ‘true’ and ‘false’ are to function as incentives then speakers must treat disputes about the truth of utterances as substantial, this is guaranteed, as long as it needs to be, by the general presumption that disagreements are substantial unless proved otherwise. Evaporative disagreements are the exceptions. There is a kind of circularity in the use of ‘true’ and ‘false’ as incentives, but it is not vicious. In order that ‘true’ and ‘false’ should function as incentives, their ascription takes the objective or declarative mode. Truth is presented as an objective property, that utterances either have or lack. We add truth to our idea of the world, to help us get on in it. But without truth, we would not have an objective or declarative mode of speech. What is distinctive about this mode of speech is that it is regarded as answerable to an external standard; that is, to truth.

8.5 PROJECTIVISM NOT VINDICATED

The second point arising from section 8.3 concerns an apparent contrast between the above case for a substantial treatment of moral disagreements and the kind of justification previously suggested – a difference that seems to echo the old idea that beliefs and desires have different ‘directions of fit’ to the world. In the previous non-evaluative cases (including the probabilistic case) we used the SBP to explain and justify the use of the normative notions of truth and falsity. We argued that because the

behavioural consequences of such commitments are broadly similar across a linguistic community, there tends to be an individual advantage in a conventional system of incentives and disincentives that encourages attempts to reach agreement. The world puts the members of a community in the same boat, as it were, and hence we tend to do better if we co-operate. In the moral case, however, the above argument may seem to reverse things – to suggest that we should behave as if an external moral reality had put us all in the same boat, in order that morality might serve its community purpose.

This is an important objection. If sustained, it would suggest that the present account is a form of the projectivist theory we discussed in chapter 4. For it would seem to admit a distinction between those areas of discourse for which truth can be cashed in terms of a world-imposed notion of correctness, and those in which there is a useful function for a pretence of world-imposed correctness.

However, I think the objection is mistaken. It is correct in noting a difference, on the proposed account, between the grounds for the application of the notions of truth and falsity in different areas of discourse (at least if their application to normative discourse is explained by the latter of the two arguments above). Indeed, I think the explanatory approach should explicitly affirm this possibility. Once we detach the explanation of the use of truth from the search for an analytic base, we should see it as possible, even probable, that across the range of functions served by the assertoric form of language, the advantages of truth-encouraged argument will not be the same in all cases. Certainly, we should expect that these cases have something in common, to explain the fact that they all employ the assertoric form and the notion of truth. But this degree of uniformity is compatible with considerable diversity in its explanatory base. The advantages of argument may vary from case to case.

The objection is mistaken, however, in so far as it assumes that in the non-evaluative cases (or some sub-class of these cases), the account provides a ‘genuine’ notion of truth, cashed in terms of ‘world-imposed correctness’. I think this mistake stems from a tendency to construe the SBP in terms of the ‘success’ account of truth. From such a point of view it seems obvious what property is shared by the same belief in different heads: that of leading to behavioural success in the same circumstances – that is, on this view, the property of being true in the same circumstances. However, in introducing the SBP I emphasized that it should not be conceived in these terms. In order to account for an evaluative notion whose function seems to be to encourage (reasoned) agreement, we needed to claim that disagreement is in general an indication of some sort of behavioural disability. That in turn seemed to require the assumption of the SBP. Without something of the kind, we could well have disagreements between speakers in cases in which each holds what from his or her own point of view is the behaviourally optimal belief. I noted that this can indeed happen in individual cases, such as when the effects of a mistaken belief are balanced by a mistake elsewhere in an agent’s system of belief. But to account for truth in terms of its propensity to encourage agreement, we need to assume that it does not happen globally.

The SBP was thus a theoretical hypothesis, needed at this point in our explanatory account of truth. As such, it derives support from the demonstration that departures from the standard use of truth can plausibly be correlated with specific reasons to favour (or at least not to penalize) disagreement, in special cases – with systematic factors in virtue of which the SBP might be expected to fail.

This route to the SBP is non-committal as to the nature and origins of notion of ‘appropriateness’ on which it depends. As just noted, there is therefore an understandable temptation to think of it in terms of truth, ‘correct representation of reality’, or something of the kind. This temptation must be resisted, however, for these are the very concepts whose use we are trying to explain. Behavioural appropriateness belongs to an explanatory but not an analytic base for the suggested account of truth. The key to the explanatory approach is the realization that these things can be kept apart.

Provided we do keep them apart, the above objection can be seen to be mistaken. Certainly there may be a difference between the senses in which the ‘appropriateness’ of evaluative and non-evaluative

judgements depends on more than individual preference or point of view. But this does not entail that in the latter case ‘appropriate’ can be equated with ‘true’. The difference lies within the scope of the explanatory approach as a whole. It does not, as for the projectivist, mark a point at which explanation takes over from analysis.

8.6 PRESCRIBING USAGE

The final point arising from section 8.3 concerns the sense in which the arguments for and against aesthetic and moral relativism can be considered to be ‘prescriptive’. It needs to be emphasized that their conclusions are ‘external’ rather than ‘internal’ prescriptions for moral and aesthetic usage. That is, they arise not from *within* a particular moral or aesthetic practice, but from an external standpoint – a point of reflection on the practice as a whole.

This distinction is not always easy to draw, but I think it has a place in many areas of discourse. Many practices yield prescriptive principles – principles that will seem *to the practitioner* to follow from the ‘ground rules’ of the practice in question. But these are internal prescriptions. They flow from the practice, and we cannot feel their force unless we already subscribe to the practice from which they stem.

External prescriptions, on the other hand, depend on an ability to detach ourselves from the practice in question, and hence to make recommendations about its conduct from some wider standpoint. This exercise is problematic in a number of ways. Even if we can achieve the required external perspective – survey usage from outside, as it were – there are difficulties with the notion of a recommendation from such a perspective. Here, as in other areas of the human and social sciences, the boundary between justification and explanation is far from precise. Naturalistic explanation tends to be presented as conservative defence.

Whether conservative or reformist, however, similar questions arise. In what sense is the recommended practice better than others? And how is it possible for us to adopt such a recommendation – to decide on a particular pattern of usage? These are difficult issues, and I have no definite answers to offer. But it does seem that we are not entirely constrained by current convention. In philosophy and in ordinary life we do have some ability to reflect on and to modify these kinds of patterns of linguistic practice; even if the process of change seems to involve not so much an explicit decision on current and possible usage as a communal shuffle in a direction we cannot quite describe.

I have suggested that some of the concerns of the non-factualist movement can be viewed in this light, as attempts to prescribe the appropriate pattern of usage for particular classes of utterances – to show, in particular, that the application of the notions of truth and falsity is sometimes inappropriate. I noted that to interpret non-factualism in this way is to separate it from its usual analytic presuppositions. However, we saw that prescriptive arguments of the above kind have a natural place in the analytic view. So it may be that we should regard the prescriptive side of non-factualism as merely a legacy of its usual analytic origins, and not try to reconstruct it on an explanatory basis.

My impression is that this is unduly pessimistic, and that there has been an aspect of non-factualism, not particularly tied to the analytic view, involving a potentially reformist process of reflection on usage. The point is incidental, however. Our main concern is with the explanatory view itself. In the remainder of this chapter I want to consider some further applications. The first of these involves the third of the main examples we discussed in part I – the case of conditionals.

8.7 CONDITIONALS

As we saw in chapter 5, the difficulties for a straightforward factualist interpretation of indicative conditionals seem to stem from the behaviour of conditionals with false antecedents. The suggestion that the indicative conditional is simply the truth-functional material conditional comes to grief on the

fact that English speakers do not take grounds for denying an antecedent as grounds for asserting a conditional in which it appears. This fact underpins Adams's hypothesis: the claim that the assertibility of an indicative conditional follows the subjective conditional probability of the consequent, given the antecedent. And Adams's hypothesis, in turn, underpins the anti-factualist arguments we discussed in chapter 5. Even if, as I claimed, these arguments are not conclusive, it is clear that the behaviour of conditionals with false antecedents makes things very much more difficult for a factualist interpretation.

From the explanatory viewpoint, however, there is a ready account of this behaviour. Suppose that the main function of the commitments we express with indicative conditionals lies in hypothetical reasoning. Crudely, the judgement that if P then Q is a kind of insurance, or contingency plan, against the possibility that P. (This seems plausible enough, but the following explanation could be modified to cope with other possibilities.) Presumably some contingency plans are better than others. Add the communal element – the SBP – and we have what it takes to make the argumentative quest for agreement a useful social practice. Individuals then stand to benefit, improving their contingency plans by comparing them to those of others. And hence there is a role for truth, the dialectical catalyst whose function is to provide the (internal) point of argument – to make disagreement seem to matter.

What then of false antecedents? How should we assess an insurance policy against an eventuality that we know has not arisen, and will not arise? That is, what bearing should it have on our assessment of a plan to deal with the contingency that P, that we learn that not-P? From the explanatory point of view, the suggestion that a false antecedent makes a conditional true (or rather provides a reason for calling it 'true') amounts, in effect, to the claim that any plan against an unrealized possibility is worthy of approbation.

I think that this convention would have several disadvantages. The best that can be said for it is that any such plan is bound to be harmless, in the immediate sense, because it will never be put into effect. But this is not to say that we should encourage each other to carry such policies. For one thing, even a redundant policy bears a cost. To encourage redundant plans would be to encourage unnecessary use of limited cognitive resources. For another thing, to encourage such redundant conditional commitments would be to devalue the incentives and disincentives on offer in the important cases – those in which such a commitment is not known to be redundant.

Most importantly, however, to reward conditional judgements on the grounds of their redundancy would be to undermine the function of such rewards in influencing *habits* of judgement. We noted that argument tends to improve not merely the commitments that form the immediate grounds for a disagreement, but also the general principles of inference that lead to such commitments. In the conditional case, the latter form of improvement seems to require that speakers not escape criticism simply in virtue of the fact that an inappropriately adopted conditional commitment turns out to be redundant (due to a false antecedent). Otherwise it would be as if we were to excuse (or worse still, commend) a habitual drunken driver, on the grounds that on the occasion in question he or she had been lucky enough to avoid an accident.

A similar range of considerations seems to explain why we should not regard conditional judgements with false antecedents as uniformly false. The first factor is different. The problem would be not that we would waste cognitive space by encouraging redundant commitments, but that we would underutilize the available resources by discouraging potentially useful commitments. (Why take on a life insurance policy that in addition to the ordinary premium imposes a heavy penalty for failing to die?) But the second and third factors are the same: this practice would devalue the incentives and disincentives in the more significant cases; and most importantly, would interfere with the optimal use of truth and falsity in encouraging the best habits of conditional judgement.

This seems to explain why (*pace* the material conditional interpretation) we do not argue about the truth of indicative conditionals whose antecedents we take to be false. Conversely, we can also

explain why we do disagree about pairs of conditionals that on the material conditional view might both be true. We are ordinarily inclined to treat ‘If P then Q’ and ‘If P then not-Q’ as incompatible, rather than as compatible claims that together imply that not-P (as we should if we were dealing with material conditionals). The contingency plan model of conditional judgement explains why this should be so. To take on both commitments would be to subscribe to a pair of plans that could not both be put into effect. Faced with the competing claims we therefore cannot accept both, unless in the sense that we affirm that neither will be put into effect. The disadvantage of the material conditional is that it makes this compromise the natural choice. Like the indexical treatment of probabilistic judgement, it defuses potentially valuable disputes. The appropriate course is therefore to treat the co-occurrence in a conversation of ‘If P then Q’ and ‘If P then not-Q’ as the basis of a substantial disagreement, and hence as *prima facie* grounds for a dispute. In this way we are provided with an incentive to call in our evidence, and to attempt to justify our conditional commitments.

In the process we may find that the initial disagreement rested on the fact that the two sides relied on different evidence, and that the combined evidence shows that not-P. We will thus be led to the conclusion that the material conditional interpretation would have made immediate. But we are led to it at the proper time. Properly conducted contingency planning may lead to the conclusion that a particular contingency will not arise, on the grounds that it would entail inconsistent consequences. That is, it may lead to the realization that such planning is inappropriate in the case in question. But this is a far cry from discounting a possibility whenever we find ourselves faced with incompatible plans for dealing with it.

Thus, it seems possible to explain in functional terms the most striking differences between the ordinary use of truth and falsity with respect to indicative conditionals and the pattern that would be appropriate for material conditionals. Of course, there are other factors that a more detailed treatment of indicative conditionals would need to take into account. One such concerns the connection between indicative and subjunctive conditionals. The focus of a disagreement about an indicative conditional will often shift to a connected subjunctive, particularly if doubt is raised about the truth of the antecedent. Suppose for example that we are arguing about whether if Rudolph is here, he is wearing a wig. If we learn that Rudolph is not here, we are inclined to go on to argue about whether if he had been here, he would have been wearing a wig. So a proper treatment along functional lines will need to consider subjunctives, as well as indicatives.

However, if we allow that in such a case the function of the original argument is partly to improve our habits of conditional judgement, there is an appealing explanation of the move to the subjunctive. For even if we find out that the immediate judgement is redundant (because Rudolph is not here), the more general potential benefits provide a reason for a continued dispute. As long as we take it that there is still a substantial disagreement between us, we are liable to reassess the principles on which our original judgments were based; and hence, on the whole, to do better in the future. This suggests that subjunctive conditionals might themselves be explained as tokens in a process of reasoning and argument whose significance lies at another level: in the general principles on the basis of which we form the contingency plans with which we meet the future.

We saw above that indicative conditionals give rise to a species of evaporative disagreement, in cases in which the effect of combined evidence is to convince the parties to a dispute that the conditional in question has a false antecedent. Subjunctive conditionals seem to be subject to a different species. As has often been noted, our assessments of counterfactual conditionals depend on background assumptions. In order to reason about what we take not to be the case, we must rely on some but not all of what we do take to be the case. To appeal to everything we currently accept would trivialize the exercise, since in particular we could appeal to the fact that not-P in reasoning about what would be the case if P.

How then do we select the background information to which we appeal? However we do it, it is clearly sensitive to context. As a result, a disagreement about a subjunctive conditional may turn out

to stem from the fact that the participants have ‘held fixed’ different parts of their common background evidence. Thus:

- (8.1) If the author of *Waverley* had been alive today, he would have been over two hundred years old (because Scott, the author of *Waverley*, was born in 1771);

and yet

- (8.2) If the author of *Waverley* were alive today, he or she would certainly be less than one hundred and thirty years old (because nobody lives longer than that).

Such cases suggest that a dispute about a subjunctive conditional can evaporate, if it turns out that the participants are taking for granted different aspects of an agreed body of background information. This behaviour seems consistent with the suggestion that the significance of subjunctive conditionals lies in the general beliefs on which they are based. Evaporation here seems to follow recognition that there is no disagreement about these underlying generalizations.

Subjunctives are a difficult case, however. No doubt there is much more to be said about the connections between on the one hand their functions in speech and thought, and on the other their usage in argument and as bearers of truth and falsity. I want to conclude these remarks by mentioning an advantage of the present explanatory approach that seems particularly prominent in the case of conditionals.

In recent discussions of indicative conditionals it has been noted that their interpretation is constrained by their evident affinities, in ordinary usage, with disjunctions on the one side and subjunctive conditionals on the other. (See particularly Jackson 1984, pp. 68–9; 1987, pp. 67–70.) These affinities are problematic for any approach that wants to draw a major semantic distinction between disjunctions and indicatives or between indicatives and subjunctives – for example, the view that disjunctions have ‘genuine’ truth conditions (given truth-functionally), whereas indicative and subjunctive conditionals are ‘projections’ of non-truth-conditional mental dispositions. Why should classes of utterances whose meanings are determined in such different ways behave so similarly? I want to emphasize that here, as in other areas of discourse, the explanatory approach offers a unitary approach. Hence it avoids these problems. At the same time, however, it has (or at any rate claims to have) the flexibility needed to account for the relevant range of variation we find in usage – variations in patterns of argument and in the application of the notions of truth and falsity.

8.8 PRESUPPOSITION

Conditionals have often been linked in philosophical discussion with assertions dependent on presuppositions – themselves a class of utterances whose truth-conditionality has been thought to be problematic. Much of the discussion of the latter class concentrates on the question as to whether a sentence such as

- (8.3) The king of France is the author of *Waverley*

can be considered to have a truth value, given that it depends on a false presupposition. I want to comment briefly on this topic, partly as a further illustration of the explanatory approach to such issues, and partly to exhibit a contrast to the conditional case.

From the explanatory point of view the question as to whether 8.3 ‘really’ has a truth value is misconceived, of course. We can only ask whether and why such an utterance is or should be treated as having a truth value. This aside, it may seem that this case ought to be similar to that of an

indicative conditional. Isn't discovering that a presupposition is false similar in effect to discovering that an antecedent is false?

I think there is actually an important difference between these cases. The difference between 8.3 and

(8.4) If there is a king of France, then he is the author of *Waverley*

is that a commitment to 8.4, unlike to 8.3, involves no commitment to the claim that there is a king of France. Other things being equal, acceptance of 8.3 can thus be expected to be accompanied by behavioural effects not shared by acceptance of 8.4: namely, those effects that would be expected to flow from the belief that there is a king of France. These effects give a point to criticism on the grounds of the failure of a presupposition, not shared by the conditional case.

This is not to say that in virtue of having such effects 8.3 *is* false when there is no king of France. It is simply to say that in virtue of these effects, the practice of describing such an utterance as 'false' in such circumstances is likely to serve a useful function. Perhaps the same function could be performed in other ways. For example, we might have a separate mode of criticism, reserved for this kind of mistake. The intuition that some philosophers have had that failure of a presupposition renders an utterance neither true nor false is perhaps evidence for the existence in usage of some degree of separation. Certainly it suggests that it would be within our capabilities to sustain such a distinction. But this is not a concession to the analytic standpoint – the basis for an account of 'real' as opposed to 'constructed' truth. For why stop here? Why not discern a separate kind of truth for every subject matter and every critical nuance? We saw in part I that what one needs to worry about, in leaping aboard the projectivist bandwagon, is not the engine but the brakes. Getting it started is easy enough. The problem is to pull the thing up where the projectivist wants to stop.

The question arises as to whether our explanatory view is simply projectivism unconstrained – the idealism that we saw to be the projectivist's most threatening danger. I shall return to this in the concluding chapter. I want to turn now to some preliminary questions. How widespread are evaporative disputes? Does the explanatory approach allow us to draw a distinction between those utterances that admit such disputes and those that do not? And if so, is there anything to be said for the idea that the latter are the genuine statements of fact?

8.9 FACTUALITY: A MATTER OF DEGREE?

Evaporative disputes are a boundary phenomenon, a manifestation of the felt inappropriateness of a continued pretence that a particular disagreement is answerable to an external standard of correctness and incorrectness. The cases we have so far considered suggest that the phenomenon is a product of the local peculiarities of particular types of utterance. Hence we might expect that in a broad range of central cases, it would be unknown. This expectation might in turn suggest a characterization of the 'genuine' statements of fact: namely, that they are those indicative utterances not prone to evaporative dispute.

Interpreted correctly, this characterization would be unobjectionable. If there is a distinction in usage between those indicative utterances that do and those that do not give rise to evaporative disputes, then by all means let us mark it. But this unobjectionable interpretation is not the source of the suggestion's appeal. This lies rather in the misconceived intuition that the suggested characterization gives us the long-sought analysis of statement of fact – a venerable confusion of taxonomy for analysis.

That aside, it seems to me that the assumption concerning usage that underlies this suggestion is significantly mistaken. There seems to be at least one potentially global source of evaporative disputes. Hence it is doubtful whether there is any class of indicative utterances conventionally immune from the patterns of usage that supposedly distinguish 'non-genuine' statements of fact. This

global source is a form of indeterminacy of meaning. In section 8.2 I mentioned the possibility that apparent disagreements may turn out to rest on the fact that the speakers concerned are using the same term with different meanings. This possibility threatened to trivialize the issue about relativism. To defend a space for an interesting relativist position, we suggested that the motivational character of moral judgement provides a fixed component in the meaning of moral terms, around which the relativist issue can be built. But the problems raised by meaning variance are not exhausted by its abortive role in this case.

Meaning variance can affect disputes in several ways, some of them relatively uninteresting. Perhaps the duller involves a term that is known to be ambiguous in ordinary usage, whose multiple meanings find their way to the heart of an argument – for example as to whether in the modern computerized hospital the doctors should be expected to grasp the rudiments of ‘terminal care’. Marginally more interesting is the case in which a dispute stems from the fact that one participant has simply misunderstood the meaning of a central term. This raises an issue as to whether (and if so, why) we should say, when the mistake becomes apparent, that this person’s original assertion was false. This issue seems to have something in common with the case of false presupposition. Indeed, it could be argued to be a sub-case of that of presupposition, on the grounds that assumptions about meaning can be treated as presuppositions. Alternatively – I think it is not the same point – it could be said that we should not expect, post-Quine, a sharp distinction between mistakes about meanings and mistakes about facts.

I want to concentrate on a third kind of case, however. It does not depend on ambiguity, in the normal sense, or on any mistake about meaning. It stems rather from a form of indeterminacy about meaning that seems endemic to language. In recent years philosophers have paid a lot of attention to the fact that language is finitely acquired. Each of us has acquired her or his linguistic knowledge in the few short years since birth – in so far, at any rate, as we did not have it at birth. This observation has been applied in a number of ways: by Chomsky and his followers, in arguing for the existence of innate linguistic ability; by Dummett, in his argument for an anti-realist conception of truth; and in the literature stemming from Wittgenstein’s treatment of the notion of following a rule. And a closely similar factor, concerning the finiteness of the evidence available to the radical interpreter, is crucial to Quine’s indeterminacy thesis, developed in *Word and Object*.

For our purposes, the significant point is that no finite level of experience can determine the application of a linguistic term to all possible cases, or exclude the possibility that two speakers – fully competent speakers, by existing community standards – will diverge in their application of the term concerned to some future case. It is true that this does not happen very often. The explanation of this fact, particularly in the light of finite acquisition, is one of the central problems in the philosophy of language. (Wittgenstein’s treatment of rule following in the *Philosophical Investigations* can be seen as an attack on one conception of what a solution would look like. Cf. McGinn, 1984, chapter 1.) Given however that it can happen, and cannot be excluded in advance, it seems that we live with the possibility that disagreements may turn out to stem from the fact that participants have made different (though equally consistent) ‘projections’ of a term from a common finite base.

What happens in cases such as this? One common if often abused response is to say that the disputed point is ‘terminological’. I think the important thing is that we recognize that neither party is *mistaken*, in either of two ways: in using the term concerned with what is the wrong meaning, by previously established standards; or in a sense that would lead us to say that they had spoken *falsely*, given their own reading of the equivocal term. The dispute thus evaporates. Neither side concedes, and yet each recognizes that there is no longer a point in insisting that the case is answerable to an external standard, that only one side can meet.

These cases reflect the intuitively appealing idea that our judgements have truth values only in so far as they are determinate in meaning. An ambiguous judgement may have a determinate truth value by default, as it were, if all its possible interpretations share the same truth value. But there is no

guarantee that this will be so, in general – ambiguities tend to show up when it is not so. And in any case, it is far from clear how to apply this notion when we are not dealing with an ambiguity that is finitely specifiable in terms of non-ambiguous atoms. To recognize ineliminable indeterminacy of meaning seems to be to recognize that our utterances are at best approximations to genuine (and therefore strictly truth-evaluable) statements of fact.

I hope it is no longer necessary to emphasize that the present approach does not claim to show that such intuitions are correct. We have offered no analysis that would afford such a judgement. Here, as elsewhere, our point is rather to observe that ordinary intuitions concerning the limits of factuality can be correlated with a certain pattern of usage – the existence of evaporative disputes – and to argue that this pattern can be explained in terms of a general account of truth, falsity and the function of argument.

The indeterminacy of meaning thus seems to be the basis for a form of universal non-factualism. That is, it offers a reason for departing from the pattern of usage characteristic of the statement form that may be applicable in any area of discourse. This topic-neutral form of non-factualism should not be confused with the form that Quine himself extracts from the indeterminacy thesis: the view that there are no genuine facts about *meaning*. (Cf. Quine 1960, p. 221.) Ignoring for a moment our scruples about the analytic viewpoint, we might say that the present view, in contrast, is that in virtue of the nature of meaning there can be no genuine facts about anything.

More correctly, the result suggests that the factuality of usage is always a matter of degree. Across the broadly fact-stating part of language – in English, roughly, the indicatives – the appropriateness of the statement form varies from case to case, both type to type and token to token. From the explanatory point of view there is nothing surprising about this conclusion. Only the analytic presumption leads us to expect a sharp distinction.

8.10 THE SOURCES OF SUBJECTIVITY

What then can be said in general about the sources of evaporative disagreement, and about its significance in other areas in which non-factualism has seemed attractive? A survey and prospectus seem in order.

Disputes evaporate when disputees decide that in virtue of the peculiarities of their own standpoints it is not appropriate for each to insist that the other is mistaken. Evaporation is a manifestation of a degree of felt subjectivity in judgement. The cases discussed above illustrate the range of potential sources of such subjectivity. On the one hand are cases in which the source lies in the context or conversational role of the judgements concerned. Probabilistic and conditional judgements seemed to be of this kind; their tendency to yield insubstantial disagreements being attributable to their distinctive roles in discourse, and hence their sensitivity to the evidential context in which they are used. On the other hand are cases in which the source of subjectivity seems to lie more in some distinctive feature of the speakers themselves – in the idiosyncrasies of their grasp of the meaning of a relevant term, for example. (A more familiar case of this kind might be the subjectivity of judgements about secondary qualities; more on this in a moment.)

The distinction between intrinsic and extrinsic sources of subjectivity is far from sharp. The sources of moral or aesthetic relativity seem to lie somewhere between the two extremes, for example. However, I think it provides sufficient structure to enable us to locate the likely basis of evaporation and hence non-factualism itself in some of the other areas in which such an approach has seemed attractive.

On the contextual side, the modal operators are the obvious candidates. The epistemological approach to possibility and necessity can be seen as an attempt to clarify the dependence of modal judgement on evidential context; even if usually under the misguided assumption that what is called for is an

analysis of modal judgement – an account of its truth conditions, or some such. Another possible candidate is causal judgement. Here, particularly, the issue is clouded by the appeal of Humean metaphysics – a route to non-factualism whose viability we disputed in part I. Metaphysics aside, however, non-factualism about causality has an appeal that rests on contextual and conversational factors. This is not the place to explore these factors in detail, but I mention two. In view of the connections between causality and probability – particularly the intuition that causes increase the probability of their effects – disagreements about single case causes seem able to evaporate when the same happens to connected judgements about probability. And the well-recognized association between causal judgement and counterfactual conditionals suggests that the sensitivity of the latter to background assumptions – about ‘the normal course of events’, for example – will also give rise to insubstantial disagreements about causality.

At the other end of the scale, I have already mentioned judgements about secondary qualities. It has long been felt that these exhibit a degree of subjectivity, attributable to their dependence on contingent features of our perceptual apparatus. The view is supported by cases of major variation between human subjects in sensory response to a given stimulus. A substance may taste bitter to some subjects yet be tasteless to others. It is interesting that our willingness to treat such a disagreement as insubstantial depends on the relative frequency of the different responses. We are inclined to say that the aberrant judgements of colour-blind subjects are indeed mistaken – that it is a *fact* that jacarandas are mauve rather than blue. But who would insist that chemical X is really bitter, when about 50 per cent of people find it tasteless? This phenomenon is easily explained from the functional viewpoint, of course; but again, it is troublesome for an account that expects a sharp underlying distinction.

What else lies at this end of the scale? I mentioned the possibility of idiosyncratic grasp of meaning – a potentially universal source of evaporative argument. My impression here is that among the determinants of meaning there might be some we would count as strongly individual – perhaps innate dispositions to categorize experience in certain ways – and others more dependent on factors external to the individual speaker, such as communal systems of convention. The latter possibility suggests a place in the scheme for another topic about which non-factualism has been popular, namely theoretical discourse. The recent popularity of scientific anti-realism stems largely from an appreciation of the difficulties of inter-theoretical discourse. Once it is recognized that proponents of different scientific theories may lack a basis for argument, it has seemed difficult to hold onto the idea that each is really making claims about the same objective world. The present view suggests that the source of the problem is an unacknowledged commitment to the notion that there is such a thing as genuinely fact-stating discourse. If we reject this notion, we are free to say that the inter-theoretical case is simply a prominent form of a phenomenon endemic to language. Theoretical discourse is not a special case; it differs only in degree from discourse of any kind. (The anti-realist might agree, of course, but interpret the result as global anti-realism. More on this in chapter 9.)

How much of non-factualism can be explained in these terms? It is difficult to make any firm pronouncement – the field is quite large, after all – but the prospects seem to me to be good. If there is a major area of uncertainty it seems to lie in the cases in which non-factualism is motivated by a combination of metaphysical and epistemological considerations. I argued in part I that a determined factualist should be untroubled by these factors. The appropriate conclusion might therefore be that in so far as it rests on such intuitions, non-factualism is simply mistaken, and should not be expected to conform to the pattern we have identified. However, there are several less dismissive interpretations of an appeal to such factors.

The first involves the interpretation of a disputed non-factualist thesis simply as a first order claim about what there is in the world, and not also as a second order claim about the nature of discourse. The practical difference here is that the weaker thesis, unlike a typical non-factualism, will recommend that we abandon the form of discourse in question. The force of the argument of chapters 4 and 5 was that if we are to retain an area of linguistic practice, the non-factualist cannot dislodge a determined defence of a factualist interpretation of that practice. This leaves it open for someone to

argue that we should discard the practice, however interpreted. Some of the appeal to metaphysics and epistemology in these contexts might be taken to be of this kind.

Secondly, it is easy to imagine that someone might appeal to metaphysical and epistemological considerations in order to establish the relatively theoretical nature of a particular area of discourse. To give a crude example, it might be said that since we don't directly perceive probabilities, probability must be a theoretical notion. Moves of this kind bring the discourse in question within the scope of a general claim about the status of theoretical language. We have seen that theoretical non-factualism can be brought within the pattern suggested by the explanatory approach. This interpretation of the appeal to metaphysics and epistemology thus brings it within the scope of our general account of non-factualism.

The last possibility is in some ways similar. There is clearly a role for metaphysical conceptual analysis in an approach of this general kind. We have already noted an example: an analysis of causality in terms of probability, if sustained, would be likely to ensure that causal judgement was similarly liable to evaporative disagreement. (This might cut both ways: the lack of such disagreement in the causal case might count against such a conceptual analysis.)

These possibilities suggest that there is a significant place for metaphysical and epistemological concerns in an explanatory approach to non-factualist concerns. But I must leave it to the reader to judge to what extent existing appeals to these factors can be interpreted in any of these ways – as, of course, to assess the approach to non-factualism as a whole.

8.11 THE SIGNIFICANCE OF MOOD

What then of the non-indicatives? Here, more than anywhere, usage may seem to call for the analyst's sharp dichotomy. Hence, as we saw, the appeal of the marker view. But here too, I think the call should be resisted. It can be countered by the strategy of indicative paraphrase, which was the basis of our argument against the marker view in chapter 3. And we can now add to that argument a possible explanation for the existence of the non-indicative moods, in terms of our account of the function of truth-evaluability.

If a non-indicative is paraphrased in the way we suggested in chapter 3, it is normally true in virtue of being uttered. At any rate, we saw that the suggested indicative paraphrase is such that its truth is normally in the gift of a speaker who utters it. In so far as non-indicatives are informative, the information they carry is not normally open to dispute. This suggests that if the linguistic functions served by the non-indicatives were to be assigned to indicatives, the argument-encouraging machinery associated with the indicative mood would *usually* be redundant in these cases. The qualification is important: first because as we'll see in a moment, such machinery would not always be redundant; but more importantly because the evidence suggests that limited redundancy is often tolerated in the interests of the general benefits of the fact-stating form. After all, there are many ordinary indicatives whose truth is similarly guaranteed by utterance – 'I have not lost the power of speech', for example.

To explain why more frequent redundancy should be disadvantageous, I think we need to mention another factor: namely, that unused machinery may be a positive liability. For one thing, the argumentative mode of speech is presumably sustained at some conceptual cost. It requires that speakers hold themselves ready to assess the utterances of others by the standards reflected in judgements of truth and falsity. Where such assessment is generally inappropriate, or predetermined by the nature of the utterance, the general conceptual cost may not be worth the very occasional benefit. For another thing, the possibility of argument might itself be disadvantageous. If non-indicatives serve linguistic functions that call for specific and often immediate responses from an audience – answering a question, meeting a request, and so on – then it might well be counter-productive to use a form of language that encourages a different kind of response: namely, an argumentative challenge to the 'correctness' of the utterance in question. This is not to say that one

cannot challenge a non-indicative utterance. Clearly one can, and we do – amongst other things on the kinds of grounds that would constitute a challenge to the truth of the kind of indicative paraphrases suggested in chapter 3. Ordered to complete a form, for example, we may object that we are not by law required to do so. The point is simply that there may be a social advantage in not actively encouraging such challenges.

These factors, taken together with the inevitable element of historical accident, are perhaps as much as we can expect of an explanation of the existence of the non-indicative moods. The striking feature of these moods is the fact that they are not taken to be truth-evaluable, in ordinary usage. The above factors bring this fact within the scope of our explanatory account of the role of truth and falsity. It would seem that the non-indicative moods occupy a region of language in which, on the whole, the disadvantages of the notion of truth outweigh its advantages. But the boundaries of this region are necessarily imprecise, and the sharpness of the mood distinction, if taken at face value, is liable to be misleading. The linguistic functions it masks embody no corresponding discontinuity, but rather a complex and multi-dimensional field of factors, varying by utterance token as well as by utterance type.

An artificial order is imposed on this field by the need for regularity in language. The rules of language cannot be sensitive to every nuance of the individual case. But we should not underestimate the considerable flexibility that does exist in natural language. In practice the mood distinction is far from sharp. As we have seen, English speakers will switch between indicative and non-indicative utterances for what are recognizably similar linguistic tasks – according, at least in part, to the perceived role for an idea of an objective standard. And within the indicative mood, we have observed the striking tendency of ordinary speakers to recognize the limits of truth and falsity, to opt for tolerance, and to allow inappropriate disputes to evaporate. Usage is very much less sharply polarized than first appearances might suggest. It is therefore a point in favour of the present explanatory approach that it does not advocate or assume a sharp distinction between fact-stating and non-fact-stating uses of language.

9

End of Enquiry

We began with the conflict between non-factualism and Fregean linguistic theory – a somewhat ironic tension, we noted, given that non-factualism was a distinctive and popular product of the mid-century linguistic movement in philosophy. In this concluding chapter I want to begin by showing how this conflict seems to be resolved by the present explanatory approach to truth and factuality. This will lead us to the question as to why, on the present account, the notion of truth should be the kind of notion that it is. We encountered this question in chapter 7, and there and elsewhere have touched on the main aspects of the answer. Here I shall try to draw together the threads. As foreshadowed in chapter 2, David Wiggins's 'marks of truth' provide a useful focus at this point. We want to confirm that the proposed functional account can explain why truth should have such characteristics. The third concern of the chapter will then be to clarify some of the connections and advantages of the functional approach, with respect to projectivism, quietism and other positions in contemporary debates about realism and objectivity. We'll finish with some comments on the scope of the functional perspective in other areas, including particularly its application to the problem of the status of folk psychology.

9.1 TRUTH: A CONFLICT RESOLVED

In chapter 2 we looked at Wiggins's attempt to ground an account of assertion on a substantial notion of truth, itself derived from a Davidsonian programme for radical interpretation. Our main criticism was that the programme of radical interpretation turns out to be insensitive to the distinction between fact-stating and non-fact-stating parts of language. The theorems of the radical interpreter's so-called truth theory do not have to mention truth at all; and if they do mention it, need only the disquotational property – the equivalence principle – in terms of which the truth predicate can be extended trivially to utterances of any kind.

The impression that Fregean general semantics requires a substantial notion of truth thus rests on an uncritical acceptance of the doctrine of the primacy of assertion. If we equate *saying* with *asserting*, then of course we equate the disquotational predicate needed in a Fregean theory of meaning with the ordinary predicate '... is true'. However, if we keep in mind that there may be sayings that are not assertings then such an equation looks highly implausible – the ordinary predicate does not extend to non-assertoric sayings. Hence the needs of radical interpretation, or of a systematic theory of meaning of the kind it requires, seem to have little bearing on the question as to why the ordinary notion of truth should be restricted in this way.

I think that the resolution of the conflict between non-factualism and the Fregean programme for general semantics thus turns on a distinction between the disquotational and the normative properties of truth. Contemporary semantic theories exploit the disquotational property. The non-factualist, however, needs to be concerned with the normative property; for only the normative property seems able to explain the limited application of the ordinary notion of truth, and hence to ensure that truth has something to do with a distinction of language into 'fact-stating' and 'non-fact-stating' parts.

A Fregean might object that one thing in favour of a truth-conditional approach to a theory of meaning is the intuition that to know the meaning of a sentence is to know its correct use – to know under what conditions it would be correct to utter that sentence (with serious intent). This intuition surely depends on the normative notion of *correctness*. So to the extent that it calls for a truth-conditional theory of meaning, it seems to require a normative notion of truth, rather than merely a disquotational notion.

However, I think this objection trades on an ambiguity in the notion of correct use. In the case of assertoric discourse this ambiguity is well recognized, if not always well heeded. The term ‘assertibility condition’ (or ‘condition for correct assertion’) can mean one of two things. Taken subjectively, it refers to something like the state of mind that would normally be considered to be evidenced by a competent speaker’s utterance of the sentence in question. According to the conventional model, a belief that P is thus the (subjective) assertibility condition for the sentence P; and there seems to be something to be said for the claim that to know the meaning of P is to know that this is the case. Taken objectively, on the other hand, ‘assertion condition’ refers to some external state of the world, whose obtaining would normally ‘justify’ the assertoric use of the sentence in question. This reading itself admits of various interpretations; for present purposes the important thing is that at least one interpretation has the effect of equating ‘correctness’ with ‘truth’. And again, there seems to be something to be said for the claim that to know the meaning of an assertoric sentence P is to know when its use would be correct in this sense – in other words, to know its truth conditions.

As long as we restrict ourselves to assertoric discourse, there is little incentive to keep in mind that there are these two readings of the principle that to know meaning is to know correct use. To the extent that they are distinguished, the objective notion has seemed of greater interest. But if we broaden our field of view to include the non-assertoric part of language, I think we should regard this as another case of taking for granted what we ought to explain. The availability of the subjective reading shows that to decline to rely on the objective reading is not to reject the principle itself. And in extending to non-assertoric discourse, the subjective reading has the generality that the objective reading lacks. In practice, indeed, the subjective notion of an assertibility condition has tended to receive attention in cases in which the notion of truth seems problematic, such as that of indicative conditionals.

We can thus aim initially at a theory of meaning that does not depend on a radical distinction between the assertoric and non-assertoric parts of language (regarding this rather as something that calls for explanation). Apart from its elegance, generality and explanatory potential, this approach has the advantage we discovered in chapter 2: it turns out to be a product of its opponents’ best efforts at following the alternative approach, given that truth theories do not in fact rely on the normative aspect of truth.

9.2 WHY IS TRUTH TRUTH-LIKE ?

This separation of the disquotational and normative aspects of truth may itself seem problematic. Why do these aspects go together? How is it that the ordinary normative truth predicate functions disquotationally? These are questions that an explanatory theory of truth should be expected to answer.

More generally, an account of truth on the proposed lines should be able to relate the features of truth as we find it to the functional role that is being claimed to explain its development in language. True, there is some flexibility in an evolutionary account of this kind. Some features of the ordinary notion may be more or less accidental. But the plausibility of any such claim seems to rest on the relative insignificance of the feature in question. The more a property seems essential to truth, the more we should expect an account of why truth should have such a property. In particular, we should expect an account of the marks of the concept of truth, in Wiggins’s sense (1980, p. 194).

In the case of the connection between the disquotational and normative aspects of truth I think the answer is fairly straightforward. To explain the disquotational property we have to explain why P and ‘P is true’ can in some sense be regarded as equivalent assertions – for example, that they can be taken to be expressions of the same commitment. On the present theory, to say that P is true is to allow to judgements that P a conventionally-generated property that we are all conditioned to want our judgements to have. Roughly, it is to ascribe worth to judgements that P. The point of this conventional system of value is to encourage argument. In making our case for these conventional

rewards, we indulge in a process whose effect tends to be to improve our behavioural commitments. But this improvement depends on the fact that (in the main, at any rate) argument is prompted by some mismatch between individual behavioural commitments, and resolved by a resolution of such a mismatch. So the process requires that we ascribe value to those judgements we are prepared to make ourselves. Only in this way can the ascription of these conventional rewards make socially unstable the appropriate cases of mismatched commitment. In other words, the role of truth in argument requires the disquotational property – it requires that we be prepared to acknowledge that P is true in the same circumstances as we acknowledge that P.

This answer bears on the problems that might be thought to flow from the fact that ascriptions of truth and falsity themselves take the grammatical form of indicative sentences, and are thus the sorts of expressions to which ‘true’ and ‘false’ can be ascribed. We touched on this point in chapter 8, mentioning in particular the objection that it might commit us to an endless hierarchy of disagreements, corresponding to iterations of the predicates ‘... is true’ and ‘... is false’. In explaining how we avoid embarking on such an escalation of ordinary disagreements, the disquotational property is part of the answer. It means, I think, that from the point of view of our behavioural commitments there is no reason to escalate. Behaviourally speaking, semantic ascent doesn’t increase the stakes.

The other part of the answer, I think, is that in the case of disagreements about truth and falsity themselves there isn’t the normal reason to *introduce* these normative terms. Truth comes into the picture to provide the incentive for argument. But once it is already in the picture – when we are disagreeing about the truth of some sentence – there is nothing to be gained by bringing it in again, by moving to second order truth ascriptions. There’s a prize for winning the argument, but no further prize for winning the prize. This is not to say that ascriptions of truth and falsity are not themselves regarded as answerable to the kind of objective standard the concept of which their use embodies. On the contrary, as we stressed in chapter 8, it is crucial to the motivating role of such evaluations (as of others) that they be regarded as objective – that there be no easy subjectivist escape hatch. It is simply to say that in the case of truth and falsity themselves, the manifestations of this felt answerability to an external standard do not have their usual importance in argument.

This brings us to Wiggins’s marks of truth. We are now in a position to explain why truth should have the characteristics that Wiggins identifies – or rather, as we’ll see, why it should have those four of the five such characteristics that are properly regarded as properties of truth.

Wiggins’s first mark is the normativity of truth. It consists in truth’s being ‘the primary dimension of assessment for sentences’ – ‘that property which sentences have *normally* to be construed as aiming to enjoy’ (1980, p. 205). This property is central, of course, to our proposed account of the function of truth in discourse. Unless we take truth as something to be aimed at, we will be given no incentive to argue by the claim that one of our utterances lacks it.

The second mark of truth concerns its attainability. Wiggins puts it as follows.

A second mark of the [truth] of a sentence *s* would appear to consist in *s*’s *having a content such that with respect to this content there should be a tendency, under favourable conditions of investigation, for disagreement to diminish and for opinion to converge in agreement.*
(1980, p. 206)

Here the relevant point would seem to be that unless speakers subscribe to some such view of truth then even its normativity will not encourage them to argue. Heroics aside, why should we argue with someone unless initially there seems some prospect that he can be brought to share our own opinion – in other words, some prospect of agreement? A degree of attainability thus seems a necessary part of our idea of truth, if truth is to serve the role we have claimed for it in encouraging argument.

Similar remarks apply to Wiggins's third and fourth marks of truth. The third mark (1980, p. 208) is that truth be considered independent of a speaker's means of recognizing whether a statement possesses it; while the fourth (p. 211) is that true sentences be considered to be true 'in virtue of something'. These would appear to be aspects of the required objectivity of truth, on which we commented a moment ago. The power of truth as a system of incentive depends on the fact that it is not taken to be able to be bestowed at will. This is what distinguishes the incentive provided by truth from the incentive provided by a system of boos and cheers. Rival views may share the available applause, but they cannot share the truth – or so, at any rate, it is useful for us to believe.

So to Wiggins's fifth claimed mark of truth (1980, p. 211): if two sentences are true then so is their conjunction. It is difficult to take this as a property of truth, rather than as an account of conjunction in terms of truth. There is a legitimate question as to why we should have a sentence-forming operator *and* such that we are prepared to assert *P and Q* when and only when we are prepared to assert both *P* and *Q* individually. Given however that we do have such an operator, the fifth mark follows directly from the disquotational property – from the fact that we are prepared to assert that *P* when and only when we are prepared to assert that *P* is true.

9.3 THE TRUTH OF TRUTH ASCRIPTIONS: SOME OBJECTIONS EVADED

I mentioned above the treatment of truth ascriptions themselves by the present explanatory account of truth. I want to draw attention to some problems that trouble analytic non-factualists at this point – problems that again the explanatory approach avoids.

For one thing, the analytic point of view leaves a non-factualist with the following dilemma. Either 'P is true' is sometimes a statement and sometimes not a statement, depending on the status of *P* itself; or all such truth ascriptions have the same status. Neither choice is particularly appealing. The former entails that truth ascriptions are not the semantically homogeneous class their form suggests. While the latter implies that 'P is true' can be a genuine statement even though *P* itself is not a statement, or *vice versa*; and hence that the equivalence principle fails. From the explanatory perspective, however, the dilemma looks entirely self-inflicted – a product of the assumption that any utterance lies on one side or other of a real semantic divide.

Arthur Fine has recently raised a related objection to a treatment of truth on projectivist lines:

The [projectivist] strategy requires a thin, unproblematic discourse on which to base its projections. To project out 'truth', the thin discourse must be truth free. But where are we to find any such thin discourse? The redundancy property of truth makes truth a part of discourse that merits the name. In particular, if we go for acceptance or coherence theories of truth we try to build truth from certain systems of judgements (or the like). But the judgement *that P*, and the judgement *it is true that P* are the same. You cannot have one without the other. That is to say, you cannot include one in the system without thereby including the other Thus there is no thin discourse from which the [projectivist] explanations of truth talk could begin. (1986, pp. 169-70)

This objection is like the above one in squeezing a non-factualist account of truth between on the one side the desire for a unified theory of truth, and on the other the equivalence principle (or redundancy property). If the projectivist is prepared to concede unification then there is again an easy way out: to offer separate accounts of 'thin' truth and 'thick' truth – separate interpretations of truth ascriptions to utterances that are themselves about, respectively, 'genuine' or 'projected' facts. (This seems to be the approach taken by Blackburn 1984, though not without reservations; see particularly pp. 256–7.)

However, I think that in Fine's case, as in the earlier one, the explanatory approach needs no such reply. Because it is committed to no prior bifurcation in language, it cannot be turned on itself in the way that an analytic theory is bound to find embarrassing. True, its functional account of the role of

truth, and fact-stating discourse in general, will itself be couched in fact-stating terms. As its proponents, we will be inclined to assert its *truth*. But there is nothing more vicious in this than there ever is in speaking or writing about the origins of language. If there were no more to Fine's objection than the observation that the projectivist must use 'true' (in effect, at any rate) in offering a constructive theory of truth, the point would be trivial and irrelevant. Its force therefore rests on the consequences of bifurcation, to which the projectivist is committed but our theory is not.

9.4 FUNCTIONAL REALISM: NOT IDEALISM ...

What then of the suggestion that the present theory is simply idealism? As we saw in chapter 4, the idealist takes over the projectivist's goal of accounting for the realist features of linguistic practice but aims to do it quite generally. Our approach thus concurs with idealism in rejecting the projectivist's bifurcation between real and projected facts, arguing that there is a single appropriate account of truth, applicable wherever truth is. Given that the idealist takes over the projectivist's constructive programme, we also agree on an explanatory as opposed to an analytic approach to truth, and perhaps on some of the details.

The main difference, I think, concerns the treatment of non-factualist claims. The projectivist aims for 'quasi-realism', as Blackburn calls it – for a demonstration on projectivist lines that we are entitled to talk *just as if* there were genuine facts of the disputed kinds. Blackburn's quasi-realist seems not to countenance differences in usage between the disputed utterances and genuine statements, aiming instead for the perfect pretence. So when the idealist challenges the distinction between pretence and reality, non-factualism, having been lashed to quasi-realism, follows it down the slope.

The present theory, on the other hand, goes out of its way to draw attention to some characteristic peculiarities of the kinds of discourse with which non-factualists have been concerned. It tries to associate these peculiarities with particular variations in the factors that seem normally to account for the functional advantages of truth and related features of usage. In the process, as well as demonstrating its ability to explain these cases, it offers some foundation for the intuitions that have usually motivated non-factualists. True, this usage-based descriptive–explanatory programme is not what non-factualists have usually had in mind. But unlike Blackburn's quasi-realist projectivism, it does offer a refuge for the non-factualist against the push to generalize the constructive or explanatory approach to truth.

In summary, I think the present approach concurs with the idealist in seeking to globalize the projectivist's constructive approach to certain areas of reality – the attempt to ground our conception of the world 'internally', so to speak, in our own patterns of reaction to our environment. Unlike the idealist, however, we find room in this for a kind of non-factualism. This aside, the term 'idealism' has associations the present account will want to avoid. It is apt in particular to suggest an analytic approach to truth, in terms of coherence, or some such – an approach quite different from ours and from that of the projectivist's opponent. The present theory is best described as *functional* or *explanatory* realism. Its goal is to explain our realist practices; as well, of course, as their non-factualist variations.

In expressing even this much sympathy for idealism I am likely to be accused of denying common sense intuitions about the reality of the external world. So I want to make it clear that my views about most aspects of the external world are the same as those of the most hard-nosed realist. We each hold true much the same class of propositions about how things are. Our professional differences concern at most a tiny fraction of such propositions: those that deal with the explanation of the fact that humans hold such views. For my part, I believe that at least in so far as holding a view about the world depends on our possession of the notion of truth or correct representation, the fact that we do so is explicable in the kind of terms here suggested. Hard-nosed realists may have different views on this topic, or none at all. But this doesn't prevent us agreeing about almost everything else (in other words, about almost everything).

In particular, it would be a mistake to think that because I have a biological explanation of the fact that we apply a notion of correctness to judgement I am unable to apply such a notion myself. That would be as absurd as supposing that a biologist who explains hunger in evolutionary terms should no longer feel the need to eat. Indeed, it would perhaps be more absurd, for whereas to hold a theory about hunger is not to feel hungry, a belief about the origins of truth is an exercise of the very biological capacity it itself concerns. So it demonstrates its author's continued ability to exercise this capacity.

9.5 ... NOT INSTRUMENTALISM ...

I should perhaps emphasize also that functional realism is not a form of instrumentalism. Indeed, it is likely to reject instrumentalism on the same grounds as it rejects projectivism. Both these approaches ask us to make a distinction between two kinds of superficially factual discourse: the genuinely factual areas, on the one hand, and the instrumental or projected areas, on the other. The present theory denies that we can draw such a distinction. The best it can offer an instrumentalist is a place in the ranks of the reconstructed non-factualists; and this only if the instrumentalist can point to the relevant sort of difference in usage between the supposed instrumental discourse (theoretical judgements in science, or whatever) and the canonical assertoric form.

All the same, I think there is a nice analogy between the instrumentalist's most powerful argument against orthodox scientific realism and an important argument for the present explanatory approach to truth. The instrumentalist's argument is described by Fine, in the paper from which I quoted above. Fine notes that

the realist must at least allow that, generally speaking, truth does lead to reliability. For instance, ... the realist offers the truth of a theoretical story in order to explain its success at a certain range of tasks. If this offering is any good at all, the realist must then allow for some intermediate connection between the truth of the theory and success in its practice. The intermediary here is precisely the pragmatist's reliability. Let us therefore replace the realist's 'truth' with the pragmatic conception, framed appropriately in terms of reliability. Then, if the realist has given a good explanatory account to begin with, we get from this pragmatic substitution a good instrumentalist account of the same phenomena. Indeed, the instrumentalist account we obtain in this way is precisely the intermediate account that the realist has himself been using. Since no further work is done by ascending ... to the realist's 'truth', the instrumental explanation has to be counted as better than the realist one. (1986, p. 154)

The instrumentalist thus accuses the realist of a redundant metatheoretical commitment.

An analogous argument can be drawn from our account of the role of truth in encouraging argument. We saw in chapter 7 that if one begins with a conception of truth as a property of belief, and asks what the point of argument could be – why we should care whether our beliefs have this property – then at first one may think of truth as an end in itself. It becomes apparent, however, that truth could at best be an intermediate goal in one's quest for something like behavioural success. On the other hand, we have also seen that progress towards behavioural success seems to require, or at least to be enhanced by, our all *behaving* as if there were such an end as truth to judgement and enquiry. If there were no real property of truth, we would be well advised to invent one. In other words, the real or substantial notion of truth now looks redundant – unnecessary either to account for the utility of argument, or to explain our *use* of a notion of truth.

This parallel is hardly surprising. After all, the redundancy argument is the hallmark of the claim to conceptual economy, itself characteristic of the general philosophical movement to which instrumentalism and the present theory both belong. Such arguments involve some characteristic risks, however. As the quasi-realist's problems illustrate, they tend to backfire, undermining the very

positions their advocates wish to preserve. Economy needs to be backed up by stability: the thin residue must be sturdy enough to stand on its own.

Fine argues, I think rightly, that instrumentalism cannot stand on its own. As he puts it, 'instrumentalism avoids the inflationary metaphysics of realism only by pumping up either epistemology or semantics' (1986, p. 170). For the semantic case, his argument is the objection to a projectivist theory of truth we noted earlier. I think that like ours, his basic objection to instrumentalism is that it involves an untenable distinction between 'genuine' and 'instrumental' areas of factual discourse.

9.6 ... AND NOT QUIETISM

In the face of these objections, however – on the one side to realism, and on the other to instrumentalism – Fine recommends that we be satisfied with what he calls 'NOA ..., the *natural ontological attitude*' (1986, p. 172). He says that

NOA allies itself with what Blackburn ... dismissively calls 'quietism'. Less pejoratively, NOA holds a 'no-theory' conception of truth NOA does not think that truth is an explanatory concept, or that there is some general thing that makes truths true. (1986, p. 175)

Fine also links NOA to Rorty's recent views on truth (on which we remarked in chapter 6), saying that Rorty's 'developing neo-pragmatism ... seems quite NOAish' (1986, p. 173, n. 16).

My use of the term 'quietism' has not been exactly Blackburn's. I used it initially for the view that declines to theorize about the fact-stating/non-fact-stating distinction, rather than for the analogous view about truth. But clearly the two views run parallel, and my sympathies are with Blackburn.

NOA is right, I think, in rejecting attempts to *analyse* truth. And it is right to dismiss recent attempts to make truth itself an explanatory notion. (We considered these attempts in chapter 6. I argued that far from supporting a realist or correspondence notion of truth, they depend on no more than the disquotational or redundancy property.) NOA is wrong, however, to conclude that there is no further role for a theory of truth. At the very least there remains the not inconsiderable task of accounting for the fact that we have such a notion at all.

The quietist is simply blind to this, a blindness perhaps explained by the tendency on all sides in these discussions to ignore non-descriptive uses of language. From this perspective we are liable to overlook the contrasts in language that draw attention to the remaining explanatory task. To notice the non-indicatives, and the issues raised by the non-factualists, is to notice not only that truth is not a universal notion in ordinary language, but also that there is considerable doubt as to the limits of its application. At this point, I think, to decline to theorize about truth would be an attitude to ontology with all the naturalness of the ostrich's famed response to challenges of a different kind.

I have urged that the most profitable approach to this explanatory task involves a conceptual retreat. We need to step backwards, asking not why truth matters but why it matters that ordinary speakers should think that it matters. We need to enquire what purpose could be served by our ordinary conviction that there is an end to enquiry. My account of the purpose of this conviction may not be the most appropriate one, but I think the stance is right. We need to examine the game rather than the goal. Concerning truth itself, in other words, I think there is more truth than usual in the pious sentiments of Grantland Rice.

For when the One Great Scorer comes

To write against your name,
He marks – not that you won or lost –
But how you played the game.

9.7 LANGUAGE AND MIND: EXPLANATORY PROSPECTS

I want to finish with some comments on the scope and potential of the general philosophical approach exemplified by this functional theory of truth. In particular I want to draw attention to its relevance to recent work on mind and language – to the philosophical end of cognitive science. Here its significance lies in the fact that it is at the same time naturalistic and non-reductive. I think it thus offers an important and undervalued alternative to recent work – a middle way for cognitive science, between the evils of non-naturalism on the one hand and the intractability of reductivism on the other. The middle way may not be the right way, but it needs to be recognized as one of the options.

The goal here is old enough: an understanding of the ordinary concepts of language and mind; of ‘truth’, ‘meaning’, ‘belief’, and so on. Increasing scientific interest in the mind has merely added weight to the view that this understanding should be sought in natural terms – in a manner compatible with the fact that we ourselves are part and product of the natural world. It is hard to object to the naturalistic constraint so stated.

So stated, however, it is neutral between reductive and explanatory accounts of the concepts of language and mind. Recent writers have generally taken it to call for the former. The task has been taken to be to give a naturalistic account of what we are talking *about* in talking of truth, belief, and the rest. Hence the philosophical industry devoted to the naturalistic analysis of these notions – an industry founded in large part on the view that the only alternative would be to give up these notions, to regard them as superseded by scientific psychology and linguistics. True, opinions differ as to the acceptability of that alternative. Not everyone agrees with Fodor that

if commonsense intentional psychology really were to collapse, that would be, beyond comparison, the greatest intellectual catastrophe in the history of our species; if we’re wrong about the mind, then that’s the wrongest we’ve ever been about anything We’ll be in deep, deep trouble if we have to give it up. (1987, p. *xii*)

For a less cataclysmic assessment see Stich (1983), chapter 11 or Churchland (1984), chapter 2.5, for example. All the same, there is a very widespread view that the fate of the folk notions rests on the success or otherwise of the reductive enterprise.

The functional approach suggests a challenge to this view. We have seen that it offers an account of the existence of the folk notions of truth and falsity, without seeking to analyse these notions in naturalistic terms. Its naturalism is that of evolutionary biology: it explains why creatures like us should have come to talk this way. But now why stop at truth? Why not look for parallel accounts of the rest of folk psychology and linguistics?

What would such a theory of intentional psychology be like? With one main difference, I think it might be very like that of Daniel Dennett, perhaps the most notable opponent of the view that the fate of folk psychology hangs on the possibility of its reduction to a physical basis. Dennett has long been concerned to save a place for the folk notions, given (as he accepts) that there is no place for them in neurophysiology. Briefly, his view is that beliefs and desires are ‘idealized fictions in an action-predicting, action-explaining calculus’ (1978, p. 30; see also 1978, chapter 1, and 1981). Clearly this has the makings of an explanatory theory in our sense. As Dennett emphasizes, the practical advantages of such a calculus in ordinary life provide the basis of an evolutionary explanation of the fact that humans have developed it.

The difference is that as the quote suggests, Dennett presents himself as an instrumentalist about folk psychology. On the present view, however, this reading of such a theory is unnecessary and unfounded. A sharp realist/instrumentalist distinction would call for a sharp factual/non-factual distinction; and that, we argued, is not to be had. This suggests that things are better than they seemed for an account of Dennett's kind. The explanatory perspective doesn't call for ontological apologies – or at least not such profuse ones, the difference between psychology and neuroscience being at most a matter of degree.

Stripped of its instrumentalism, Dennett's view thus seems very much in keeping with our general explanatory approach. However, it may then seem too good to be true. If it isn't instrumentalist, if it doesn't deny that there really are beliefs and desires, then how does it evade the strong intuition that there ought to be a physical account of what such things *are*? How in other words does it distinguish itself from reductivism?

The answer will be in two parts. The first step is to defuse the physicalist intuitions – to appeal to factors in virtue of which it seems implausible that ascriptions of psychological attitudes could be equated with ascriptions of physical states. (Not surprisingly the most comprehensive accounts of these factors are those of the opponents of realism about folk psychology; see, particularly, that of Stich 1983.) The second step is then to deny that the first calls for an anti-realist view of folk-psychological attitudes. The case here draws on the argument of this book, in a way that I hope will by now be familiar. Briefly, the anti-realist construal rests on an application of an untenable fact/non-fact distinction. All areas of discourse are potentially non-factual to some degree, but the factors responsible may vary from case to case. The peculiarities of folk-psychological discourse seem to distinguish it in this respect from talk about neurophysiology, with the result that we shouldn't expect a reduction. But only a philosophical misconception about truth could support the conclusion the anti-realist wants to draw from this: namely, that only the claims of neurophysiology are genuinely factual, truth-bearing, or whatever.

This approach to the problem of folk psychology in cognitive science thus applies the explanatory approach at two levels. It applies it first to the folk notions themselves, recommending an explanatory treatment in the same evolutionary vein as our functional theory of truth. The result, plausibly, is Dennett's account of the origins of intentionality. The second application then concerns the interpretation of this theory of psychology. The issue is the factual status of folk-psychological claims, and here an explanatory view of truth affords no basis for a sharp distinction. It was our disenchantment with the possibility of such a distinction that led us to the explanatory approach.

I don't want to sound too sanguine about the prospects for explanatory accounts of language and mind. Nor do I want to suggest that intuitions about ontological and theoretical priority can simply be dismissed out of hand. It would be much more in keeping with the present approach to try to explain these intuitions and the phenomena of theory change – to explain them in terms of our general conception of the function of factual discourse. Clearly usage is not immutable in the face of such considerations, and it would be nice to be able to explain why it changes when it does change.

For the moment, however, I want simply to urge some critical attention to the terms in which the current debate about the status of folk psychology is being phrased. The questions as to whether there are 'really' beliefs and desires, as to whether ordinary ascriptions of such attitudes have genuine truth values, are themselves phrased in folk-linguistic terms. Truth, fact and reality are folk notions. Their status is as much at issue as that of the folk-psychological notions – to which, evidently, they are closely related. Hence it is a mistake, I think, to expect or assume different kinds of answers in the psychological and linguistic cases. In particular it is a mistake to allow an unexamined philosophical conception of truth to govern one's interpretation of the issue about the status of folk psychology. There is a kind of consistency constraint here. If we're physicalists about belief then we ought to interpret the question of the truth status of belief ascriptions in terms of a physicalist notion of truth.

But if we take the explanatory approach in both cases then we deserve to be judged by the standards of that approach.

It thus turns out to have been important to treat truth first. To accept an explanatory account of truth is to constrain one's understanding of any other theory in the same general spirit. Explanatory theories have in the past been crippled by their failure to apply their insights to truth itself (think of the problems of projectivism, for example). The promise of the explanatory approach lies in considerable measure in its ability to explain the terms in which the traditional issues have been phrased.

REFERENCES

- Adams, E. 1965. The logic of conditionals. *Inquiry*, 8, 166–97.
- 1975. *The Logic of Conditionals*, Dordrecht: Reidel.
- Albury, R. and Slezak, P. (eds) 1988. *Computers Brains and Minds: Papers in Cognitive Science*, Dordrecht: Kluwer Academic Publishers.
- Austin, J. L. 1962. *How to Do Things with Words*, Oxford: Oxford University Press.
- Bach, K. 1975. Performatives are statements too. *Philosophical Studies*, 28, 229–36.
- Bar-Hillel, Y. (ed.) 1966. *Proceedings of the 1964 International Congress for Logic, Methodology and Philosophy of Science*, Amsterdam: North Holland.
- Blackburn, S. W. 1971. Moral realism. Casey (ed.) 1971, 101–24.
- 1984. *Spreading the Word*, Oxford: Oxford University Press.
- 1985. Supervenience revisited. Hacking (ed.) 1985, 47–67.
- Bouveresse, J. and Parret, H. (eds) 1981. *Meaning and Understanding*, Berlin: de Gruyter.
- Braithwaite, R. B. 1953. *Scientific Explanation*, Cambridge: Cambridge University Press.
- 1966. Why is it reasonable to base a betting rate upon an estimate of chance? Bar-Hillel (ed.) 1966, 263–73.
- Carlstrom, I. and Hill, C. 1978. Review of Ernest Adams, *The Logic of Conditionals* [Adams 1975] *Philosophy of Science*, 45, 155–8.
- Casey, J. (ed.) 1971. *Morality and Moral Reasoning*, London: Methuen.
- Churchland, P. 1984. *Matter and Consciousness*, Cambridge, Mass.: MIT Press.
- Davidson, D. 1968. On saying that. *Synthese* 19, 130–46. Reprinted in Davidson 1984, 93–108.
- 1974. Belief and the basis of meaning. *Synthese*, 27, 309–24. Reprinted in Davidson 1984, 141–54.
- 1984. *Inquiries into Truth and Interpretation*, Oxford: Oxford University Press.
- and Harman, G. (eds) 1972. *Semantics of Natural Language*, Dordrecht: Reidel.
- Dennett, D. 1978. *Brainstorms*, Cambridge, Mass.: MIT Press.

- 1981. Three kinds of intentional psychology. Healey (ed.) 1981, 37–61.
- Devitt, M. 1984. *Realism and Truth*, Oxford: Basil Blackwell.
- Dummett, M. A. E. 1958. Truth. *Proceedings of the Aristotelian Society*, 59, 141–62.
- 1973. *Frege: Philosophy of Language*, London: Duckworth.
- 1978. *Truth and Other Enigmas*, London: Duckworth.
- Ellis, B. 1980. Truth as a mode of evaluation. *Pacific Philosophical Quarterly*, 61, 85–99.
- Evans, G. and McDowell, J. (eds) 1976. *Truth and Meaning: Essays in Semantics*, Oxford: Oxford University Press.
- Field, H. 1972. Tarski's theory of truth. *Journal of Philosophy*, 69, 347–75. Reprinted in Platts (ed.) 1980, 83–110.
- 1978. Mental representation. *Erkenntnis*, 13, 9–16.
- Fine, A. 1986. Unnatural attitudes: realist and instrumentalist attachments to science. *Mind*, 95, 149–79.
- Fodor, J. 1985. Fodor's guide to mental representation. *Mind*, 94, 76–100.
- 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*, Cambridge, Mass.: MIT Press.
- Grover, D., Camp, J. and Belnap, N. 1975. A prosentential theory of truth. *Philosophical Studies*, 27, 73–125.
- Habermas, J. 1970. Towards a theory of communicative competence. *Inquiry*, 13, 360–75.
- Hacking, I. (ed.) 1985. *Exercises in Analysis: Essays by Students of Casimir Lewy*, Cambridge: Cambridge University Press.
- Hamblin, C. L. 1987. *Imperatives*, Oxford: Basil Blackwell.
- Harman, G. 1986. *Change in View: Principles of Reasoning*, Cambridge, Mass.: MIT Press.
- Healey, R. (ed.) 1981. *Reduction, Time and Reality*, Cambridge: Cambridge University Press.
- Hesse, M. 1978. Habermas' consensus theory of truth. *Proceeding of the Philosophy of Science Association 1978*, 2, 373–96.
- Hume, D. 1978. *Treatise of Human Nature*, ed. L. A. Selby-Bigge, 2nd edn with revisions by P. H. Nidditch, Oxford: Oxford University Press.
- Jackson, F. 1984. Two theories of indicative conditionals: reply to Brian Ellis. *Australasian Journal of Philosophy*, 62, 67–76.
- 1987. *Conditionals*, Oxford: Basil Blackwell.

- Kripke, S. 1982. *Wittgenstein on Rules and Private Language*, Oxford: Basil Blackwell.
- Leibniz, G. W. 1951. *Theodicy: Essays on the Goodness of God, the Freedom of Man and the Origin of Evil*, ed. A. Farrer, London: Routledge & Kegan Paul.
- Lewis, D. 1972. General semantics. Davidson and Harman (eds) 1972, 169–218. Originally published in *Synthese*, 22 (1970), 18–67.
- 1976. Probabilities of conditionals and conditional probabilities. *Philosophical Review*, 85, 297–315.
- Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*, Harmondsworth: Penguin.
- McDowell, J. 1978. Are moral requirements hypothetical imperatives? *Proceedings of the Aristotelian Society, Supplementary Volume*, 52, 13–29.
- 1981. Anti-realism and the epistemology of understanding. Bouveresse and Parret (eds) 1981, 225–48.
- McGinn, C. 1977. Semantics for nonindicative sentences. *Philosophical Studies*, 32, 301–11.
- 1982. The structure of content. Woodfield (ed.) 1982, 207–58.
- 1984. *Wittgenstein on Meaning*, Oxford: Basil Blackwell.
- Mellor, D. H. 1971. *The Matter of Chance*, Cambridge: Cambridge University Press.
- 1988. How much of the mind is a computer? forthcoming in Albury and Slezak (eds) 1988.
- Olshewsky, T. M. (ed.) 1969. *Problems in the Philosophy of Language*, New York: Holt, Rinehart & Winston.
- Palmer, F. R. 1986. *Mood and Modality*, Cambridge: Cambridge University Press.
- Parkinson, G. (ed.) 1982. *Marx and Marxisms*, Cambridge: Cambridge University Press.
- Pettit, P. 1982. Habermas on truth and justice. Parkinson (ed.) 1982, 207–28.
- Platts, M. (ed.) 1980. *Reference, Truth and Reality*, London: Routledge & Kegan Paul.
- Price, H. 1981. *The Problem of the Single Case*. Ph.D. thesis, Cambridge University.
- 1983a. Sense, assertion, Dummett and denial. *Mind*, 92, 174–88.
- 1983b. Does ‘Probably’ modify sense? *Australasian Journal of Philosophy*, 61, 396–408.
- 1986. Conditional credence. *Mind*, 95, 18–36.
- 1987. Truth and the nature of assertion. *Mind*, 96, 202–20.
- Putnam, H. 1978. *Meaning and the Moral Sciences*, London: Routledge & Kegan Paul.

- Quine, W. V. O. 1960. *Word and Object*, Cambridge, Mass.: MIT Press.
- 1970. *Philosophy of Logic*, Englewood Cliffs, NJ: Prentice-Hall.
- Ramsey, F. P. 1978. *Foundations*, ed. D. H. Mellor, London: Routledge & Kegan Paul.
- Rorty, R. 1982. *Consequences of Pragmatism: Essays 1972–1980*, Brighton: Harvester.
- Schiffer, S. R. 1972. *Meaning*, Oxford: Oxford University Press.
- Stalnaker, R. 1970. Probability and conditionals. *Philosophy of Science*, 37, 64–80.
- Stenius, E. 1967. Mood and language-game. *Synthese*, 17, 254–74.
- Stich, S. 1983. *From Folk Psychology to Cognitive Science: The Case Against Belief*, Cambridge, Mass.: MIT Press.
- Strawson, P. F. 1949. Truth. *Analysis*, 9. Reprinted in Olshewsky (ed.) 1969, 610–23.
- 1950. Truth. *Proceedings of the Aristotelian Society, Supplementary Volume*, 24, 129–56.
- van Straaten, Z. (ed.) 1980. *Philosophical Subjects: Essays Presented to P. F. Strawson*, Oxford: Oxford University Press.
- Wiggins, D. 1980. What would be a substantial theory of truth? van Straaten (ed.) 1980, 189–221.
- Williams, B. A. O. 1973. *Problems of the Self*, Cambridge: Cambridge University Press.
- Woodfield, A. (ed.) 1982. *Thought and Object*, Oxford: Oxford University Press.

INDEX of 1st Edition

Adams, E., 85, 99

Adams's hypothesis, 85, 99–112, 185

analytic approach *see* assertion, explanatory approach, marker view, non-factualism

anti-realism, 1

see also non-factualism

approval problem, 76, 82

see also confidence problem

argument

advantages over agreement *per se*, 156–7

behavioural utility of, 15, 134–5, 145, 151–5, 156–7

role of truth in, 15, 134–5, 153–5, 180–1

see also disagreement, dissent

assertion

belief-based analysis of, 14, 41–2, 63–4, 69, 75, 85–6, 99

primacy of, 7, 203

supposed characteristics of, 43–4, 57–9

truth-based analysis of, 13, 19–40, 93

see also non-factualism, non-indicatives

Austin, J. L., 50, 55

Bach, K., 52

Blackburn, S., 67, 76, 77–82, 127, 209, 210, 213–14

Braithwaite, R.B., 72

Carlstrom, I., 102, 107–11, 112–13

Chomsky, N., 193

Churchland, P., 216

commands *see* assertion, marker view, non-indicatives

conditionals, 99–114, 185–9

subjunctive, 188–9

see also Adams's hypothesis, factualism, non-factualism

confidence problem, 72, 74, 82

see also approval problem

Davidson, D., 8, 9, 33, 35, 36, 37

Dennett, D., 216–18

desire

- catalytic view of, 90
- in cognitive theory, 94–9
- Humean view of, 87–94

Devitt, M., 138–42

direction of fit, 181

disagreement

- evaluative, 173–80
- evaporative, 161, 195–8
- evidential, 158, 173
- inferential, 158, 173
- probabilistic, 159–60, 162–7
- substantial, 161
- see also* argument, dissent

dissent

- and negation, 149
- triggers of, 146–9
- see also* argument, disagreement, truth

Dummett, M., 21, 29, 30, 31–2, 43–5, 57, 131, 193

Ellis, B., 155–7

emotivism, 1, 6

- see also* moral judgement, non-factualism

ethics

- see* moral judgement, non-factualism

Evans, G., 8, 36

explanatory approach, 14–16, 118–19

- contrasted to analytic approach, 4, 118–19
- evolutionary character of, 15, 125–30, 216
- to folk psychology, 215–18
- to linguistic distinctions, 119–20, 122–30
- naturalistic character of, 4, 15, 215
- prospects for, 215–18
- to social class, 120–2
- see also* non-factualism, truth

explicit performatives, 44, 49, 50–3

factualism

- about conditionals, 99–114
- about meaning, 11
- about moral judgement, 64–5, 76–7
- about probability, 63–5, 72–5

radical, 75, 77, 82, 84, 88, 99
 universal, 43, 44, 49, 53, 56, 59

factuality *see* assertion, explanatory approach, factualism, non-factualism

fatalism, 103–7

Field, H., 137–44

Fine, A., 208–9, 212–14

Fodor, J., 94–5, 98, 216

folk psychology, 16, 94, 216–18

Fregean general semantics

- affinity with marker view, 12
- conflict with non-factualism, 3, 7–11, 20, 202–4
- primacy of assertion in, 7–10
- sense /force distinction in, 9, 34–5
- treatment of non-indicatives, 8–9, 12
- use of truth in, 7– 8

functional realism, 211

- naturalism of, 215–16
- not idealism, 210–11
- not instrumentalism, 211–13
- not projectivism, 210
- not quietism, 214
- see also* explanatory approach

Habermas, J., 157

Hamblin, C., 53

Harman, G., 163–5

Hesse, M., 157

Hill, C. *see* Carlstrom, I.

Hume, D., 6–7, 67, 76, 86, 87–94

IDA (imagined dislike argument), 88–9, 90

imperatives *see* assertion, marker view, non-indicatives

indexicality

- and conditionals, 107
- and probability, 106–7, 167–9

instrumentalism, 211–13, 215–16

interrogatives *see* assertion, marker view, non-indicatives

Jackson, F., 112–13, 189

Kripke, S., 11

language of thought, 95–7, 128

Leibniz, G., 104

Lewis, D., 44, 50, 52, 102–7, 111–13

McDowell, J., 8, 33–6, 38, 40, 91

McGinn, C., 51, 98, 138, 140–4, 193

Mackie, J. L., 65–6

marker view

affinity with Fregean semantics, 12

analytic presuppositions of, 12

characterization of, 11

recognitional requirement, 28 sceptical argument against, 43–59

truth-based version of, 19, 22, 28

meaning *see* Fregean general semantics, radical interpretation, truth

Mellor, D. H., 73, 96, 104

moral judgement

disagreement about, 173–80

functional role of, 16, 76–7

regulative role of, 16, 179

supervenience of, 77–82

see also disagreement, factualism, non-factualism

negation

and dissent, 149

as a mood indicator, 23

NOA (natural ontological attitude), 213–14

non-cognitivism, 1

see also non-factualism

non-factualism

about aesthetics, 173–80

about causality, 196

about conditionals, 85, 99–114, 185–9

about everything, 191–5

about folk psychology, 217–18

about meaning, 11

about modality, 196

- about morality, 76–7, 173–4
- about presupposition, 190–1
- about probability, 63–4, 159–60, 162–7, 177–8
- about secondary qualities, 196
- about theory, 197
- about truth, 180–1, 208–9
- analytic version of, 2
- characterization of, 1–2
- conflict with Fregean semantics, 3, 7–11, 20
- explanatory approach to, 136, 161–2, 172, 184
- history of, 6–7
- sources of, 195–8
- and universal factualism, 61–2

non-indicatives

- deception by means of, 44, 45–7
- indicative paraphrases of, 49, 54–6
- non-applicability of truth to, 135
- significance of, 199–201
- see also* marker view

normativity, 131–4

- see also* moral judgement, truth

Northern Line, 123

Palmer, F., 11–12

paratactic analysis, 35

Pettit, P., 157

presupposition, 190–1

Price, H., 37, 63, 101

probability

- disagreement about, 159–60, 162–7, 177–8
- functional role of, 70–2, 165–6
- and non-final judgement, 165
- and rationality, 72–4
- supervenience of, 82–4
- see also* disagreement, factualism, non-factualism

projectivism, 67–8, 79, 181–3, 208–9, 210

- and idealism, 68–9, 79, 209–10

Putnam, H., 138, 144

quasi-realism, 210, 213

queerness, argument from, 65–6

questions *see* assertion, marker view, non-indicatives

quietism, 14, 133, 213–14

Quine, W. V. O., 26, 193, 194–5

rabbit, 163

radical interpretation, 32–3, 37, 39, 202–3
empirical grounds for, 37–8

Ramsey, F. P., 26, 85, 139

rat *see* rabbit

realism *see* factualism, functional realism, non-factualism

relativism, 173–80, 183

Rice, G., 215

Rorty, R., 133, 214

RTM (representational theory of mind), 95–8

SBP (same boat property), 152–3, 159, 160, 167, 177, 178, 182–3, 185

Schiffer, S., 51–3

SEA (simple ethical ascription), 64
see also factualism, moral judgement, non-factualism

SPA (simple probability ascription), 63
see also factualism, non-factualism, probability

statement of fact *see* assertion, explanatory approach, factualism, non-factualism

statementhood, 2
see also assertion, explanatory approach, factualism, non-factualism

Stenius, E., 22–4, 28, 29, 30

Stich, S., 96–7, 216, 217

STM (syntactic theory of mind), 96–7

Strawson, P. F., 26, 31, 39, 130, 137

subjunctives *see* conditionals

supervenience *see* moral judgement, probability

Tarskian truth theory, 32, 137
see also Fregean general semantics, truth

truth

amen theory of, 26, 39, 130–1, 137
disquotational view of, 26, 137–44, 203
equivalence property of, 26, 39, 205–6
and evaporative argument, 170–1
explanation by means of, 137–4
explanatory approach to, 4, 15–16, 119–20, 130–1, 137, 205–7
functional role of, 4, 15–16, 130–1, 132, 135–6, 145, 205–7
marks of, 20, 40–1, 202, 206–8
normativity of, 131, 203, 205–6
redundancy view of, 26
role in argument of, 134–5, 153–5, 180–1
substantial theory of, 21
truth-like nature of, 205–7
see also explanatory approach, non-factualism

universal factualism *see* factualism, universal

Wiggins, D., 19–20, 31, 32, 40–1, 206–7

Williams, B., 22, 26, 31

Wittgenstein, L., 11, 29, 193